Il contributo dell'accessibilità per sordi alla resocontazione

By Carlo Eugeni & Silvia Velardi (University of Leeds, UK; IULM University, Italy)

Abstract & Keywords

English:

In our age of artificial intelligence, technology pervades every single human activity and human-machine interaction is now daily practice. In the area of diamesic translation, this has led to important distinctions in workflows, greatly reducing human intervention in many fields including institutional ones. However, and while it is clear that technology is here to stay, the stages that marked this journey often lie forgotten in end-of-project reports, published online and then removed at the end of the project itself. In an attempt to reconstruct the path that led to the pervasive role of technology in the production phases of translation, from speech to writing, this article suggests a quadripartition of human-machine interaction and illustrates two important Italian research projects, which implemented speech recognition technology in the production of pre-recorded, real-time subtitles. On the basis of these experiments, this article highlights the advancement of automation in the field of accessibility on one hand, which guarantees speed and accuracy=, and on the other underlines the contribution that accessibility has made to society in terms of flexibility and of the transparency of democratic and legislative processes.

Italian:

Nell'era dell'intelligenza artificiale, la tecnologia pervade ormai ogni singola attività umana e l'interazione uomo-macchina è prassi quotidiana. Nell'ambito della traduzione diamesica, essa ha portato a importanti distinzioni dei flussi di lavoro, riducendo notevolmente l'intervento umano in molti ambiti, compreso quello istituzionale. Ma se è chiaro che la tecnologia è qui per restare, le tappe che hanno segnato questo percorso sono spesso dimenticate in relazioni di fine progetto pubblicate online e poi rimosse con la fine del progetto stesso. In un tentativo di ricostruzione del percorso che ha portato la tecnologia a pervadere le fasi produttive della traduzione dal parlato allo scritto nella stessa lingua e in lingua straniera, questo articolo propone una quadripartizione dell'interazione uomo-macchina e illustra due importanti progetti di ricerca italiani, che hanno implementato la tecnologia del riconoscimento del parlato nella produzione di sottotitoli preregistrati e in tempo reale. Sulla base di questi esperimenti volti all'accessibilità, si è potuto notare l'avanzata dell'automazione nell'ambito dell'accessibilità, che riesce anche a garantire rapidità e accuratezza da un lato e il contributo che l'accessibilità ha dato alla società in termini di flessibilità e trasparenza dei processi democratici e legislativi dall'altro.

Keywords: interazione uomo-macchina, resocontazione parlamentare, sottotitoli per sordi, respeaking, traduzione intralinguistica, human-machine interaction, parliamentary reporting, subtitling for the deaf, intralinguistic translation

1. Introduzione

Nell'era dell'intelligenza artificiale, la tecnologia pervade ormai ogni singola attività umana. Hewett *et al.* definiscono questa interdipendenza tra l'uomo e la tecnologia interazione uomo-macchina e, più specificamente, l'implementazione di sistemi di calcolo interattivi ad uso umano (Hewett *et al.* 1992). Nell'ambito della traduzione basti pensare ai *Computer-Aided Translation tools* e alla loro evoluzione in *Human-Aided Translation tools* e *Fully-Automated Translation tools* (Eugeni e Gambier 2023), questi ultimi all'origine del *Machine Translation Post Editing*. Nell'elaborazione del linguaggio naturale, il riconoscimento automatico del parlato ha ridotto notevolmente l'intervento umano negli ambiti del *Media Indexing*, della resocontazione, della sottotitolazione e perfino dell'interpretariato (Eugeni 2019, Pagano 2020, Spinolo e Amato 2020, Romero-Fresco 2023) dove sono stati introdotti i concetti ibridi di *written interpreting* (Eugeni e Bernabé 2021) e *AI-enhanced computer-assisted interpreting* (Fantinuoli 2023). Nonostante la qualità del lavoro della macchina rispetto a quella del lavoro umano sia ancora molto dibattuta, il verdetto è chiaro: la tecnologia è qui per restare e il mercato non può che adeguarsi ad esso. Ma come si è arrivati a questo risultato? Quali sono state le tappe che hanno segnato questo percorso? Qual è stato l'impatto della tecnologia sulla vita di tutti i giorni? Come potrà ulteriormente evolvere l'interazione uomo-macchina in settori affini alla traduzione come la resocontazione parlamentare e la verbalizzazione del processo penale? In particolare, l'accessibilità, spesso beneficiaria della tecnologia e dell'adattamento di testi pensati per un pubblico più ampio, può contribuire a "restituire il favore" alla società, permettendole di capitalizzare sui successi dei progetti per l'accessibilità? Il presente articolo si propone di rispondere a queste domande, elaborando brevemente una ripartizione più dettagliata dell'interazione uomo-macchina (§2). Successivamente, verranno illustrati i tre progetti in cui si sono messi a confronto la produzione del lavoro umano con quello della macchina in due ambiti distinti ma affini, la sottotitolazione per sordi preregistrata (§3) e in tempo reale (§4). Infine, verrà illustrato il contributo in questi ambiti alla resocontazione parlamentare e la verbalizzazione del processo penale, tramite una piattaforma multifunzionale che garantisca accuratezza, trasparenza e accessibilità (§5).

2. Interazione Uomo-Macchina nella traduzione diamesica

Come appena introdotto, l'intervento della tecnologia gradualmente riduce l'intervento umano. Nell'ambito della traduzione diamesica, o traduzione intra-linguistica dal parlato allo scritto sotto forma, per esempio, di sottotitoli per sordi, resoconti parlamentari o trascrizione in tempo reale, "technological investments have reduced the place of humans to such an extent that the profession is impossible without the former, but not without the latter" (Eugeni, 2020: 22). Se ignoriamo il contributo della tecnologia nello sviluppo delle tecniche utilizzate (respeaking e velotipia) per produrre le forme di traduzione diamesica analizzate e ci concentriamo solo sulle fasi di produzione di una traduzione (nel caso dei sottotitoli le fasi sono trascrizione, eventualmente traduzione, correzione e formattazione), risulta chiara la ripartizione dei flussi di lavoro secondo il grado di interazione uomo-macchina presente, indipendentemente dalla tecnologia usata per produrre i testi di arrivo (Eugeni 2019), dalla lingua (Pagano 2020) e dal modo traduttivo (Fantinuoli 2023):

Human-Only, in cui le fasi fondamentali del processo traduttivo sono assegnati nella loro interezza a uno o più professionisti. Nella traduzione diamesica, i professionisti trascrivono, eventualmente traducono e correggono, limitando l'uso della tecnologia alla tecnica per produrre (respeaking, stenotipia, tastiera QWERTY) e visualizzare (software di videoscrittura) i sottotitoli;

Computer-Aided, in cui il processo traduttivo è gestito dal professionista, che si avvale della traduzione per una minoranza delle fasi principali del processo traduttivo. Nella traduzione diamesica, il professionista trascrive il testo e usa la tecnologia per una o più fasi più facilmente automatizzabili come la correzione, la formattazione o, più spesso la traduzione;

Human-Aided, in cui il professionista vede il suo contribuito diminuire rispetto a quello della macchina nel processo traduttivo. Nella traduzione diamesica, il *live editor*/*scopist* si limita a correggere la trascrizione ed eventualmente la traduzione prodotte della macchina;

Computer-Only, in cui il processo traduttivo è interamente assegnato a uno o più software. Nella traduzione diamesica, un software di riconoscimento automatico del parlato trascrive immediatamente il testo di partenza e un altro eventualmente lo traduce, lo corregge e lo manda in onda, senza che un professionista lo corregga o adegui alle esigenze degli utenti.

3. Interazione Uomo-Macchina nella sottotitolazione preregistrata

Nell'ambito del progetto CLAST**[1]**, sono stati messi a confronto cinque video in lingua italiana, di cui due sottotitolati in italiano per sordi**[2]** da professionisti e due sottotitolati automaticamente. Anche un quinto video è stato sottotitolato dall'inglese in italiano in modalità automatica. Obiettivo della sperimentazione era la comprensione dei sottotitoli automatici intralinguistici e interlinguistici da parte degli utenti sordi rispetto a quella degli udenti. I video sono stati inizialmente valutati con la tassonomia IRA**[3]** e sono stati sottoposti a un gruppo di 16 sordi segnanti e a un gruppo di 16 udenti di pari livello socio-culturale. Dopo ogni video, a ciascun partecipante è stato chiesto di rispondere a test di comprensione e a un *think-aloud protocol*, per poterne valutare la ricezione e la percezione.

3.1. Materiali

Al fine di valutare la qualità dei sottotitoli prodotti automaticamente tramite la tecnologia sviluppata durante il progetto CLAST, sono stati selezionati cinque video di una durata compresa tra i 2 e i 3 minuti:

un documentario in italiano divulgativo sulla città di Siena nel Rinascimento;

un documentario in italiano divulgativo sulla città di Firenze nel Rinascimento;

un notiziario in italiano tecnico sullo sviluppo economico nell'Italia contemporanea;

un notiziario in italiano tecnico sull'e-commerce;

un notiziario in inglese tecnico sui fondamenti dell'economia.

La scelta dei tipi testuali (documentario divulgativo e notiziario economico) è stata dettata dalla necessità di testare i due profili sviluppati nel progetto CLAST: il modello divulgativo e il modello giornalistico. Per poter valutare l'efficacia della sottotitolazione automatica rispetto a quella manuale, dei due video a carattere divulgativo, il primo è stato sottotitolato automaticamente (*Computer Made*) dall'italiano in italiano (video 1) e l'altro manualmente (*Human-Only*), sempre dall'italiano in italiano (video 2). La stessa operazione è stata fatta con i primi due notiziari di ambito economico (video 3 e 4). Il quinto video è stato sottotitolato automaticamente (*Computer-Only*) dall'inglese in italiano (video 5). Non sono state usate forme di interazione uomo-macchina intermedie (*Computer-Aided* o *Human-Aided*) per la produzione dei sottotitoli.

Per quanto riguarda la qualità della sottotitolazione, si rende evidente che, applicando la tassonomia concettuale IRA, la qualità del video 1 è pari all'88 per cento, mentre quella del video 2 è pari al 100 per cento. Una differenza inferiore tra i due video è stata riscontrata tra il video 3 e il video 4. Il video 3 ha una qualità del 95 per cento contro il 100 per cento della qualità del video 4. La qualità del video 5 è invece dell'85 per cento. Pur essendo di una qualità inferiore al 95 per cento considerata come accettabile nell'elaborazione della tassonomia IRA, i video 1 e 5 sono comunque stati testati.

