Alma Mater Studiorum Università di Bologna

DIPARTIMENTO DI INTERPRETAZIONE E TRADUZIONE

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

TESI DI LAUREA

In Interpretazione Dialogica, Multimediale e a Distanza in Spagnolo

Interpreti in formazione e intelligenza artificiale: trascrizione automatica di Google Meet e CAI *tool* di SmarTerp a confronto

Una prospettiva didattica da *User Experience*, il confronto delle rese e i dati di *eye-tracking*

CANDIDATO

Gabriele Giudice

RELATRICE

Prof.ssa Nicoletta Spinolo

CORRELATRICE

Prof.ssa Mariachiara Russo

Sessione luglio 2024

Anno Accademico 2023/2024

A Camilla e Celeste.
Perché siate voci giuste
al fianco degli ultimi,
abbracciando le paure.

"We are computer tacked onto creature […]. And the point isn't to denigrate one, or the other, any more than a catamaran ought to become a canoe. The point isn't that we're half lifted out of beastliness by reason and can try to get even further through force of will. The tension is the point. Or, perhaps to put it better, the collaboration, the dialogue, the duet."
(Christian, 2011, The Most Human Human: A Defence of Humanity in the Age of the Computer, p. 70)

Indice

Riassunto

Nel punto di contatto tra intelligenza artificiale e interpreti in formazione emerge la necessità di favorire una prospettiva didattica che permetta un rapporto interprete-macchina sempre più ergonomico e una facile integrazione dei neolaureati nel mercato. In questo studio, seguendo un *design* quasi-sperimentale e un approccio di ricerca *mixed-method*, vengono messi a confronto il CAI *tool* di SmarTerp e la trascrizione automatica di Google Meet, indagandone i seguenti aspetti: (a) il diverso impatto sull'accuratezza e la fluidità di interpreti in formazione, (b) la percezione della loro diversa usabilità, (c) la diversa distribuzione dell'attenzione sull'input visivo, (d) la modalità di utilizzo e (e) le potenziali prospettive didattiche.

A tal fine, è stata condotta un'indagine sperimentale in cui 11 partecipanti hanno interpretato in modalità simultanea due discorsi nelle rispettive condizioni sperimentali. I dati raccolti sono stati di sette tipi. Durante la simultanea sono stati infatti raccolti dati di *eye-tracking*. Dopo ogni simultanea è stata condotta un'attività di indagine retrospettiva ed è stato proposto un questionario proveniente dal campo di ricerca della *user experience* (UEQ) (Schrepp et al., 2017). Dopo aver trascritto le rese, queste sono state valutate globalmente, calcolando un punteggio relativo alla loro intelligibilità e informatività, seguendo la metodologia proposta da Tiselius (2009). Successivamente, sono stati calcolati i *task success rate* specifici per ogni *task* interpretato, seguendo i principi metodologici avanzati da Frittella (2023). Inoltre, dopo aver trascritto le retrospezioni, sono state analizzate le categorie emerse. I dati di *eye-tracking* sono stati invece analizzati in due modalità: (1) si è cercato di determinare mediante un'analisi quantitativa se nelle due condizioni sperimentali determinati stimoli fossero prevalentemente guardati in fase di percezione o di monitoraggio; (2) sono stati analizzati i valori relativi al *dwell time* nelle aree di interesse preconfigurate.

Le conclusioni tratte dalla ricerca presentata suggeriscono l'inevitabile necessità di una prospettiva didattica che permetta una gestione consapevole di un input visivo complesso, come quello presente nelle due tecnologie proposte.

Resumen

En el punto de contacto entre la inteligencia artificial y las intérpretes en formación, surge la necesidad de fomentar una perspectiva didáctica que permita, por una parte, una relación intérprete-máquina cada vez más ergonómica y, por otra, una fácil integración de los recién licenciados en el mercado. En este estudio, siguiendo un diseño cuasi-experimental y un enfoque de investigación de métodos mixtos, se comparan la herramienta CAI de SmarTerp y la transcripción automática de Google Meet, investigando los siguientes aspectos: (a) su diferente impacto en la precisión y fluidez de las intérpretes en formación, (b) la percepción de su diferente usabilidad, (c) la diferente distribución de la atención en el input visual, (d) el modo en que se utilizan y (e) las potenciales perspectivas docentes.

Para ello, se llevó a cabo una investigación experimental en la que 11 participantes interpretaron simultáneamente dos discursos en las respectivas condiciones experimentales. Los datos recogidos fueron de siete tipos. Se recogieron datos de seguimiento ocular durante la interpretación simultánea. Después de cada simultánea, se llevó a cabo una tarea retrospectiva y se propuso un cuestionario procedente del ámbito de la investigación de la experiencia del usuario (UEQ) (Schrepp et al., 2017). Tras la transcripción de los discursos interpretados, estos se evaluaron globalmente calculando una puntuación para su inteligibilidad e informatividad, siguiendo la metodología propuesta por Tiselius (2009). Posteriormente, se calcularon las tasas de éxito específicas de cada tarea interpretada, siguiendo los principios metodológicos avanzados por Frittella (2023).