I video sono stati sottoposti a tutti i partecipanti con l'audio a un volume udibile a tutti gli udenti e con i sottotitoli. Ogni diverso trattamento tra i due gruppi è stato considerato inutile o fuorviante e pertanto scartato. Va sottolineato che la componente visiva dei primi due video ha una valenza ancillare rispetto alla componente acustica, perché la comprensione delle immagini dipende sempre dalla corretta percezione della componente acustica verbale. La stessa caratteristica è stata riscontrata nel video sottotitolato interlinguisticamente (video 5), per quanto quest'ultimo contenga informazioni di tipo scritto-verbale. La componente visiva dei video 3 e 4, invece, svolge un ruolo diverso, perché auto-esplicativa: fornisce informazioni comprensibili senza fruire la relativa componente acustica verbale.

3.2. Partecipanti

Target principale del progetto CLAST sono stati gli utenti sordi. Per questo motivo si è scelto di collaborare con l'Ente Nazionale Sordi di Trento. Oltre che per motivazioni legate alla vicinanza dalla sede della ricerca, la scelta di testare i sottotitoli su una popolazione segnante è stata dettata dalla necessità di testare l'efficacia dei sottotitoli sul tipo di pubblico che maggiormente si allontana dallo standard linguistico dell'utente medio dei video in questione. In linea con sperimentazioni simili svolte a livello nazionale**[4]** e internazionale**[5]**, la scelta del gruppo target è ricaduta sui sordi segnanti. Il *focus group* era composto da 17 sordi segnanti gravi o profondi, di cui 9 donne e 8 uomini, impiegati, di età compresa tra i 30 e i 70 anni e con scolarizzazione tra le scuole medie e le scuole superiori. Dal gruppo iniziale, sono stati esclusi i risultati di un volontario che aveva scelto più opzioni di risposta a quasi tutti i quesiti, alterando così l'esito del test.

Nonostante tutti si fossero dichiarati bilingui, il *focus group* era composto da soci dell'ENS – sordi prelinguali o perilinguali – la cui lingua materna è la lingua dei segni italiana (LIS). Il *control group* era composto da altrettante persone (16), scelte sulla base del medesimo profilo socio-culturale dei 16 sordi rimanenti del *focus group*: 9 donne e 7 uomini impiegati, di età compresa tra i 30 e i 70 anni e scolarizzati fino alle superiori.

A latere, preme ricordare che, da uno studio condotto dall'ente regolatore dei *broadcaster* britannici Ofcom, che gli utenti del servizio di sottotitolazione dei teletext britannici solo per il 20 per cento sono composti da ipoacusici o cofotici. Il restante 80 per cento è rappresentato principalmente da stranieri (immigrati, studenti, amici di britannici temporaneamente nel Regno Unito), abitanti che lavorano in ambienti rumorosi o di cittadini britannici emigrati all'estero - o dei loro discendenti - che vogliono mantenere un contatto con la lingua materna scritta e parlata (Ofcom 2006, 2013). Pur con differenze di ordine sociale e demografico, lo studio britannico non può essere ignorato nella valutazione della qualità dei sottotitoli automatici oggetto del presente studio di ricerca.

3.3. Questionari

Sei questionari sono stati sottoposti ai partecipanti, a ognuno dei due gruppi che si sono sottoposti al test, in linea con quanto già definito nel progetto DTV4All: il primo questionario è stato proposto ai partecipanti prima della presentazione del lavoro e richiedeva dati

riguardanti le generalità principali dei partecipanti (genere, età, scolarizzazione, impiego e metodo di comunicazione), il loro rapporto con la lingua scritta e la lettura (siti web, sottotitoli TV, *social network*) e le loro preferenze in termini di tipi testuali (film e fiction, documentari divulgativi, notiziari economici, programmi sportivi).

Gli altri cinque questionari sono stati sottoposti ai partecipanti dopo ogni video visionato ed erano composti da sei domande ciascuno, incentrate sulle due grandi componenti semiotiche del testo: acustica e visiva. Le tre domande di ogni questionario incentrate sulla componente acustica riguardavano i macro-argomenti trattati nel video, mentre quelle incentrate sulla componente visiva riguardavano immagini di supporto al testo ma utili alla sua comprensione o immagini auto-esplicative come luoghi descritti, dati (parole e numeri) sovrimpressi o grafici.

Alla fine della compilazione dei questionari si è proceduto alla somministrazione di un *Think-Aloud Protocol* individuale e poi a un confronto collettivo sulle criticità dell'esperimento. I risultati sono stati raccolti e analizzati sia globalmente, sia nel dettaglio.

3.4. Risultati

Non essendo state rilevate differenze sostanziali dovute al genere, all'età o al livello di scolarizzazione, si procederà ora ad analizzare i dati secondo le variabili più interessanti. Dopo un'analisi generale dei risultati ottenuti dai sordi e dagli udenti, si entrerà nel dettaglio dei singoli video in termini assoluti (percentuale di risposte esatte) e relativi (percentuale di risposte esatte rispetto alla qualità dei sottotitoli). Poi si prenderanno in considerazione le macro-variabili più interessanti, a partire da quelle di maggiore interesse per il presente studio, come la modalità di produzione del sottotitolo (automatico vs. manuale). Successivamente si procederà all'osservazione dei dati riguardanti il tipo di programma (documentario divulgativo vs. notiziario economico). Questo ci permetterà di capire se e in che misura il tipo testuale dei video influenza la percezione e la ricezione di un video. In seguito, si analizzeranno i due tipi di domande contenute nei questionari: quelle relative alla comprensione delle informazioni veicolate dalla componente acustica e quelle relative alla comprensione delle informazioni veicolate dalla componente visiva. Questo permetterà di comprendere se, indipendentemente dalla qualità, i sottotitoli automatici comunque agevolano la comprensione delle immagini oppure no. Infine, si analizzeranno i dati relativi al video sottotitolato automaticamente dall'inglese in italiano.

3.4.1. Sordi vs Udenti

In generale, le risposte degli udenti sono state migliori di quelle date dai sordi, con una media di 4,2 risposte corrette su 6 per i primi *vs* 3,8 per i secondi. Questo dimostra che la comprensione dei video da parte di chi riesce contemporaneamente a percepire la componente acustica e quella visiva su due canali distinti è superiore a quella di chi deve utilizzare un unico canale (quello visivo) per percepire informazioni di tipo sia visivo che acustico (cfr. Paivio 1986). Questa operazione comporta una minore possibilità di concentrarsi anche sulle immagini, non riuscendo quindi a percepire correttamente l'interrelazione tra parlato e immagini, contrariamente al pubblico udente. Si consideri inoltre che i partecipanti al *focus group* non sono madrelingua italiani e i messaggi del video sono tutti veicolati sul canale visivo, che non è comunemente deputato alla comunicazione orale. Per comprendere, quindi, se i sottotitoli automatici hanno influito negativamente su questo risultato, occorre analizzare l'esito delle risposte alle domande di tipo testuale e di tipo visivo.

Utenti	Video 1	Video 2	Video 3	Video 4	Video 5	Media
Sordi	3,2/6	3,9/6	3,9/6	4,6/6	3,3/6	3,8/6
Udenti	4,2/6	4,7/6	4,6/6	4,2/6	3,4/6	4,2/6

Tabella 1: Risposte corrette ai questionari suddivise per singolo video

Dai dati esposti nella Tabella 1 emerge quanto espresso precedentemente sulla comprensione generale degli udenti rispetto ai sordi. Sistematicamente, la comprensione degli udenti dei diversi video è maggiore (video 1, 2, 3, 5) rispetto a quella dei sordi. Da una più approfondita analisi, tale divario risulta maggiore nel caso dei video sottotitolati automaticamente (+ 1 nel video 1 e + 0,7 nel video 3), mentre in uno dei due video sottotitolati manualmente il divario è - 0,4 nel video 4, anche se il video 2 non conferma tale costante. Considerata questa eccezione, è necessario controllare i dati più nel dettaglio per comprendere se l'ipotesi che ne deriva – i sottotitoli manuali sono più adeguati alla comprensione rispetto a quelli automatici – è fondata o il dato deriva da altre ragioni (cfr. § 3.4.3).

Un altro dato che salta all'occhio riguarda la maggiore comprensione da parte dei sordi del video 4, ossia il notiziario economico sottotitolato manualmente. Come si vedrà dall'analisi delle risposte alle domande di tipo testuale e le domande di tipo visivo (§ 3.4.5), questo dato è la risultante della maggiore comprensione da parte dei sordi della componente visiva. Visto che il notiziario in questione è il più ricco di dati a video, questo ci porta a tre prime conclusioni:

> anche gli udenti soffrono di sovraccarico cognitivo qualora la componente video veicoli non solo informazioni secondarie (immagini che accompagnano il testo pronunciato oralmente), ma anche primarie (in questo caso i dati sull'e-commerce);
>
> di contro, la maggiore abilità a usare la vista per comprendere le informazioni offre ai sordi un vantaggio sugli udenti nel comprendere testi con prevalenza di informazioni visive;
>
> oltre a permettere una maggiore comprensione della componente scritta rispetto ai sottotitoli prodotti automaticamente, i sottotitoli prodotti manualmente permettono di lasciare il tempo agli utenti di ottenere informazioni anche dalla componente visiva.

Altro dato degno di nota che sarà analizzato successivamente (§ 3.4.6) riguarda i dati dei sordi riferiti al video in inglese sottotitolato automaticamente, che registra dei risultati vicini a quelli degli udenti. Questo dato è interessante perché sordi e segnanti sono, in questo caso, posti su un livello di esposizione al video particolarmente più simile rispetto alla ricezione degli altri quattro video, dato che la loro conoscenza dell'inglese parlato è più limitata.

3.4.2. Comprensione assoluta vs. Comprensione relativa

Un altro elemento importante da analizzare riguarda la comprensione relativa. I dati sopra riportati, infatti, riguardano la comprensione assoluta, vale a dire la comprensione del singolo video indipendentemente dalla maggiore o minore qualità dei sottotitoli. Per comprendere quanto la qualità del sottotitolo influisca sulla comprensione generale del prodotto, però, è più interessante conoscere la qualità relativa del sottotitolo. Per qualità relativa si intende quanto i sottotitoli, in base alla loro accuratezza misurata con IRA, riescono a veicolare i messaggi ai quali si riferiscono. La comprensione relativa (CR), ossia il numero relativo di risposte esatte alle domande

poste, viene ricavato moltiplicando la comprensione assoluta (CA) per il quoziente qualitativo dei sottotitoli (qqS), che si ottiene dividendo 100 per la qualità dei sottotitoli calcolata da IRA (qIRA), secondo la formula CR = CA * qqS, dove qqS = 100/qIRA, quindi CR = CA * 100/qIRA.