Además, tras transcribir las retrospectivas, se analizaron las categorías que surgieron. Los datos de seguimiento ocular, por su parte, se analizaron de dos formas: (1) se intentó determinar mediante un análisis cuantitativo si en las dos condiciones experimentales determinados estímulos eran observados predominantemente en la fase de percepción o de seguimiento; (2) se analizaron los valores del tiempo de permanencia en las áreas de interés preconfiguradas.

Las conclusiones derivadas de la investigación presentada sugieren la inevitable necesidad de una perspectiva didáctica que permita el manejo consciente de un input visual complejo, como el presente en las dos tecnologías propuestas.

Abstract

At the point of contact between artificial intelligence and trainee interpreters, the need emerges to foster a teaching perspective that enables an increasingly ergonomic interpreter-machine relationship and an easy inclusion of new graduates into the market. In this study, following a quasi-experimental design and a mixed-method research approach, SmarTerp's CAI tool and Google Meet's automatic transcription are compared, investigating the following aspects: (a) their different impact on the accuracy and fluency of trainee interpreters, (b) the perception of their different usability, (c) the different distribution of attention on the visual input, (d) the way they are used and (e) the potential teaching perspectives.

To this end, an experimental investigation was conducted in which 11 participants simultaneously interpreted (SI) two speeches under the respective experimental conditions. The data collected were of seven types. Eye-tracking data were also collected during the simultaneous interpretation. After each SI task, a retrospective task was conducted and a questionnaire from the user experience research field (UEQ) was proposed (Schrepp et al., 2017). After transcribing the renditions, these were assessed globally by calculating a score for their intelligibility and informativeness, following the methodology proposed by Tiselius (2009). Subsequently, task-specific success rates were calculated for each interpreted task, following the methodological principles developed by Frittella (2023). After transcribing the retrospectives, the emerging categories were analyzed. On the other hand, the eye-tracking data were analyzed in two ways: (1) by means of a quantitative analysis, an attempt was made to determine whether in the two experimental conditions certain stimuli were predominantly watched in the perception or follow-up phase; (2) dwell time values were analyzed in the pre-configured areas of interest.

The conclusions drawn from the presented research suggest the inevitable need for a didactic perspective that allows for the conscious handling of complex visual input, such as that of two proposed technologies.

Introduzione

"Io sono io e la mia circostanza": con queste parole il filosofo spagnolo José Ortega y Gasset faceva riferimento alla forza che lega la nostra essenza a tutto ciò che la circonda, ossia a tutto l'ambiente attorno a noi che è ricco di limiti e libertà. La motivazione di questo studio viene proprio dal timore del ricercatore nei confronti dell'intelligenza artificiale e dalla volontà di accogliere questa paura e dialogarci. La scarsa consapevolezza dei limiti dell'intelligenza artificiale e la minaccia più volte sentita relativamente alla potenziale sostituzione dell'interprete da parte dei nuovi sistemi di intelligenza artificiale costituiscono infatti i due fattori alla base di questo timore. In un panorama sempre più promettente dal punto di vista tecnologico e in seguito alla costatazione dell'ottima qualità dei sistemi di riconoscimento vocale già esistenti, lo studio vuole essere un piccolo contributo circa la relazione tra interpreti e sistemi di intelligenza artificiale.

Seguendo un *design* quasi-sperimentale, lo studio si prefigge l'obiettivo di mettere in luce le differenze nella percezione dell'usabilità di due strumenti (la trascrizione automatica di Google Meet e il CAI *tool* di SmarTerp), cercando di evidenziarne la diversa utilità e i possibili impatti sull'accuratezza e la fluidità nelle rese di interpreti in formazione. Pertanto, utilizzando gli strumenti di valutazione della *user experience*, lo studio desidera approfondire le necessità di interpreti in formazione, con il desiderio di promuovere un'attitudine di dialogo tra interprete e intelligenza artificiale in cui quest'ultima possa amplificare la libertà propriamente umana dell'interprete. Prima di procedere con la sintesi del contenuto di questo elaborato, è necessario chiarire che i sottotitoli di Google Meet non sono pensati per offrire supporto all'interprete durante l'interpretazione simultanea, ma rappresentano uno strumento che potrebbe essere realisticamente utilizzato in questo contesto come CAI *tool*.

Il primo capitolo verte sull'inquadramento teorico dello studio. Dopo aver descritto brevemente il panorama attuale relativamente alle tecnologie per l'interpretazione (e in particolare per quanto riguarda i CAI *tool*), verrà proposto un approfondimento circa l'importanza della valutazione dell'usabilità per promuovere un rapporto uomo-macchina che sia ergonomico ed efficiente. Successivamente, verrà proposta una riflessione circa la complessità dei processi cognitivi in atto nel corso dell'interpretazione simultanea, cercando di osservare l'interferenza delle nuove tecnologie durante l'interpretazione simultanea. Verranno poi presentati i limiti e le potenzialità dei sistemi di riconoscimento vocale automatico, indagandone l'applicabilità durante l'IS. Terminata la panoramica generale, lo studio entrerà nei dettagli dei due sistemi oggetto dello studio sperimentale: il CAI *tool* SmarTerp e la trascrizione automatica di Google