Da questo emerge che la CR è inversamente proporzionale alla qualità dei sottotitoli. Vale a dire che, a parità di CA, minore è la qualità dei sottotitoli e maggiore sarà la CR. Di fatto, si tratta dunque di comprendere quanto i rispondenti abbiano davvero capito del video, nonostante eventuali errori nei sottotitoli. Saranno considerati soltanto i sottotitoli visionati dai sordi, poiché gli udenti, al netto di irrilevanti casi, hanno potuto percepire acusticamente la componente testuale. Calcolare anche per gli udenti la comprensione relativa significherebbe falsare la realtà del processo cognitivo degli udenti, pur riconoscendo l'influenza della presenza dei sottotitoli nel processo di visualizzazione dei video da parte degli udenti stessi (cfr. Romero-Fresco 2105 e Bianchi *et al.* 2020). Da segnalare è anche la scelta delle componenti da considerare. Analizzare la sola componente testuale vorrebbe dire riconoscere ai sottotitoli un'influenza sulla sola componente testuale. Studi scientifici hanno ampiamente dimostrato che la qualità e la densità dei sottotitoli hanno un impatto decisivo anche sulla visione delle immagini (*ibidem*). In particolare più è denso il sottotitolo, minore sarà il tempo dedicato alla visione delle immagini da parte degli utenti. Di converso, maggiore sarà la qualità dei sottotitoli maggiore saranno il tempo e l'attenzione dedicati alla visione delle immagini.

Comprensione	Video 1	Video 2	Video 3	Video 4	Video 5
Assoluta	3,2/6	3,9/6	3,9/6	4,6/6	3,3/6
Relativa	3,6/6	3,9/6	4,1/6	4,6/6	3,9/6

Tabella 2: Comprensione assoluta e relativa dei sordi suddivisa per singolo video

I dati sulla CR illustrati in Tabella 2 mostrano una prima evidenza: al relativizzare della comprensione, il divario tra la sottotitolazione automatica e quella manuale si riduce notevolmente. Nel caso dei primi due video, esso si riduce di oltre la metà (da 0,7 a 0,3), mentre nel caso dei secondi due rimane pressoché invariato, per via della maggiore qualità dei sottotitoli automatici del video 3. Sorprende altresì il balzo qualitativo del video 5, che arriva a rivaleggiare con il video 2, sottotitolato manualmente. Questo non significa che i sottotitoli interlinguistici automatici siano efficienti quanto quelli intra-linguistici manuali, tuttavia dimostra il potenziale di sottotitoli prodotti dalla doppia automazione della trascrizione e della traduzione. Con l'evolversi della ricerca in questo ambito, l'impiego efficiente di sottotitoli automatici, intra-linguistici o interlinguistici, è sempre più a portata di mano.

3.4.3. Sottotitoli automatici vs. Sottotitoli manuali

I sottotitoli prodotti manualmente (video 2 e 4) sono più efficienti di quelli prodotti automaticamente (video 1, 3 e 5) nel veicolare messaggi di tipo sia acustico che visivo. Questo risulta non solo dai dati assoluti (§ 3.4.1), ma anche da quelli relativi (§ 3.4.2) e da quelli scorporati per componente semiotica (§ 3.4.4). Nella Tabella 3, i dati riferiti ai sottotitoli automatici comprendono solamente la media delle risposte corrette riguardanti i video sottotitolati dall'italiano in italiano (video 1 e 3), perché il video 5 introduce la variabile della traduzione interlinguistica, che non è presente nei sottotitoli manuali.

Comprensione	Sottotitoli Automatici	Sottotitoli Manuali
Assoluta	3,5/6	4,2/6
Relativa	3,8/6	4,2/6
Testuale	3,3/6	4,4/6
Visiva	3,8/6	4/6

Tabella 3: Comprensione dei sordi suddivisa per tipo di sottotitoli

Dai dati della Tabella 3 emergono diverse realtà degne di nota. Innanzitutto, in termini assoluti, risulta chiaro che i sottotitoli prodotti manualmente garantiscono una maggiore comprensibilità dei messaggi veicolati dai video usati nella sperimentazione (+ 0,7 domande), pari a una differenza dell'11,7 per cento Tuttavia, visto che la qualità del riconoscimento automatico del parlato è in progressivo miglioramento e che la qualità della componente testuale impatta anche sulla fruibilità della componente visiva, quindi del video nel suo insieme, si è ritenuto interessante rapportare la comprensione dei sordi alla quantità effettiva di informazioni ricevute dal video sottotitolato automaticamente. In questo caso, le informazioni effettivamente veicolate dai video sottotitolati automaticamente non corrispondono al 100 per cento delle informazioni veicolate nei video sottotitolati manualmente, ma a una percentuale di volta in volta variabile. Da questo calcolo della comprensione relativa dei video sottotitolati automaticamente, emerge che il divario con i video sottotitolati manualmente, prima netto, si rivela molto inferiore (0,4 risposte, pari al 6,7 per cento.

Un altro dato interessante, che si può leggere da due prospettive opposte, riguarda la componente testuale. Le risposte alle domande sulla componente testuale sono quelle che mostrano il maggiore divario tra i video sottotitolati automaticamente (3,3 risposte corrette su 6) e quelli sottotitolati manualmente (4,4 su 6):

> da un lato, 1,1 risposte su 6 dimostrano che l'area di intervento diretto della sottotitolazione risente negativamente dell'automazione rispetto a un trattamento professionale. Questo può sembrare banale, ma è sempre interessante quantificarne la portata, che in questo caso è del 18,3 per cento.
>
> dall'altro lato, le stesse 1,1 risposte su 6 di differenza dicono che i sottotitoli automatici non impattano negativamente sulla ricezione del video che traducono, perché il divario in termini relativi, ma anche assoluti, è inferiore.

La seconda ipotesi è confermata dai dati riguardanti la comprensione della componente visiva dei video. In termini teorici, una buona sottotitolazione permette agli spettatori di comprendere la componente acustica del video, lasciando loro il tempo di concentrarsi sulle immagini. Di conseguenza, sottotitoli qualitativamente inferiori non solo veicolerebbero meno informazioni, ma comprometterebbero la visione del video nel suo insieme, rallentando la lettura dei sottotitoli e lasciando allo spettatore meno tempo per guardare le immagini e cogliere le informazioni che veicolano. Di fatto, invece, emerge che la relativa bassa qualità dei sottotitoli automatici non comporta un perdita di informazioni altrettanto importante.

3.4.4. Componente acustica vs. Componente visiva

A completamento dell'analisi precedente sui sottotitoli automatici rispetto ai sottotitoli manuali (§3.3), è forse utile controllare non solo i dati riferiti ai singoli video, ma anche i dati riferiti agli utenti udenti. Questa necessità si impone per poter corroborare o confutare le ipotesi fin qui esposte e riferite alle due componenti in questione. Una maggiore presenza delle domande corrette in una delle due categorie potrebbe, infatti, dipendere non solo dalla qualità dei sottotitoli, ma anche da una intrinseca maggiore o minore trasparenza delle informazioni oggetto dei singoli questionari. Neanche in questa analisi sono stati considerati i dati relativi al video 5 (cfr. § 3.4.6), perché alterano il rapporto con la fruibilità del video da parte degli udenti.

Utente	Componente	Video 1	Video 2	Video 3	Video 4
Sordi	Acustica	3/6	4,5/6	3,6/6	4,4/6
	Visiva	3,4/6	3,3/6	4,2/6	4,8/6
Udenti	Acustica	4,3/6	4,7/6	4,5/6	4,4/6
	Visiva	4,1/6	4,7/6	4,7/6	4/6

Tabella 4: Comprensione suddivisa per tipo di componente semiotica e tipo di utente

Dalla Tabella 4 emerge un dato che parrebbe contraddire quanto esposto: non esiste correlazione alcuna tra qualità del sottotitolo e comprensione testuale. Infatti, i dati riferiti alla comprensione della componente audio rispetto a quelli riferiti alla componente video non sembrano presentare regolarità, né per ognuna delle due categorie di utenti, né all'interno dello stesso video. A voler forzare l'analisi, escludendo quindi i dati contraddittori, è possibile trarre tre parziali conclusioni, non tutte valide, perché in contraddizione le une con le altre:

- se si ignorano i dati riferiti al video 2, è possibile riscontrare una maggiore comprensione da parte dei sordi della componente visiva veicolata dai sottotitoli rispetto alla comprensione della componente acustica. Se questa ipotesi fosse suffragata dai dati, questo significherebbe che il tempo dedicato alla componente visiva è superiore a quello dedicato alla lettura dei sottotitoli, che quindi non monopolizzerebbero l'attenzione dello spettatore, neanche in caso di sottotitoli automatici, che superano in termini quantitativi quelli prodotti manualmente. Questa ipotesi è parzialmente avvalorata da una caratteristica intrinseca dei primi due video: la secondarietà delle immagini rispetto al testo esclusivamente veicolato tramite la componente acustica;

- se si ignorano i dati riferiti al video 3, un'ipotesi immediata ma poco probabile è l'andamento diametralmente opposto della comprensione delle componenti semiotiche da parte dei sordi rispetto alla comprensione degli udenti. Infatti, ad eccezione del video 3, ogni volta che la comprensione della componente acustica da parte dei sordi risulta inferiore rispetto a quella visiva, negli udenti si ha una situazione ribaltata (comprensione della componente acustica superiore a quella visiva). Allo stesso modo, ogni volta che la comprensione della componente acustica da parte dei sordi risulta superiore rispetto a quella visiva, negli udenti la comprensione della componente acustica risulta inferiore a quella visiva. Questa ipotesi potrebbe essere dettata dalla scarsa influenza dei sottotitoli sulla visione dei programmi da parte degli udenti;

- se si ignorano i dati riferiti al video 4, emerge una ipotesi più ragionevole della precedente per quanto banale: la qualità dei sottotitoli è direttamente proporzionale alla comprensione delle informazioni veicolate dai sottotitoli stessi e sul tempo dedicato alla componente video. Se si analizzano i dati dei soli sordi, si noterà, infatti, che i video sottotitolati manualmente comportano una maggiore comprensione della componente acustica contrariamente a quanto accade nei video sottotitolati automaticamente, che contengono errori e/o troppo testo.