Meet. Dopo aver illustrato l'architettura di base per entrambi i sistemi, lo studio affronterà una delle frontiere nel campo dell'interpretazione automatica: la sostituzione della voce umana tramite sistemi *speech-to-speech*. Terminata la descrizione delle potenzialità delle nuove tecnologie, verrà proposta una riflessione relativa all'utilizzo delle tecnologie di *eye-tracking* nel campo degli *interpreting studies*, fornendo alle lettrici alcuni strumenti e termini di base per poter comprendere meglio l'analisi dei dati relativi al movimento oculare che verranno presentati nel corso dello studio sperimentale. Dato che lo studio si concentra su interpreti in formazione, verrà proposto in seguito un breve *excursus* teorico relativamente alla didattica delle tecnologie nella formazione delle interpreti[1], osservando lo stato dell'arte e potenziali lacune. Inoltre, considerato che lo studio verrà proposto per la combinazione linguistica italiano-spagnolo, verrà proposto un approfondimento sulla particolarità di questa combinazione, con particolare enfasi sulle strutture dissimmetriche e le sfide lessicali.

Il secondo capitolo affronta la metodologia utilizzata nello studio sperimentale. Dopo aver presentato gli obiettivi e i quesiti di ricerca, verrà illustrato l'approccio sperimentale utilizzato, in funzione dei dati raccolti, congiuntamente al relativo *design* sperimentale. Verrà poi descritta la preparazione dei materiali sperimentali, con particolare enfasi sulle caratteristiche dei discorsi, la preparazione dei video, lo svolgimento delle retrospezioni e la modalità di raccolta dei dati di *eye-tracking*. Il capitolo presenterà successivamente il criterio di reclutamento delle partecipanti, la loro profilazione e la formazione asincrona che è stata erogata ad ogni partecipante prima dell'attività sperimentale. Terminata la descrizione delle caratteristiche delle partecipanti coinvolte, verrà sintetizzato il protocollo sperimentale. Dopo aver illustrato la preparazione dell'attività sperimentale e la modalità di svolgimento della stessa, verrà descritta sinteticamente la preparazione di tutti i dati per l'analisi. Al termine del capitolo, verrà brevemente descritto come il materiale sperimentale è stato validato nel corso di un esperimento pilota.

Il terzo capitolo presenta l'analisi e la discussione dei dati relativi all'accuratezza delle rese e all'*User Experience Questionnaire* (o UEQ). Per ogni tipologia, i dati verranno prima analizzati separatamente per le due condizioni sperimentali e, successivamente, verranno inseriti nell'ottica di un'analisi comparativa. Per quanto riguarda l'analisi relativa all'accuratezza delle rese, seguendo la metodologia proposta da Frittella (2023), verranno prima riportati i risultati dei *task success rate*, che mirano a valutare la percentuale di successo nella resa dei *task* inseriti

[1] Con il fine di perseguire l'impegno del ricercatore per la costruzione di un'università più inclusiva, nel corso dello studio sperimentale (e, quindi, a partire dal secondo capitolo) verrà utilizzato il femminile sovraesteso; i termini femminili usati in questo testo si riferiscono a tutte le persone.

strategicamente nei discorsi. Successivamente, tale metodologia verrà integrata con quella proposta da Tiselius (2009) e verrà proposta la valutazione dell'intelligibilità e dell'informatività di tutte le rese, che consente di avere una più ampia valutazione circa l'accuratezza informativa e la comprensibilità delle stesse. Sia per i *task success rate* che per l'intelligibilità e l'informatività delle rese, verrà proposta una breve analisi relativa alla significatività statistica delle differenze emerse nelle rispettive analisi comparative. Al termine del capitolo, verranno riportati i dati relativi alla percezione dell'usabilità degli strumenti, ottenuti tramite appositi questionari (Laugwitz et al., 2008).

Il quarto capitolo mira invece a illustrare l'analisi e la discussione dei dati di *eye-tracking* e delle attività di retrospezione. In particolare, nella prima parte verranno discussi i dati relativi al movimento oculare delle partecipanti in corrispondenza di stimoli predefiniti, in modo da investigare se gli strumenti siano stati utilizzati prevalentemente in fase di comprensione o in fase di monitoraggio della resa. Successivamente, verranno analizzati i dati quantitativi di *eye-tracking*, ossia quelli relativi al tempo trascorso dall'occhio delle partecipanti nelle singole aree di interesse preconfigurate, in modo da avere una comprensione più ampia circa la modalità con cui gli strumenti vengono utilizzati durante l'IS. Nella seconda parte del quarto capitolo verranno invece descritte qualitativamente le retrospezioni svolte dopo ogni prova di interpretazione simultanea; in particolare, verranno descritte in funzione delle categorie di analisi più frequentemente emerse dalle partecipanti: dalla percezione della latenza dei sistemi alla necessità di una prospettiva didattica.

Nelle conclusioni si discuteranno i dati emersi, nel tentativo di rispondere ai quesiti di ricerca e cercando di delineare possibili sviluppi futuri. Inoltre, le conclusioni presenteranno i limiti dell'indagine proposta nello studio.