3.4.5. Documentari culturali vs. Notiziari economici

Un'interessante analisi è quella riferita alla comprensione dei due tipi di video (documentario culturale e notiziario economico) in base alle preferenze dei singoli rispondenti, ai quali è stato chiesto di indicare i programmi preferiti. L'ipotesi è che la comprensione dei singoli programmi dipenda dalla maggiore o minore affinità del singolo utente con il tipo di programma in questione.

Utenti	Preferenze	Componente acustica documentario culturale	Componente visiva documentario culturale	Componente acustica notiziario economico	Componente visiva notiziario economico
Sordi	Cultura	4,4/6	3,7/6	3,2/6	4,1/6
	Economia	3/6	2,9/6	4,8/6	4,9/6
Udenti	Cultura	4,7/6	4,5/6	4,2/6	4,2/6
	Economia	4,3/6	4,3/6	4,6/6	4,4/6

Tabella 5: Comprensione suddivisa per tipo di programma e tipo di utente

L'ipotesi appare immediatamente e inequivocabilmente suffragata dai dati riportati in Tabella 5. Sia i sordi, sia gli udenti sembrano comprendere meglio sia la componente acustica (veicolata o meno dai sottotitoli), sia la componente visiva del tipo di programma che preferiscono guardare in TV o su Internet. Per quanto riguarda i sordi, questo sembra valere per i programmi sottotitolati manualmente e per i programmi sottotitolati automaticamente.

Questo ci porta a un'ulteriore considerazione: in caso di programma con il quale lo spettatore ha maggiore affinità, la qualità del sottotitolo ha una minore influenza sulla comprensione. L'effetto positivo della maggiore affinità di un utente sordo con il programma sembra esserci anche in termini di tempo di lettura (più rapidi) e di tempo a disposizione per guardare le immagini.

Guardando i dati con maggiore attenzione, si può notare infine che la variazione sia più forte nei sordi. La ragione dipende dal maggiore sforzo cognitivo compiuto dai sordi nell'usare la vista come solo canale sensoriale per accedere a entrambe le componenti dei video. Tuttavia, una domanda sorge spontanea: sordi e udenti comprendono meglio un programma preferito perché hanno davvero colto il senso del filmato o perché conoscevano le risposte *ex ante*? Certamente una conclusione che può essere tratta è che nel caso di video preferito, la relativa minore qualità dei sottotitoli automatici rispetto a quella dei sottotitoli manuali avrà un effetto minore sulla comprensione, dato che l'utente userà la propria conoscenza per sopperire a eventuali carenze dei sottotitoli. Questo offre ai sottotitoli automatici una più ampia gamma d'impiego.

3.4.6. La sottotitolazione interlinguistica automatica

L'ultima analisi riguarda i dati sulla sottotitolazione automatica dall'inglese in italiano. I dati riguardanti il video 5 sono particolarmente interessanti per diverse ragioni. Innanzitutto perché udenti e sordi sono stati messi su un simile livello**[6]**, quindi tutti i partecipanti hanno utilizzato i sottotitoli automatici per poter rispondere alle domande di comprensione poste nel questionario.

Comprensione	Sordi	Udenti
Acustica	3,1/6	3,7/6
Visiva	3,5/6	3,1/6
Generale	3,3/6	3,4/6

Tabella 6: Comprensione del video 5 suddivisa per tipo di comprensione e tipo di utente

Dai dati della Tabella 6, emerge una sensibile differenza tra la media dei dati riferiti alla comprensione degli altri video (più evidente negli udenti), spiegabile con l'introduzione della doppia automazione (della trascrizione e della traduzione). Da qui emerge una prima conclusione: quando sordi e udenti sono posti su un livello di relativa parità rispetto alla visione di un filmato (perché entrambi i gruppi devono sostanzialmente dipendere dai sottotitoli per comprenderlo), i risultati in termini di comprensione tendono a equivalersi.

É interessante, tuttavia, notare i dati scorporati per tipo di domande al fine di comprendere come, nel caso dei sordi, la comprensione del video sia maggiormente dettata da una prevalenza di risposte corrette alle domande sulla componente visiva, mentre negli udenti prevalgono le risposte alle domande su informazioni veicolate acusticamente. Questo ci porta a tre conclusioni:

> dconsiderato che i sordi sono più abituati a usare la vista per comprendere un video, essi tendono a trarre vantaggio da questa maggiore abilità, quando sono messi su un livello di relativa parità con gli udenti, riuscendo a cogliere più informazioni provenienti dalla componente visiva del video e più agevolmente;
>
> anche in caso di sottotitoli non perfetti, i sordi si fanno meno distrarre dall'errore e procedono a cogliere il senso delle informazioni veicolate dai sottotitoli per potersi concentrare sulle immagini il più possibile;
>
> gli udenti sono più disturbati dall'errore nei sottotitoli**[7]** rispetto ai sordi, perché tendono a concentrarsi più sulla correttezza grammaticale dei sottotitoli e meno sul senso.

Dall'analisi delle risposte al questionario sui dati personali emerge che quasi la metà degli udenti ha usato i sottotitoli per completare le loro limitate conoscenze della lingua inglese. Risulta quindi chiaro il motivo della maggiore prevalenza della comprensione alle domande sulla componente acustica (che maggiormente veicola informazioni di tipo verbale) rispetto a quella visiva.

3.5 Discussione

La sperimentazione che ha coronato il percorso del progetto CLAST ha portato alla luce interessanti risultati degli sforzi compiuti nello sviluppo dei software di trascrizione e traduzione automatiche. La prima evidente e scontata conclusione riguarda la maggiore comprensione dei video da parte degli udenti, specialmente se sottotitolati automaticamente. Questo dato conferma la maggiore difficoltà dei sottotitoli di permettere l'accessibilità di informazioni pensate per un pubblico udente. Tuttavia esso cela un interessante e inaspettato aspetto, cioè la maggiore comprensione dei sordi del video 4, che comporta due ipotesi interessanti:

> la maggiore abilità dei sordi a usare la vista offre loro un vantaggio sugli udenti nella comprensione di programmi la cui componente visiva svolge un ruolo prevalente;
>
> i sottotitoli automatici non inficiano la comprensione delle informazioni veicolate dalle immagini, neanche qualora queste fossero collegate a sottotitoli non corretti.

Queste considerazioni valgono non solo in termini assoluti, ma anche e soprattutto in termini relativi, cioè quando si rapporta la media delle risposte corrette all'effettiva quantità di informazioni veicolata dai video sottotitolati. Da questo calcolo della CR dei video sottotitolati automaticamente, si deduce che il divario con i video sottotitolati manualmente è meno importante della metà.

Considerando le differenze tra comprensione delle informazioni veicolate dalla componente acustica e informazioni veicolate da quella visiva, emergono altre due conclusioni fondamentali:

sia i sordi, sia gli udenti comprendono meglio sia la componente acustica (veicolata o meno dai sottotitoli), sia la componente visiva del tipo di programma che preferiscono.

per quanto riguarda i sordi, questo sembra valere sia per i programmi sottotitolati manualmente, sia per quelli sottotitolati automaticamente.

Pertanto, in caso di programma con il quale lo spettatore ha maggiore affinità, la qualità del sottotitolo ha una minore influenza sulla comprensione dello stesso, dato che userà la propria conoscenza per sopperire a eventuali carenze dei sottotitoli. Per quanto riguarda i sottotitoli interlinguistici prodotti automaticamente, qualora gli udenti ne avessero bisogno per comprendere le informazioni contenute nei video, essi creano una situazione che permette maggiormente di comparare le prestazioni di sordi e udenti. In termini di comprensione tendono, infatti, a equivalersi, mostrando una verità sostanziale: l'abitudine a utilizzare i sottotitoli rende i sordi meno distratti da eventuali errori in essi contenuti. Di contro, gli udenti sono maggiormente disturbati dall'errore nei sottotitoli rispetto ai sordi, perché si concentrano di più sulla loro correttezza formale.

Da questa analisi, i sottotitoli automatici intra-linguistici e interlinguistici, mostrano di assolvere alla loro funzione di veicolo di informazioni, specialmente in caso di video che tratta argomenti affini allo spettatore. Sarà quindi interessante capire quanto lo sviluppo di questa tecnologia abbia influito su altre discipline, come ad esempio la sottotitolazione in tempo reale.

4. Interazione Uomo-Macchina nella sottotitolazione in tempo reale

Nel 2018, la Città Metropolitana di Roma ha avviato il progetto Tirone per l'accessibilità universale delle sedute consiliari: un servizio di sottotitolazione intralinguistica (da Italiano a Italiano) e interpretariato in Lingua dei Segni Italiana (LIS) in tempo reale delle sedute consiliari trasmesse in *streaming***[8]**. L'obiettivo è renderle fruibili al pubblico sordo *in primis* e a tutte le persone che dovessero trovare utile o necessario accedervi tramite i sottotitoli in italiano o la LIS.

Il principio ispiratore del Progetto Tirone è il concetto di progettazione universale così come interpretato dalla Fondazione ASPHI onlus: "approccio incentrato sull'utente (...), al fine di non proporre una soluzione unica per tutti, piuttosto un prodotto capace di fornire diverse alternative per soddisfare (meglio se automaticamente, apprendendo e adattandosi) l'insieme di abilità, requisiti e preferenze dei singoli utenti**[9]**. Come mostrato dalla Figura 4, il processo di produzione del servizio di accessibilità del progetto Tirone avviene in più fasi e su più tracce, ma il servizio viene fornito in un unico *stream* disponibile sul canale YouTube del Comune di Roma.

Il presente paragrafo fornisce un'analisi quali-quantitativa della fase sperimentale del progetto e delle raccomandazioni fornite da una consultazione pubblica con i rappresentanti delle associazioni di sordi oralisti e segnanti, che hanno aggiunto il punto di vista dell'utenza finale ai dati oggettivi. Dopo una breve illustrazione del metodo seguito nel progetto Tirone (§4.1) e dell'analisi quali-quantitativa dei sottotitoli e dell'interpretariato LIS delle sedute consiliari (§4.2), seguiranno le raccomandazioni dei tecnici e il punto di vista dell'utenza finale sorda, raggruppate per garantire una maggiore usabilità del presente lavoro (§4.3) e uno sviluppo della tecnologia in materia (§5).

4.1. Metodo

In questo progetto si è seguito un metodo parzialmente diverso da quello utilizzato nel progetto CLAST. In particolare, sono stati selezionati 3 campioni di 10 minuti ciascuno relativi alla diretta in streaming di tre sedute del consiglio capitolino in maniera del tutto casuale. L'audio originale è stato trascritto e suddiviso in unità concettuali. Le unità concettuali sono state paragonate ai sottotitoli e all'interpretariato LIS. Grazie alla tassonomia IRA già descritta precedentemente, è stato possibile valutare la qualità dei sottotitoli (§4.2.1) e dell'interpretariato (§4.2.2). Rispetto all'analisi precedente, si è proceduto in questo caso ad applicare la tassonomia IRA per intero, permettendo così di comprendere come sono state rese (ripetizione o alterazione, ulteriormente suddivisa in riduzione, correzione o errore marginale) o non rese (omissione o travisamenti, ulteriormente suddivisi in errori grammaticali e lessicali) le unità concettuali in questione. Come nello studio precedente, è stata presa in considerazione la soglia del 95 per cento come soglia minima dell'accuratezza, corrispondente al 98 per cento della tassonomia NER (Romero Fresco e Martínez 2015) usata da Ofcom. Dall'Ofcom è stato ripreso anche il concetto di qualità dei servizi di accessibilità, vale a dire una resa del maggior numero possibile di unità concettuali, utilizzando (nel caso dei sottotitoli) il maggior numero possibile di parole del testo di partenza (Ofcom 2013). Invece di procedere a un test sulla ricezione dei sottotitoli, si è proceduto poi con una consultazione pubblica alla quale hanno preso parte le associazioni di categoria. Durante l'incontro sono stati illustrati i dati e discussi i risultati in termini di punti di forza e opportunità di sviluppo del servizio (§4.3).

4.2. Risultati

In questo paragrafo saranno brevemente illustrati i dati relativi alla qualità oggettiva dei sottotitoli in tempo reale e dell'interpretariato in LIS di tre sedute di un consiglio capitolino risultante dall'applicazione della tassonomia IRA precedentemente illustrata.

4.2.1. La qualità dei sottotitoli intralinguistici

All'interno del progetto Tirone, la piattaforma utilizzata per l'accessibilità delle sedute consiliari prevede una doppia modalità di produzione dei sottotitoli in tempo reale:

assistita: un respeaker detta i sottotitoli al software di riconoscimento del parlato che li trascrive e un live editor li corregge;

siretta: il software produce i sottotitoli automaticamente e il live editor li corregge.

Dall'analisi dei sottotitoli prodotti secondo la modalità assistita, emerge subito che la quantità delle unità concettuali rese nei sottotitoli è del 96,2 per cento, superiore al criterio minimo utilizzato come riferimento (Figura 1).

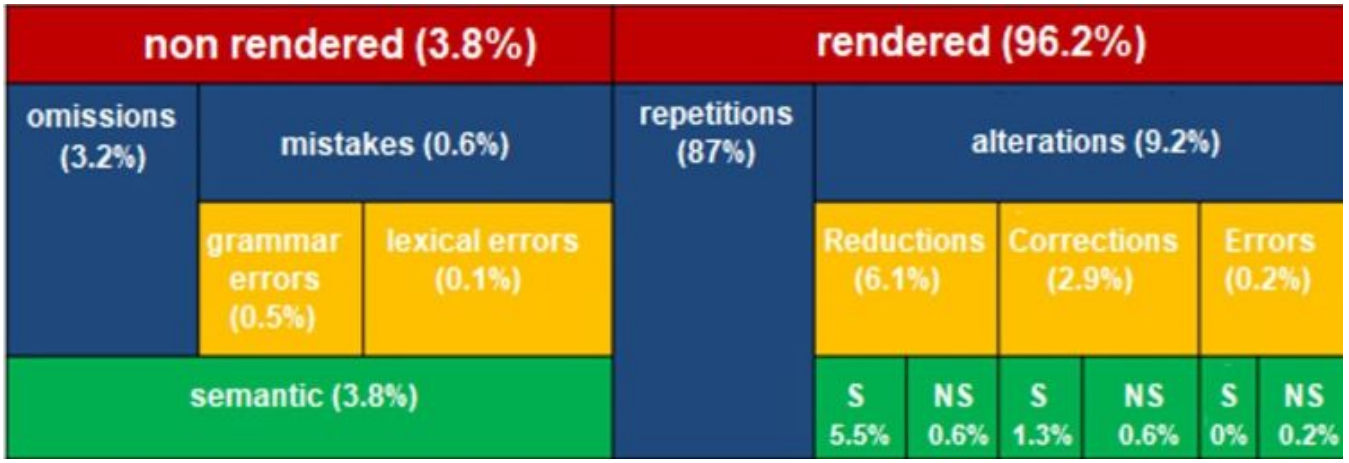

Figura 1: Risultati dell'analisi della qualità dei sottotitoli prodotti dal respeaker

A uno sguardo più approfondito ai risultati dell'analisi condotta (Figura 1), emerge anche che i sottotitoli sono stati prodotti con un numero di ripetizioni molto alto (87 per cento), conformemente alle linee guida Ofcom. Per quanto riguarda le alterazioni, o modifiche rispetto al testo di partenza, esse rappresentano un decimo (9,2 per cento) delle strategie adottate dai sottotitolatori per giungere al risultato finale. Tra queste, preponderanti sono le riduzioni (omissione di parole ridondanti o compressione di pensieri complessi) e le correzioni (principalmente grammaticali) dell'oratore da parte dei sottotitolatori. Trascurabili sono gli errori che non influiscono sulla comprensibilità dei sottotitoli. Quanto alle unità concettuali che non sono passate nei sottotitoli (3,8 per cento), notiamo che sono principalmente costituite dalle omissioni dell'unità concettuale stessa. Si tratta principalmente di unità che garantiscono la transizione da un concetto all'altro con una portata semantica marginale. Tuttavia la coesione testuale risulta essere intaccata, seppur in maniera molto circostanziata. Gli errori riscontrati sono pochi, ma in questo caso compromettono la comprensibilità dei sottotitoli. Infine il ritardo registrato è di 3,9 secondi in media. Nel dettaglio esso è di 3,4 secondi nel caso di sottotitoli contenenti ripetizioni del testo pronunciato e di 4,5 quando il sottotitolatore lo modifica. Tale ritardo è conforme a quanto richiesto dai parametri di qualità previsti dal Televideo – RAI (6 secondi).

Se si mettono questi dati a confronto con quelli dei sottotitoli prodotti tramite la seconda modalità (diretta), emergono alcune conclusioni, di cui alcune saltano subito all'occhio, mentre altre richiedono una maggiore attenzione (Figura 2).

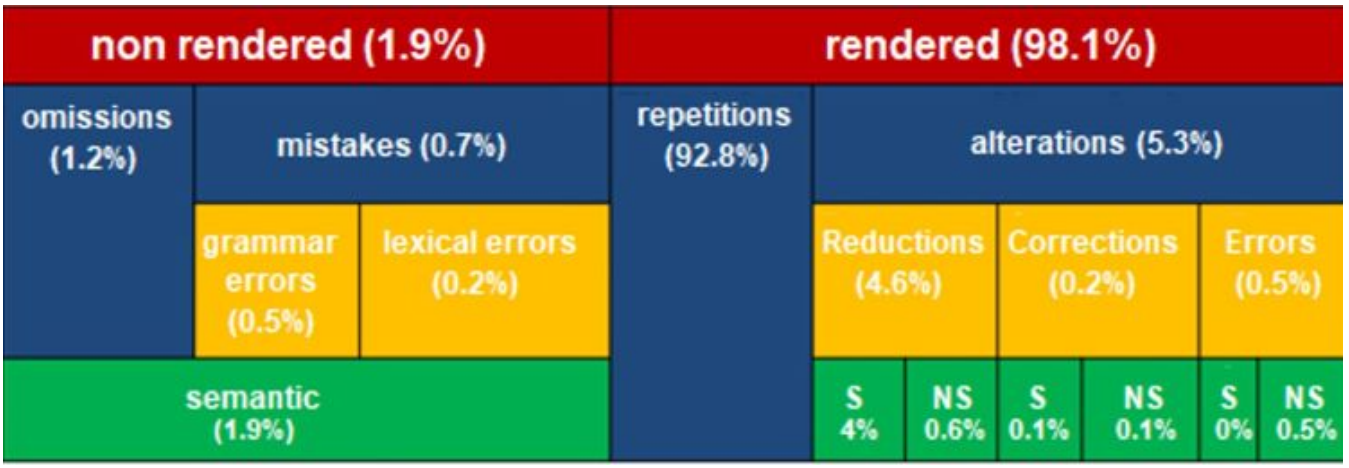

Figura 2: Risultati dell'analisi della qualità dei sottotitoli prodotti dalla macchina

In primis, il numero di ripetizioni prodotte con questa modalità è superiore rispetto a quelle prodotte dalla modalità assistita dal respeaker, nonostante l'intervento del live editor. Questo avviene per due motivi fondamentali: 1) nonostante provi un approccio verbatim alla sottotitolazione, il respeaker non riesce, come invece fa la macchina, a ripetere tutte le parole dell'originale e quindi 2) seleziona quelle essenziali e tralascia quelle ridondanti, agevolando così la leggibilità dei sottotitoli. Il secondo motivo si applica anche per giustificare la maggiore presenza di altre forme di riduzione nella sottotitolazione diretta, come la compressione nei sottotitoli prodotti dal respeaker. Un ulteriore risultato evidente riguarda il maggior numero di errori presenti nella modalità diretta rispetto a quella assistita, dovuti intuitivamente a una minore qualità della trascrizione proveniente dalla macchina. Questa non è chiaramente imputabile alla macchina, ma alla qualità del input: mentre un respeaker è abituato a dettare in maniera professionale alla macchina, il politico che viene trascritto è meno attento alla presenza del software di riconoscimento del parlato e quindi produce un testo meno chiaro sia foneticamente che grammaticalmente. Tuttavia, la differenza tra errori prodotti nella modalità assistita ed errori prodotti nella modalità diretta non è così sostanziale (0,1 per cento nel caso di errori che alterano l'unità concettuale e 0,3 per cento nel caso di errori che non alterano l'unità concettuale). Un dato che invece sembra controintuitivo riguarda le unità concettuali rese. Stante la seconda motivazione appena menzionata, l'ipotesi è che il respeaker produca sottotitoli maggiormente accurati. Se questo è vero in termini di qualità dell'input, i dati dimostrano che la seconda modalità garantisce una resa più alta delle unità concettuali rispetto alla seconda (98,1 per cento e 96,2 per cento rispettivamente). Questo perché il respeaker, come risultante della prima motivazione, non riesce a tenere sempre il passo dell'oratore non solo perché l'oratore parla più velocemente di quanto riesca il respeaker a dettare, ma anche per ragioni che pertengono a una bassa qualità del discorso da sottotitolare (in termini grammaticali e fonetici) e alla stanchezza del respeaker. Visto che ne consegue una riduzione quantitativa del testo iniziale, anche il numero di unità concettuali non rese risulta maggiore, sia che esse comportino una mancata resa dell'unità concettuale in questione (3,2 per cento in modalità assistita contro il 1,2 per cento nella modalità diretta) oppure no (6,1 per cento di riduzioni in modalità assistita contro il 4,6 per cento in diretta).