CAPITOLO 1. Dalle tecnologie per l'interpretazione alle sfide della combinazione italiano-spagnolo

1.1 Le tecnologie per l'interpretazione: panorama attuale e orizzonti futuri

Il dinamico avvicendarsi delle innovazioni tecnologiche non può non avere influenze significative sull'attività dell'interprete, che vede sempre più il costante bisogno di formarsi per avvalersi delle tecnologie che vengono in suo aiuto. L'effetto dirompente delle tecnologie sulla professione dell'interprete non è certo nuovo (Fantinuoli, 2018a, p. 1); si possono infatti scorgere nel passato due innovazioni significative: (1) l'introduzione di sistemi cablati per la trasmissione del parlato (che hanno portato alla nascita dell'interpretazione simultanea) e (2) l'avvento di Internet (Fantinuoli, 2018a, p. 2). Queste due svolte hanno permesso la nascita e il consolidamento di modalità di interpretazione simultanea nuove. Infatti, anche se il processo cognitivo che prevedeva l'attività traduttiva in sovrapposizione all'attività dell'ascolto non era nuovo agli interpreti (in quanto era già presente l'interpretazione sussurrata, o *chuchotage*), il processo di Norimberga è stato un'eccellente cassa di risonanza che ha permesso di dare visibilità alla nuova tecnologia che prevedeva un'attrezzatura basata su sistemi cablati. Questo ha comportato un impatto significativo non solo sullo status sociale e sull'autopercezione degli interpreti, ma anche sulle modalità possibili di interpretazione simultanea, riducendo i tempi di realizzazione della prestazione per un numero molto più ampio di utenti, rispetto allo *chuchotage* (Fantinuoli, 2018a, p. 2). Come evidenzia Baigorri Jalón (2004), l'arrivo della nuova tecnologia ha comportato un prezzo da pagare per gli interpreti: la paura degli interpreti stessi di una perdita nella qualità della resa e il timore di un peggioramento del prestigio associato alla professione per lo spostamento dal palcoscenico, dove spesso figuravano a fianco di diplomatici, a cabine. Questi timori sembrano riecheggiare ancora oggi, quando la tecnologia entra in dialogo con la professione dell'interprete; tuttavia, nonostante le reticenze, l'avvento del *World Wide Web* riduce notevolmente costi e tempi dell'indispensabile *ad-hoc knowledge acquisition* (Gile, 2009, p. 144): la preparazione terminologica e il materiale per fare pratica per il miglioramento nel lungo termine della tecnica consentono una maggiore facilità di accesso alle fonti che porta a una riduzione della distanza linguistica e del divario di conoscenze tra interprete e oratore (Fantinuoli, 2018a, p. 2).

Le tecnologie possono mediare, generare oppure supportare l'interpretazione (Braun, 2019). Le tecnologie che mediano hanno portato l'interprete in *setting* precedentemente inaccessibili: nel

corso della conferenza dell'Organizzazione Internazionale del Lavoro del 1928, nuove tecnologie di mediazione vengono inaugurate e testate per fornire il servizio di interpretazione simultanea e, solo recentemente, hanno visto un terreno fertile le tecnologie per l'interpretazione da remoto, che consentono *set-up* tecnologicamente variegati (Ziegler & Gigliobianco, 2018, p. 121). Le tecnologie che invece generano interpretazione sono quelle che forniscono servizi di comunicazione interculturale senza il coinvolgimento diretto degli interpreti umani (Prandi, 2023, p. 27): il sopraggiungere di tecnologie avanzate nel campo del riconoscimento vocale, della trascrizione automatica (tramite tecnologie *speech-to-text*) e della sintesi vocale ha portato a una maggior forza di queste tecnologie. Le tecnologie che invece supportano l'interpretazione sono quegli *hardware* o *software* che gli interpreti utilizzano per far fronte a un incarico, ne sono esempi il tablet, la penna digitale (Orlando, 2010, p. 77) e i software utilizzati per la gestione terminologica (Prandi, 2023, p. 33). In questa svolta in cui le tecnologie entrano in contatto diretto con l'attività dell'interprete, è cruciale il ruolo dei motori di ricerca, il cui requisito funzionale comune è quello di fornire un servizio che comprenda le esigenze dell'utente, reagendo di conseguenza (Zavras & Finn, 2017).

Circoscrivere l'impatto dei dispositivi digitali e del Web a una maggior comodità dell'interprete sarebbe assai riduttivo; diventa dunque di prim'ordine mettere in luce le conseguenze su tutte le particolarità dei sotto-processi dell'attività di preparazione dell'interprete: comprensione, rievocazione e mantenimento in memoria delle conoscenze vedono cambiamenti dovuti anche alla diversa disposizione degli elementi paratestuali, ossia di ciò che complementa l'input verbale-auditivo (Ross et al., 2017). Quando osserviamo l'evoluzione del legame tra tecnologie e interpretazione, prevalgono tre campi di interesse: l'interpretazione assistita da computer (in inglese, *computer-assisted interpreting*, o CAI); l'interpretazione da remoto (*remote interpreting*, o RI) e l'interpretazione automatica (*machine interpreting*, o MI) (Fantinuoli, 2018a, p. 4).

L'interpretazione assistita da computer è definita da Fantinuoli (2018b, p. 155) come una modalità di traduzione orale in cui un interprete umano si avvale di un software progettato per supportare e facilitare alcuni aspetti dell'attività interpretativa per aumentare qualità e – in minor misura – produttività. Alleggerire i compiti più impegnativi nella fase di preparazione (quali la ricerca terminologica) e migliorare l'esperienza lavorativa dell'interprete durante l'attività interpretativa stessa diventano dunque i motivi evidenti alla base della creazione dei *software* per l'interpretazione assistita da computer (Fantinuoli, 2018a, p. 4).

L’interpretazione a distanza fa invece riferimento a un concetto ampio (e non monolitico) che viene frequentemente utilizzato per riferirsi a modalità di comunicazione mediata da un interprete e realizzata attraverso le tecnologie dell’informazione e della comunicazione in cui interprete e utenti non sono nella stessa sede (Fantinuoli, 2018a, p. 4). Nonostante le differenze specifiche dei *setting* dell’interpretazione di conferenza e dell’interpretazione dialogica, lo spettro delle applicazioni possibili dell’interpretazione a distanza sembra ampliarsi sempre di più: allo stato attuale vediamo un’applicazione che va dall’interpretazione telefonica all’interpretazione di conferenza. L’interpretazione automatica (anche nota come traduzione automatica del parlato, traduzione *speech-to-speech* o *machine interpreting*) fa riferimento alla tecnologia che consente la traduzione di testi parlati da una lingua all’altra per mezzo di un programma informatico (Fantinuoli, 2018a, p. 5).

Come chiarisce Fantinuoli (2018a, p. 5), se l’obiettivo dell’interpretazione assistita da computer e dell’interpretazione a distanza è quello di entrare in un dialogo creativo e proficuo con l’interprete umano, quello dell’interpretazione automatica mira a sostituirlo, considerato che la sua ambizione è combinare: il riconoscimento vocale automatico (o *automatic speech recognition*, o ASR) per trascrivere il discorso orale in testo scritto, la traduzione automatica (o *machine translation*, o MT) e la sintesi vocale (o *speech-to-text synthesis*, o STT) per generare una versione udibile nella lingua di arrivo (Fantinuoli, 2018a, p. 5); talvolta, la sintesi vocale in output potrebbe essere sostituita dalla creazione di sottotitoli in tempo reale (Prandi, 2023, p. 28).

Attualmente, i due modelli principali di MI sono i sistemi a cascata (o *cascading systems*) e quelli *end-to-end*. Nei sistemi a cascata, ASR e traduzione automatica vengono eseguiti consecutivamente e il parlato deve essere segmentato. I sistemi *end-to-end* non richiedono la fase intermedia dell'ASR. Quest'ultima possibilità ha iniziato a essere approfondita solo di recente, ma sta già raggiungendo la qualità dei sistemi a cascata (Prandi, 2023, p. 28). La rapida evoluzione del paradigma connessionista (noto anche come, *neural paradigm*, o come reti neurali artificiali), che è rilevante per tutte le componenti principali dell’interpretazione automatica, potrebbe portare a un ulteriore progresso notevole e rapido (Braun, 2019, p. 274); in particolare, quelle reti neurali artificiali che possono imparare da compiti eseguiti precedentemente e che possono spostare l’attenzione secondo la rilevanza di un elemento nel discorso originale potrebbero avere il potenziale di rendere l’interpretazione automatica più simile rispetto a quella umana (Seligman et al., 2017, p. 44). Prandi (2023, p. 29) afferma che (a differenza della comunicazione scritta) il discorso orale presenta un grado più alto di

spontaneità e ambiguità, che le macchine ancora non sono in grado di affrontare senza l'assistenza umana; inoltre, le macchine hanno notevoli limiti nell'anticipazione del contesto e nella capacità inferenziale.

Nella rassegna delle possibili conseguenze dovute alle svolte tecnologiche del progresso, la MI viene da più fonti dichiarata come una minaccia alla professione dell'interprete, ma Fantinuoli (2018a, p. 7) ritiene che, d'altro canto, l'introduzione dei CAI *tool* non sia da considerarsi problematica, in quanto con questi si prospetta un'applicazione mirata sui singoli micro-processi dell'attività interpretativa (Fantinuoli, 2018a, p. 7). Al contempo, la difficile prevedibilità nel lungo termine del progresso dell'interpretazione da remoto e dell'interpretazione automatica è motivo di non poca preoccupazione: le nuove opportunità lavorative in nuovi segmenti di mercato raggiungibili tramite RI potrebbero giungere mano nella mano con un deterioramento delle condizioni lavorative, in cui la mercificazione e la spersonalizzazione dell'interprete potrebbero aprire le porte a una spirale di declino economico e di de-professionalizzazione del mercato (Fantinuoli, 2018a, p. 7).

Prima di procedere con l'inquadramento teorico del presente studio sperimentale, è necessario chiarire quali tecnologie dell'informazione e della comunicazione (o ICT) verranno affrontate in questo elaborato. Fantinuoli (2018b) propone una classificazione delle ICT nel campo dell'interpretazione, in cui il livello di interazione dell'ICT con l'interprete (e, quindi, con l'intera attività interpretativa) rappresenta il fattore che distingue i due gruppi: il gruppo delle ICT *process-oriented* e quello delle ICT *setting-oriented*. Il primo comprende i sistemi di gestione della terminologia e i software per l'estrazione terminologica e l'analisi dei corpora: sono tecnologie *process-oriented* in quanto sono progettate per offrire supporto all'interprete durante i diversi sotto-processi dell'attività interpretativa (prima, durante e dopo l'incarico). Il secondo comprende i dispositivi, gli strumenti e i software che accompagnano il processo interpretativo, sono un esempio i dispositivi per l'interpretazione da remoto, le *console* all'interno delle cabine e le piattaforme per l'esercitazione (Fantinuoli, 2018b, p. 155). L'estrema generalizzazione della classificazione ci consente di considerare le tecnologie affrontate in questo lavoro come prevalentemente ICT *process-oriented*.