4.2.2. La qualità dell'interpretariato in LIS

Passando all'analisi dell'interpretariato in LIS, la tassonomia IRA è stata adattata in maniera da contemplare la differenza tra i sottotitoli e l'interpretariato (i primi sono intra-linguistici, cioè dall'italiano all'italiano, la seconda è interlinguistica). A tal fine, le ripetizioni sono state sostituite con le traduzioni complete, cioè traduzioni che non trascurano alcun elemento dell'originale, fatte salve le differenze grammaticali tra le due lingue (Figura 3). L'analisi mostra che la quantità delle unità concettuali rese è superiore al minimo richiesto (95,4 per cento) e che le unità concettuali rese sono state prodotte con un numero di traduzioni complete molto alto (88,1 per cento), conformemente alle linee guida Ofcom sulle ripetizioni.

Figura 3: Risultati dell'analisi della qualità dell'interpretariato LIS

Quanto alle alterazioni rispetto al testo di partenza, esse sono comprensibilmente inferiori rispetto ai sottotitoli (7,3 per cento) e sono composte perlopiù da riduzioni, vale a dire omissione di parole ridondanti o compressione di pensieri complessi. Assenti le correzioni e trascurabili, seppur in numero maggiore, gli errori che non influiscono sulla comprensibilità del segnato. Per quanto riguarda le unità concettuali che non sono passate nell'interpretariato (4,6 per cento), notiamo anche qui che esse sono principalmente costituite dalle omissioni dell'unità concettuale stessa. Anche in questo caso si tratta di unità che garantiscono la transizione da un concetto all'altro con una portata semantica marginale. La sola differenza con i sottotitoli è rappresentato dalla maggior presenza di errori, che potrebbero essere imputabili alla difficoltà di comprensione del testo dell'oratore, alla maggiore complessità dell'operazione di interpretariato rispetto a quella di ripetizione o ancora all'assenza di una strumentazione che permetta l'ascolto in cuffia da parte dell'interprete. Infine il ritardo registrato è di 0,9 secondi in media. Nel dettaglio, il ritardo è di 0,8 secondi quando il testo non presenta difficoltà e di 1,3 quando sono necessari interventi che riducono il testo originale.

4.2.3. La qualità dei sottotitoli interlinguistici

In uno studio simile condotto dalla Federazione Internazionale di Elaborazione dell'Informazione e della Comunicazione Intersteno, si è portato avanti il *Communication Project* (iniziato nel 2017, cfr. Eugeni *et al.* 2018), volto alla valutazione del migliore flusso di lavoro possibile in termini di immediatezza, economicità e accuratezza per la produzione di sottotitoli interlinguistici (da inglese a italiano o francese) in tempo reale (ILS) di riunioni di varia natura, quali conferenze, assemblee e consigli. Lo studio ha testato i seguenti flussi di lavoro di produzione di ILS nell'ambito del Congresso Intersteno 2019:

1. un interprete traduce la riunione in un'altra lingua e uno stenotipista trascrive (flusso di lavoro *Human-Only*);
2. un respeaker interlinguistico produce i sottotitoli direttamente in lingua straniera (flusso di lavoro *Human-Only*);
3. un professionista trascrive la riunione parola per parola e una macchina traduce (flusso di lavoro *Computer-Aided*);
4. un professionista intralinguistico trascrive la riunione semplificando la sintassi e una macchina traduce (flusso di lavoro *Computer-Aided*);
5. un software ASR trascrive la riunione e un live editor corregge eventuali errori (flusso di lavoro *Human-Aided*).

Si noti che il flusso ILS 3 era stato già testato nel 2017 (Manetti 2018) in un contesto simile. A variare, in questo caso, sono il professionista (nel *Communication Project* – ILS 3a – un respeaker, nell'altro caso – ILS 3b – un velotipista), la coppia linguistica dei sottotitoli valutata (dall'inglese in italiano in ILS 3a e dall'inglese in francese in ILS 3b), la qualità del software di traduzione automatica intuitivamente superiore in ILS 3a e le variabili testate (in ILS 3b solo il numero di parole al minuto e la qualità). Per ognuno di questi cinque flussi, sono state testate 4 variabili:

fedeltà relativa dei sottotitoli misurata in numero di parole prodotte rispetto a quello originale (WORDS);

velocità di scorrimento dei sottotitoli misurata in numero di parole al minuto (WPM);

accuratezza dei sottotitoli misurata in numero di idee concettuali rese (IRA);

ritardo dei sottotitoli misurati in secondi tra l'occorrenza di un'idea concettuale e la comparsa a schermo del relativo sottotitolo (DELAY).

Come mostra la Tabella 7, a un primo sguardo risulta chiaro che il flusso più fedele al testo di partenza è ILS 5 (è anche il solo flusso *Human-Aided* testato); in termini di velocità di lettura, che dipende dal numero di parole al minuto del discorso originale; il flusso che produce il maggior numero di parole e anche una qualità dei sottotitoli minore è ILS 3b, *Computer-Aided*; mentre quello più facile da leggere perché produce meno parole al minuto è ILS 1, che è *Human-Only* e anche il più accurato; il flusso con il ritardo minore è ILS 3a, altro *flusso Human-Aided*. In altre parole, in termini assoluti e senza distinzione tra i vari flussi di lavoro, quelli *Human-Only* sono più accurati ma anche meno fedeli, quelli *Computer-Aided* sono più rapidi ma anche meno accurati e quelli *Human-Aided* sono più fedeli ma anche meno rapidi.

	ILS 1	ILS 2	ILS 3a	ILS 3b	ILS 4	ILS 5
IRA	97,3%	95,8%	91,6%	71,2%	92,1%	86,9%

WORDS	79,6%	82,2%	87,7%	n.a.	86%	95,5%
WPM	110	114	121	141	118	132
DELAY	4,3"	1,8"	1,3"	n.a.	3,8"	5,1"

Tabella 7: Valutazione della qualità dei sottotitoli interlinguistici

Se si guardano i dati un po' più nel dettaglio, ci si può rendere conto che i flussi *Human-Only* (ILS 1 e ILS 2) sono gli unici due flussi che garantiscono un'accuratezza professionalmente accettabile (97,3 per cento e 95,8 per cento). Per farlo, però, hanno bisogno di un importante lavoro editoriale a opera dei professionisti coinvolti che devono possedere competenze di interpretariato e di sottotitolazione in tempo reale. Questo può risultare nella produzione di un numero di parole di circa il 20 per cento inferiore rispetto a quelle pronunciate del testo originale (79,6 per cento e 82,2 per cento) e quindi una velocità di lettura del sottotitolo inferiore (110 e 114 wpm). Se, da un lato, questo garantisce una maggiore leggibilità del sottotitolo, dall'altro può anche creare dissonanza cognitiva negli utenti. Da segnalare è anche il ritardo maggiore per ILS 1 (4,3"). Tuttavia, se un solo professionista (in ILS2 un respeaker interlinguistico) si occupa sia della traduzione sia della trascrizione, il ritardo si riduce sostanzialmente (1,8"). Un aspetto da considerare, non misurato nella Tabella 7, è l'investimento da parte dell'eventuale cliente, che oltre a una squadra di professionisti per turno, deve anche considerarne una squadra per ogni coppia linguistica di cui dovesse avere bisogno, perché i flussi *Human-Only* sono chiaramente *language dependent*. Tra i due flussi, l'uso di un respeaker interlinguistico risulta chiaramente più conveniente per via del numero minore di professionisti coinvolti, che comporta meno ritardo e meno spesa.

Per quanto riguarda i flussi *Computer-Aided* (ILS 3a, ILS 3b e ILS 4), i dati che saranno qui analizzati riguardano solo ILS 3a e ILS 4, dato che ILS 3b (71,2 per cento) è stato testato in un periodo in cui le tecnologie della trascrizione e della traduzione erano intuitivamente meno avanzate. Essi garantiscono un'accuratezza comunque superiore al 90 per cento (91,6 per cento e 92,1 per cento). Questa qualità, comunque accettabile anche se inferiore al 95 per cento stabilita come soglia minima nella sottotitolazione intralinguistica misurata con IRA, comporta un intervento editoriale minore rispetto ai flussi *Human-Only* (87,7 per cento e 86 per cento) e quindi una velocità di lettura del sottotitolo superiore, anche se di poco (121 e 118 wpm). Questo crea comunque una buona leggibilità del sottotitolo (stabilita da Rai tra le 120 e le 180 parole al minuto**[10]**) e un'accuratezza superiore a quella dei flussi *Human-only* (87,7 per cento e 86 per cento). Quanto al ritardo, ILS 3a (e intuitivamente anche ILS 3b) garantisce il ritardo più basso (1,3") mentre ILS 4 implica un ritardo maggiore, probabilmente dovuto al fatto che il sottotitolatore non si limita a ripetere il testo di partenza (come in ILS 3a e 3b), ma deve pensare a semplificare il testo di arrivo. Per questo motivo, ci si sarebbe aspettati una produzione di parole rispetto al testo di partenza, e quindi anche di parole al minuto, di molto inferiore a quelle prodotte da ILS 3a. Se la velocità di produzione del testo di partenza può spiegare differenze nella compressione o nel numero di parole al minuto, essa non può giustificare entrambe. Quindi, se risulta abbastanza chiaro che il respeaker in ILS4 non abbia operato una semplificazione del testo di partenza maggiore rispetto a quanto fatto dal respeaker in ILS 3a, le ragioni potrebbero essere molteplici. Tra le più probabili, il respeaker non è riuscito a semplificare il testo di partenza e il testo di partenza non aveva bisogno di ulteriore semplificazione. In questo specifico ambito può risultare utile dare uno sguardo ai dati relativi a ILS 3b che mostrano chiaramente che il testo di arrivo aveva una velocità di produzione di molto superiore a quella degli altri flussi (141 wpm). In termini di spesa, i flussi *Computer-Aided* somigliano a ILS 2 per economicità ma garantisce una maggiore scalabilità del servizio, dato che non serve una squadra per ogni coppia linguistica di cui dovesse avere bisogno, perché i flussi *Computer-Aided* sono *language independent*.

Quanto al solo flusso *Human-Aided* analizzato (ILS 5), esso garantisce la minore accuratezza di tutti gli altri flussi. Tuttavia, se si guardano i dati di ILS 4, si nota che la qualità, in questo caso è notevolmente superiore, dato particolarmente incoraggiante se si considera che la distanza le variabili tra i due sono non solo la distanza di due anni tra un esperimento e l'altro (2017 e 2019), ma anche il flusso di lavoro, che in ILS 5 non prevede una gestione del flusso stesso da parte del professionista. Un ulteriore dato evidente riguarda la fedeltà del testo di arrivo rispetto a quello di partenza (95,5 per cento). Incrociando questo dato con il precedente, risulta chiaro che l'automazione, da una parte implica una trascrizione e una traduzione del 100 per cento del testo di partenza; e dall'altra implica una serie di errori che vengono corretti o eliminati dal *live editor/scopist*. Quest'ultimo dato viene anche confermato dall'alto tasso di fedeltà relativa (95,5 per cento) e dal ritardo con cui i sottotitoli vengono rilasciati (5,1"). In un contesto professionale, i sottotitoli prodotti in modalità *Human-Aided* sono anche intuitivamente migliori di sottotitoli prodotti in modalità *Computer-Only*, nonostante interessanti sviluppi anche in questo ambito (Romero-Fresco 2015). A far sorgere alcune perplessità sono la qualità dell'input, che può comportare diverse rese qualitative (Pagano 2020), e la conseguente impossibilità del professionista di gestire il flusso di lavoro. In termini economici, se la spesa per i software risulta ammortizzabile sul lungo periodo, la modalità *Human-Aided* richiede un numero di squadre di sottotitolazione pari a quello delle coppie linguistiche richieste.

4.3. Consultazione pubblica

Dalla tavola rotonda organizzata con il gruppo di esperti e con i rappresentanti dell'utenza sorda, è emerso che il servizio di sottotitolazione e di interpretariato in LIS soddisfa i criteri tecnici previsti dal bando e i tecnici di qualità previsti dall'Ofcom. Tuttavia, è stata proposta una serie di migliorie per aumentare la qualità tecnica dei sottotitoli e dell'interpretariato, così come la fruibilità degli stessi da parte dell'utenza finale. Per quanto riguarda l'interpretariato, si è raccomandato un ingrandimento del riquadro contenente l'interprete LIS e uno sfondo che garantisca una maggiore visibilità dei segni. Inoltre è stato consigliato l'uso delle cuffie da parte degli interpreti per meglio sentire l'originale. Quanto ai sottotitoli, è stato raccomandato l'uso di tre righe per garantire una maggiore leggibilità delle frasi (spesso lunghe) in essi contenute. Tra i desiderata la creazione di un'applicazione per poter seguire più agevolmente le sedute tramite cellulare, che personalizzi il servizio di accessibilità con un'opzione che permetta di scegliere tra interpretariato LIS, sottotitoli, diretta e una combinazione delle tre. L'applicazione dovrebbe anche offrire la possibilità di regolare il ritardo dei sottotitoli e dell'interpretato in LIS in maniera da garantire una maggior sincronizzazione tra il servizio di accessibilità e la diretta streaming. Un'ultima raccomandazione ha riguardato la possibilità di riutilizzare i sottotitoli integrandoli nel processo di resocontazione (§5).

5. Interazione Uomo-Macchina nella resocontazione

Capitalizzando sul concetto di progettazione universale e, tra le altre, la raccomandazione sul riutilizzo dei sottotitoli in tempo reale per la produzione di resoconti, la piattaforma utilizzata per la resa accessibile delle sedute capitoline può essere impiegata in un flusso di lavoro ideale che si rifà al concetto di verbale multimediale, proposto in Italia in seno al gruppo di ricerca del Governo ForumTAL,**[11]** con l'obiettivo di digitalizzare al massimo il processo penale e garantire così una Giustizia più rapida ed efficacie. Il verbale multimediale crea un flusso di lavoro in cui la verbalizzazione passa da *Computer-Aided* a *Human-Aided* e il verbale da oggetto ultimo della verbalizzazione a flessibile strumento del processo penale stesso (Figura 4).

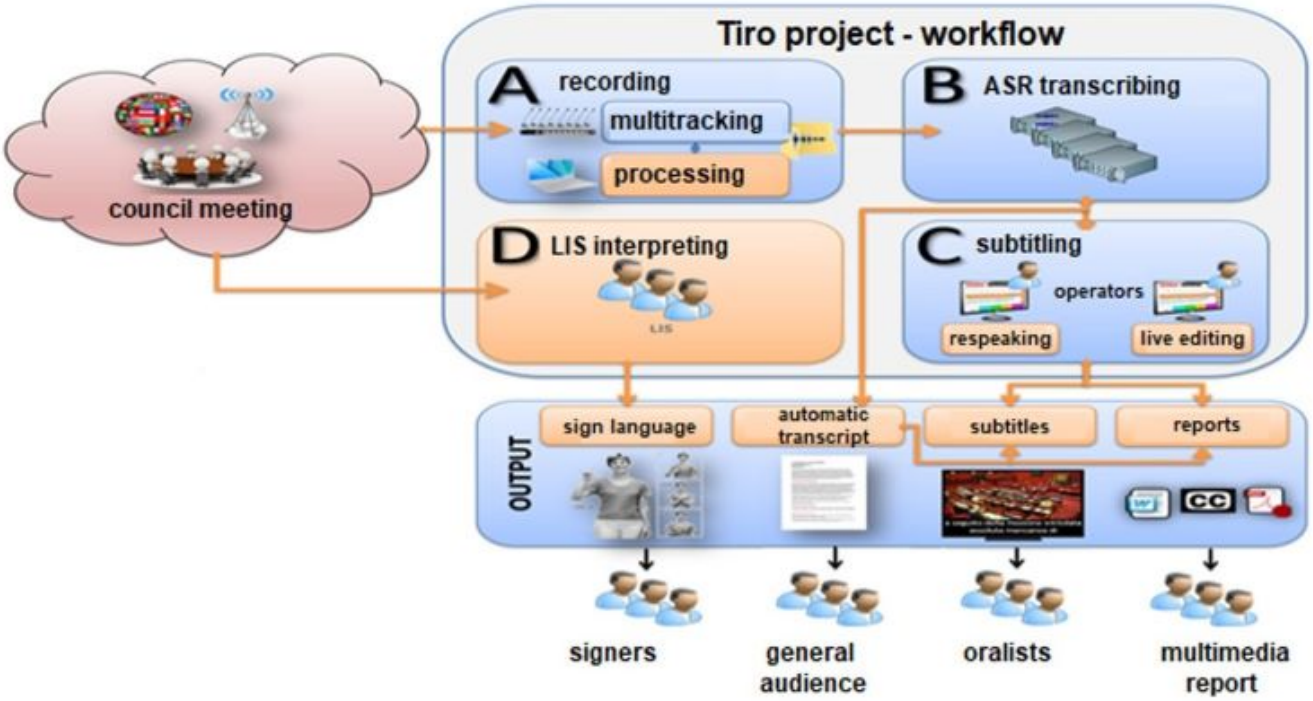

Figura 4: Flusso di lavoro per un resoconto multimediale

Il flusso di lavoro del resoconto multimediale parte con la registrazione multitraccia della seduta simultaneamente al suo svolgimento (fase A del flusso di lavoro). Oltre alla registrazione, la piattaforma che gestisce il flusso di lavoro che porta - tra gli altri servizi - al resoconto multimediale, procede anche all'elaborazione dei dati in ingresso e alla loro trascrizione automatica tramite software di riconoscimento del parlato (fase B). L'elaborazione dei dati serve da riferimento per la sottotitolazione che passa per un ulteriore postazione di lavoro. Quest'ultima permette alla squadra di sottotitolazione di scegliere tra una delle due modalità di produzione dei sottotitoli summenzionate (assistita o diretta) l'uso diretto della trascrizione prodotta nella fase B o l'inserimento dei sottotitoli da parte del respeaker nel caso in cui la qualità della trascrizione non permettesse una facile e accurata impaginazione dei sottotitoli in tempo reale da parte del live editor (fase C). Queste tre fasi corrono parallele all'interpretariato in LIS (fase D) che viene registrato in un'altra traccia rispetto a quella dei sottotitoli, entrambi trasmessi nella stessa traccia o in tracce diverse rispetto a quella della seduta da rendere accessibile.

Dalle fasi di produzione si arriva alla ricezione dei singoli servizi offerti da parte degli utenti finali, che possono scegliere se attivarli in combinazione con altri servizi:

i segnanti avranno la possibilità di accedere all'interpretariato in LIS prodotto nella fase D simultaneamente all'incontro, con o senza il fisiologico ritardo tra i due;

gli oralisti, sordi e non, avranno la possibilità di accedere

alla trascrizione automatica prodotta dalla macchina nella fase B,

ai sottotitoli intralinguistici prodotti in modalità respeaking diretta o automatica in fase C simultaneamente all'incontro, con o senza il fisiologico ritardo tra i due,

ai sottotitoli interlinguistici prodotti automaticamente dopo la modalità diretta o assistita della fase C e simultaneamente all'incontro, con o senza il fisiologico ritardo tra i due,

al resoconto multimediale, contenente il video della seduta e il resoconto, le cui sezioni e parole sono state corrette e indicizzate e sincronizzate con il video stesso.

Conclusioni

L'interazione uomo-macchina nell'ambito della traduzione diamesica ha portato a importanti distinzioni dei flussi di lavoro in *Human-Only, Computer-Aided, Human-Aided Translation* o *Computer-Only*. Grazie all'intelligenza artificiale, le soluzioni *Computer-Only* un tempo neanche considerate, sono già realtà in contesti comunicativi anche molto importanti come le istituzioni europee. Nel tentativo di ricostruire le tappe che hanno portato a questo risultato e per comprendere l'impatto della tecnologia sulla vita di tutti i giorni e il suo potenziale sviluppo, il presente studio di ricerca ha illustrato alcuni progetti condotti in contesti principalmente italofoni. Nell'analizzare e raffrontare sottotitoli preregistrati e in tempo reale, sia intralinguistici che interlinguistici, prodotti principalmente o esclusivamente dall'uomo o dalla macchina, sono emersi interessanti spunti di riflessione non soltanto sulla rapida evoluzione della tecnologia e della sua applicazione nella traduzione diamesica, ma anche in ottica di contributo delle tecnologie assistive alla evoluzione della società nel suo complesso, in una specie di processo inverso rispetto a quanto si è assistito negli ultimi decenni. In particolare, una piattaforma multifunzionale destinata all'accessibilità che riesce anche a garantire rapidità, accuratezza e trasparenza dei processi democratici e legislativi di un Paese è la dimostrazione di quanto l'accessibilità non possa più essere considerata una questione marginale della società, ma intrinsecamente parte della stessa.