1.1.1 Interpretazione assistita da computer (o *Computer-Assisted Interpreting*)

Fantinuoli (2018b, p. 155) definisce i CAI *tool* come tutti i tipi di programmi per computer progettati e sviluppati appositamente per assistere gli interpreti in almeno uno dei diversi

sottoprocessi dell'interpretazione alleggerendone il peso in fase di preparazione (*pre-booth*), di interpretazione (*in-booth*) e di follow-up (*post-booth*), migliorandone produttività e qualità della *performance* (EABM Project, 2018). Come chiarito in §1.1, i CAI *tool* offrono un supporto all'interprete nel suo lavoro e non sostituiscono l'interprete umano. Lo sviluppo attuale dei *software* si concentra sulle due fasi *pre-booth* e *in-booth*, mentre la fase di *follow-up* è il risultato di quanto appreso dall'interprete durante le due fasi precedenti (Frittella, 2023, p. 47). Nella fase di *pre-booth*, i CAI *tool* intendono migliorare l'efficienza della preparazione degli interpreti, permettendo un'agevole raccolta terminologica e un'approfondita acquisizione delle conoscenze contestuali e contenutistiche utili ai fini dell'incarico, tramite la creazione e la gestione di glossari (Frittella, 2023, p. 47).

Nella fase *in-booth*, i CAI *tool* permettono di avere un accesso rapido alle informazioni raccolte in fase di preparazione (oppure ricorrendo a banche dati e dizionari *online*), per garantire una maggior accuratezza. La complessità dell'interpretazione simultanea (o IS) non può, tuttavia, essere ricondotta unicamente alla presenza di terminologia specialistica. Diversi possono essere i possibili *problem trigger* (Gile, 2009), elementi potenzialmente problematici che possono condurre a tassi di errore più alti rispetto all'andamento medio di una prestazione. Oltre alla terminologia specialistica, anche i numeri, gli acronimi e i nomi propri ne sono un esempio (Frittella, 2023, p. 48).

Architettura, funzionalità e obiettivi dei singoli CAI *tool* hanno portato a una classificazione in diversi gruppi, o meglio in diverse generazioni (Fantinuoli, 2018a). I CAI *tool* di prima generazione sono programmi progettati per gestire la terminologia in un modo che sia di facile intuizione per l'interprete (Fantinuoli, 2018b, p. 164). Questi CAI *tool* sono stati inizialmente progettati per aiutare l'interprete nella fase di raccolta terminologica, prevedendo anche un uso *in-booth* che permette una ricerca simile a quella che potrebbe essere una ricerca in un *file* di testo o di calcolo. Altri esempi di CAI *tool* di prima generazione sono i dizionari digitali, i motori di ricerca e i convertitori delle unità di misura.

I CAI *tool* di seconda generazione offrono un approccio più esteso relativamente alla terminologia e alle conoscenze per l'attività interpretativa e alle funzionalità dei CAI *tool* di prima generazione si aggiungono funzionalità avanzate che vanno oltre la gestione terminologica di base come, ad esempio, delle funzionalità per organizzare il materiale testuale e recuperare informazioni da corpora o altre risorse (Fantinuoli, 2018b, p. 165). Questi CAI *tool* mirano a garantire un miglioramento dell'attività *in-booth*, ad esempio rendendo più agevole, dinamica e rapida la ricerca terminologica. Una funzionalità ricorrente è, ad esempio,

la modalità *progressive search* con cui automaticamente il CAI *tool* interroga altri glossari dell'interprete se il termine non viene trovato nel glossario che si sta interrogando, quando in modalità *in-booth* (Fantinuoli, 2018b, p. 166). Alcuni esempi di CAI *tool* di seconda generazione sono quelli utilizzati per la gestione terminologica basati su sistemi di riconoscimento vocale automatico.

L'introduzione dell'intelligenza artificiale nei CAI *tool* ha aperto la strada anche a una terza generazione (EABM Project, 2018). In questa generazione, la fase *pre-booth* vede un'automatizzazione che rende possibile la compilazione automatica di glossari partendo dal materiale che l'interprete utilizza in fase di preparazione. In secondo luogo, il riconoscimento vocale (noto anche come riconoscimento vocale automatico, *automatic speech recognition*, o ASR) viene combinato con l'intelligenza artificiale (IA), permettendo dunque la ricerca automatica nel glossario e la risoluzione di molti dei *problem trigger* presenti nel discorso: numeri, nomi propri, acronimi e tecnicismi vengono dunque riconosciuti dell'IA, automaticamente rinvenuti nei glossari a disposizione e proiettati tradotti sullo schermo dell'interprete (Frittella, 2023, p. 49).

Frittella (2023, p. 49) mette in luce i due limiti più evidenti di questi CAI *tool*: la latenza (ossia il ritardo da quando viene pronunciata la parola a quando compare sullo schermo dell'interprete) e l'accuratezza (spesso minata da tassi di errore molto elevati). La causa di questi due limiti può essere rinvenuta nel fatto che i CAI *tool* di terza generazione sono CAI basati su sistemi a cascata basati su due moduli di ASR e IA indipendenti. Lavorano quindi su due fasi temporalmente distinte: riconoscimento/trascrizione del discorso originale ed estrazione dei *problem trigger*/conversione in una rappresentazione grafica sullo schermo dell'interprete (Frittella, 2023, p. 50).

Latenza e inaccuratezza potrebbero, tuttavia, essere superati dai CAI *tool* di *quarta generazione*, ossia da CAI *tool* basati su sistemi *end-to-end* (Gaido et al., 2021). Onde evitare di confondere CAI di potenziale quarta generazione con eventuali sistemi di interpretazione automatica (o MI), Feng (2018) chiarisce che la differenza tra CAI e MI risiede nel fatto che, nel caso dei CAI, la traduzione automatica viene utilizzata come riferimento da parte degli interpreti umani, mentre nella MI la traduzione automatica viene riprodotta tramite sintesi vocale.

Nel tentativo di giungere a una definizione dei CAI *tool* circoscritta e mirata, che metta in evidenza le principali caratteristiche di essi, Guo et al. (2023) offrono una definizione

interessante, che vede intersecarsi tre caratteristiche: l'immediatezza, l'ampiezza dell'ambito di applicazione e l'influenza sui processi cognitivi.

La prima caratteristica è dovuta al fatto che l'interpretazione è una forma di traduzione in cui viene prodotta una resa in un'altra lingua sulla base della presentazione di un enunciato in una lingua di partenza ed è pertanto un'attività traduttiva in tempo reale finalizzata alla fruizione immediata (Pöchhacker, 2016, p. 10,11). La seconda caratteristica si deve al fatto che qualsiasi *software*, applicazione per dispositivo cellulare o dispositivo digitale può essere utilizzato come CAI *tool* (Guo et al., 2023, p. 91). Per quanto riguarda, invece, l'influenza sui processi cognitivi, essa si deve alla multidimensionalità delle ripercussioni dei CAI *tool* sui processi cognitivi dell'interprete, talora riducendo lo stress cognitivo. La complessità di questi impatti verrà affrontata più avanti nel corso di questo lavoro.

1.1.1.1 Progettazione, sviluppo e ricezione dei CAI *tool*

Passando in rassegna i CAI *tool* che sono stati sviluppati nel corso degli anni, citiamo *software* di prima generazione, quali Interplex[2], Terminus, Interpreters' Help[3], LookUp[4] e DolTerm; anche se solo Interplex e Interpreters'Help sono attualmente disponibili in commercio. Questi CAI *tool* sono stati sviluppati per stoccare e raccogliere dati terminologici da *database* e si differenziano rispetto ai sistemi di gestione terminologica tendenzialmente utilizzati dai traduttori in quanto hanno una funzionalità dedicata per la modalità *in-booth*; tuttavia, nessuno di questi CAI presenta un algoritmo di ricerca avanzato che prende in considerazione i limiti temporali dell'attività interpretativa, ad esempio, relativamente alla correzioni degli errori di pronuncia o alla modalità *progressive search* per svolgere una ricerca terminologica in uno o più glossari contemporaneamente (Fantinuoli, 2016, p. 44). Tra i CAI di seconda generazione troviamo Intragloss e InterpretBank[5]. Intragloss è prevalentemente mirato alla fase di preparazione dell'incarico, con la particolarità di essere basato sull'interazione tra i testi utilizzati nella fase di *pre-booth* e il *database* terminologico: una volta individuato il termine in un documento, cerca la traduzione su risorse *online* come glossari, *database* o dizionari (Fantinuoli, 2016, p. 44). InterpretBank è stato inizialmente lanciato sul mercato tra i CAI *tool*

[2] https://interplex.com/ (Ultimo accesso: 29 aprile 2024)
[3] https://interpretershelp.com/ (Ultimo accesso: 29 aprile 2024)
[4] https://www.lookup-web.de/index.php (Ultimo accesso: 29 aprile 2024)
[5] https://interpretbank.com/site/ (Ultimo accesso: 29 aprile 2024)

di seconda generazione, ma grazie alle costanti integrazioni oggi si trova tra i CAI *tool* di terza generazione. Ad oggi, i principali CAI *tool* di terza generazione sul mercato sono InterpretBank ASR, InterpreterAssist di Kudo[6] e SmarTerp[7]. Tra questi, è proprio InterpretBank (Fantinuoli, 2017) che ha aperto la strada a questa nuova generazione: in esso ASR e tecnologia basata su IA sono combinati per interrogare automaticamente i glossari e mostrare i numerali. Al momento, gli interpreti possono utilizzare il CAI *tool* di terza generazione di InterpretBank in due modalità. La modalità *Digital Boothmate* è una funzione basata su *cloud* che cerca automaticamente ed evidenzia la terminologia, i nomi propri e i numeri mentre si interpreta simultaneamente in cabina, ascoltando l'oratore originale. Si crea un glossario in *InterpretBank Desktop* e dall'installazione del *Desktop* si può avviare una sessione.