Riferimenti bibliografici

Bianchi, Francesco, Eugeni, Carlo e Grandioso, Luisa (2020) "Verbatim vs. adapted subtitling and beyond. An empirical study with deaf, hard-of-hearing and hearing children", in Lingue e Linguaggi, vol. 36 **http://siba-ese.unisalento.it/index.php/linguelinguaggi/article/view/20822**

Cutugno, Francesco e Paoloni, Andrea (2013) "Proposta del ForumTAL sul verbale multimediale di atti giudiziari", in Eugeni, Carlo e Zambelli, Luigi (a cura di) *Respeaking*, Specializzazione on-line, vol. 1, pp. 61-63, **https://accademia-aliprandi.it/public/specializzazione/respeaking.pdf**

Eugeni, Carlo e Gambier, Yves (2023) *La traduction intralinguistique – les défis de la diamésie*, Timisoara – Editura Politehnica.

Eugeni, Carlo (2007) "Il rispeakeraggio televisivo per sordi: per una sottotitolazione mirata del TG", in Intralinea, vol. 9. **http://www.intralinea.org/archive/article/1638**

Eugeni, Carlo (2017) "La sottotitolazione intralinguistica automatica: Valutare la qualità con IRA", in CoMe 2 (1), pp. 102-113, **http://comejournal.com/wp-content/uploads/2017/12/EUGENI-2017.pdf**

Eugeni, Carlo (2019) "Technology in court reporting – capitalising on human-computer interaction". In Topal, Şevket e Yaniklar, Cengiz (a cura di) *I. Uluslararasi adalet kongresi Bildiri kitabi*, Rize: Recep Tayyip Erdoğan Üniversitesi, pp. 853-861 **https://drive.google.com/file/d/1wbATfxDgaRixgK1LnDJpdiTu2RASkkR5/view**

Eugeni, Carlo (2020) "Human-Computer Interaction in Diamesic Translation – Multilingual Live Subtitling", in Dejica, Danca, Eugeni, Carlo e Dejica-Cartis, Anca (a cura di) *Translation Studies and Information Technology - New Pathways for Researchers, Teachers and Professionals*, Timişoara: Editura Politehnica, TSS, pp. 19-31 **www.researchgate.net/publication/345803837_Human-Computer_Interaction_in_Diamesic_Translation_Multilingual_Live_Subtitling**

Eugeni, Carlo e Bernabé, Rocio (2021) "Written Interpretation: When Simultaneous Interpreting Meets Real-Time Subtitling", in Seeber, K. (a cura di) *100 Years of Conference Interpreting – A Legacy*, Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 93-109.

Eugeni, Carlo, Rotz, Allen e Checcarelli, Alessandra (2018) "Il "Communication Project dell'Intersteno". Per una comunicazione internazionale facilitata", in SpeciaLinguaggi, Numero 1. Retrieved from **https://specialinguaggi.accademia-aliprandi.it/2018/01/01/il-communication-project-dellintersteno-per-una-comunicazione-internazionale-facilitata/**

Fantinuoli, Claudio (2023) "Towards AI-enhanced computer-assisted interpreting", in Corpas Pastor, Gloria e Defrancq, Bart (a cura di) *Interpreting Technologies – Current and Future Trends*, Amsterdam e Philadelphia: John Benjamins, pp.46-71. **https://www.claudiofantinuoli.org/docs/ivitra.37.03fan.pdf**

Hewett, Timothy, Baecker, Ronald, Card, Stuart, Carey, Tom, Gasen, Jean, Mantei, Matilyn, Perlman, Gary, Strong, Gary e Verplank, William (1992) *ACM SIGCHI curricula for human-computer interaction*, Broadway: ACM, **https://www.researchgate.net/publication/234823126_ACM_SIGCHI_curricula_for_human-computer_interaction**

Manetti, Ilenia (2018) "L'interaction homme-machine – analyse d'un cas de sous-titrage interlinguistique semi-automatisé", in CoMe III, vol. 1, pp. 57-69 **https://comejournal.com/wp-content/uploads/2019/06/CoMe-III-1-2018.-Completo-web.pdf**

Ofcom (2006) *Television access services – Review of the Code and guidance*. **https://www.ofcom.org.uk/__data/assets/pdf_file/0016/42442/access.pdf**

Ofcom (2013) The quality of live subtitling. London: Office of Communications. **https://www.ofcom.org.uk/consultations-and-statements/category-1/subtitling**

Pagano, Alice (2020) "Verbatim vs. Edited live parliamentary subtitling", in Dejica, Daniel, Eugeni Carlo e Dejica-Cartis Anca (a cura di) *Translation Studies and Information Technology - New Pathways for Researchers, Teachers and Professionals*, Timişoara: Editura Politehnica, TSS, pp. 32-44.

Paivio, Allan (1986) *Mental representations: a dual coding approach*, Oxford: OUP, Oxford.

Romero-Fresco, Pablo e Martínez, Juan (2015) Accuracy Rate in Live Subtitling: The NER Model. In J. Díaz-Cintas & R. Baños Piñero (eds.), Audiovisual Translation in a Global Context. Mapping an Ever-changing Landscape, London: Palgrave, 28-50.

Romero-Fresco, Pablo (2015) (a cura di) *The reception of subtitles for the deaf and hard-of-hearing In Europe*, 1st edition. Bern, Berlin, Brussels, Francoforte, New York, Oxford e Vienna: Peter Lang.

Romero-Fresco, Pablo (2023) "Interpreting for access – The long road to recognition", in Zwischenberger, C., Reithofer, K. e Rennert, S. (a cura di) *Introducing New Hypertexts on Interpreting (Studies) – A tribute to Franz Pöchhacker*. Amsterdam: John Benjamins Publishing Company, pp.236-253, **https://doi.org/10.1075/btl.160.12rom**

Spinolo, Nicoletta e Amato, Amalia (2020) (a cura di) *inTRAlinea Special Issue: Technology in Interpreter Education and Practice*, **https://www.intralinea.org/specials/article/2520**

Note

[1] Il progetto CLAST (*Cross Language Automatic Subtitling Technology*), coordinato da PerVoice, si è concluso nel gennaio 2018, è stato finanziato dalla LP6/99 della Provincia di Trento ed era volto alla realizzazione di un sistema *cloud* per l'automatizzazione della produzione di sottotitoli in lingua originale e in lingua straniera e il doppiaggio.

[2] Per produrre i sottotitoli sono state seguite le linee guida del Televideo Rai, che limitano il testo nei sottotitoli tra i 10 e i 15 caratteri al secondo, secondo la durata del sottotitolo. Le linee guida di Televideo-Rai sono disponibili online **https://www.rai.it/dl/doc/2020/10/19/1603121663902_PREREGISTR_22_feb_2016_-_Norme_e__Convenzioni_essenziali_per_la_composiz...%20-%20Copia.pdf**

[3] La tassonomia IRA (*Idea-unit Rendition Assessment*) si basa sul principio dell'unità concettuale come unità minima di analisi, vale a dire il periodo, la frase o ogni altro concetto di senso compiuto grammaticalmente identificabile. La valutazione consiste, in un primo momento, nel comprendere se i sottotitoli rendono o non rendono l'unità concettuale in questione. La percentuale della qualità si ottiene moltiplicando il numero di unità concettuali rese per 100 diviso il numero di unità concettuali contenute nel testo di partenza (cfr. Eugeni 2017). In mancanza di uno standard nazionale di riferimento per la qualità dei sottotitoli e dell'interpretariato in LIS, è stata fissata la soglia dell'accuratezza al 95 per cento, corrispondente al 98 per cento della tassonomia NER (Romero Fresco e Martínez 2015) utilizzata da Ofcom.

[4] Si veda, per esempio, il progetto SALES, condotto dal 2005 al 2006 presso l'università di Bologna e volto all'inclusione sociale delle persone sorde tramite sottotitoli di programmi televisivi in diretta prodotti con il respeaking. Il sito web del progetto non è più

disponibile. Per informazioni sui contenuti, cfr. Eugeni 2007.

[5] Si veda, per esempio, il progetto DTV4All, coordinato dal 2011 al 2012 dall'Universitat Autònoma de Barcelona e volto allo studio della ricezione dei sottotitoli preregistrati da parte delle persone sorde. Il sito web del progetto non è più disponibile. Per ulteriori informazioni sui contenuti, cfr. Romero-Fresco 2015.

[6] I 16 udenti hanno dichiarato che la loro conoscenza dell'inglese non permetteva loro una vera comprensione del video, per quanto 7 di loro abbiano dichiarato di riuscire a compensare con i sottotitoli la loro comprensione dell'inglese parlato. Dai dati non emergono significative variazioni rispetto a chi ha dichiarato di non avere tale abilità.

[7] Questa conclusione è confermata dal *Think Aloud Protocol*, dal quale emerge che gli udenti, più hanno perso più tempo dei sordi a leggere i sottotitoli contenenti errori perché cercavano di ricostruire il senso sintattico della frase perdendo di vista quello semantico. Tuttavia alcuni udenti hanno dichiarato di perdere molto tempo nella lettura anche dei sottotitoli corretti perché non abituati a vedere un video sottotitolato. Questo è confermato dallo studio dell'Unione Europea sul potenziale dei sottotitoli per l'apprendimento delle lingue straniere in Europa. Per maggiori informazioni si veda la relazione "Etude sur l'utilisation du sous-titrage" dell'EACEA al link **https://op.europa.eu/fr/publication-detail/-/publication/afc5cf17-02f8-459c-b238-1890ee5cca2b** (ultimo accesso 29 febbraio 2024).

[8] Cfr. **https://www.youtube.com/channel/UCCQygqTutREYU9AbTHfHhTw**

[9] Cfr. **https://asphi.it/2013/09/16/progettazione-universale-o-universal-design-o-design-for-all/**

[10] Per ulteriori informazioni, cfr. le linee guida tecniche disponibili online:

https://www.rai.it/dl/doc/2020/10/19/1603121663902_PREREGISTR_22_feb_2016_-_Norme_e__Convenzioni_essenziali_per_la_composiz...%20-%20Copia.pdf

[11] Il sito web del forumTAL non è più disponibile. Per informazioni sul progetto, cfr. Cutugno e Paoloni (2013).