Nella seconda modalità (*Artificial Boothmate*), la terminologia specialistica e i numerali vengono organizzati in due diverse colonne e ogni nuovo *item* viene evidenziato non appena compare sullo schermo. In questo caso, il CAI *tool* è stato adattato appositamente per l'interpretazione consecutiva, con l'obiettivo di migliorare efficienza e accuratezza. Il *tool* vede l'integrazione della trascrizione automatica sullo schermo dell'interprete. Nello spazio a destra dello schermo, l'interprete può prendere le proprie note di consecutiva (concetti, parole chiave) per rafforzare la comprensione e in vista di un'interpretazione coerente e adeguata del messaggio originale.

Per quel che concerne Kudo InterpreterAssist, si tratta di un CAI *tool* che comprende due funzioni principali, applicabili in contesti di interpretazioni simultanee da remoto: uno strumento di creazione automatica del glossario e un sistema di suggerimenti in tempo reale per la terminologia specialistica, i numeri e i nomi propri. Questo CAI *tool* consente la suddivisione del lavoro in quattro fasi principali: (1) la creazione di un progetto mirato al tema dell'evento; (2) la valutazione facoltativa delle risorse multilingue automaticamente generate; (3) la condivisione di risorse con i colleghi e (4) l'interpretazione su piattaforma RSI con i suggerimenti in tempo reale (Fantinuoli et al., 2022).

Considerati gli sviluppi recenti nei CAI *tool*, è indispensabile prendere atto – tuttavia – della limitata ricezione, spesso circoscritta a singoli interpreti che per interesse personale hanno utilizzato uno di questi CAI *tool*. Frittella (2023, p. 53) riferisce come, al momento della stesura del suo contributo, non risultassero adozioni di CAI *tool* da parte di agenzie e istituzioni e ritiene che un passo decisivo per un'adozione quantitativamente rilevante sarà dettato

[6] https://kudoway.com (Ultimo accesso: 29 aprile 2024)
[7] https://smarterp.me (Ultimo accesso: 29 aprile 2024)

dall'integrazione dei CAI *tool* all'interno delle piattaforme di RSI. La ricezione dei CAI tool ha spesso avuto reazioni contrastanti da parte degli interpreti: inevitabilmente, l'introduzione di nuove tecnologie nel processo interpretativo comporta nuovi poli di interesse nella complessità dell'attività dell'interpretazione per la presenza di nuove fonti di input visivo, andando potenzialmente ad avere un impatto sul carico cognitivo stesso (Frittella, 2023, p. 52). Jones (2014) mette in luce la potenziale alienazione conseguente all'introduzione delle nuove tecnologie, che potrebbe far perdere di vista il compito primario dell'interprete: trasferire il significato e agevolare la comunicazione.

Un'altra preoccupazione legata alla ricezione da parte degli interpreti è di natura prevalentemente economica: l'aumento dei CAI *tool* sul mercato potrebbe portare a una diminuzione del costo dei servizi di interpretazione (Frittella, 2023, p. 53). L'ultima preoccupazione mossa dagli interpreti nella fase di ricezione dei CAI *tool* è legata alla paura che questi possano sostituire gli interpreti umani stessi (Frittella, 2023, p. 53): una preoccupazione legata in generale all'IA. Dinanzi a questa reticenza, è sufficiente ribadire la differenza tra quelle tecnologie che vogliono sostituire gli esseri umani e quelle che vogliono offrire loro un supporto (e tra quest'ultime rientrano i CAI *tool*). Christian (2011, p. 70) ricorda che è proprio questa tensione costante tre essere umano e macchina che è il punto controverso; o, per meglio dire, la collaborazione, il dialogo, il duetto.

Frittella (2023, p. 54) ipotizza tra i motivi della scarsa ricezione il *design* dell'interfaccia utente, che nei diversi CAI *tool* è stato spesso configurato a partire dalle idee dei singoli sviluppatori, i quali spesso si sono trovati a prendere decisioni circa la disposizione dei diversi elementi dell'interfaccia utente seguendo la loro sola intuizione e un'euristica oscura e vaga, priva di studi sperimentali che ne rafforzassero la validità (Frittella, 2023, p. 54). Se è vero che Defrancq & Fantinuoli (2021, p. 4) affermano la necessità di un'interfaccia ergonomicamente adeguata all'interprete, è altrettanto vero che una definizione di ergonomia (specifica per l'interfaccia utente dei CAI *tool*) non viene fornita tramite studi specifici in questo campo, a causa della novità portata dagli stessi CAI *tool*. Come Fantinuoli (2018b, p. 164) suggerisce, la scarsa diffusione e il poco tempo trascorso dall'arrivo dei CAI *tool* sono stati fattori cruciali nel determinare la scarsa maturità delle prime proposte di *design* dell'interfaccia utente. Al momento, le funzionalità presentate si basano più sulle abitudini e sulle idee dello sviluppatore-interprete che sui bisogni della comunità di professionisti.

1.1.2 La chiave per l'ergonomia nel rapporto uomo-macchina: tra usabilità e *usability engineering*

Con il concetto di usabilità (o *usability*, in inglese), si intende il grado con cui un prodotto (inteso come sistema interattivo) può essere utilizzato da utenti specifici per raggiungere – in un contesto d'uso ristretto - determinati obiettivi con efficacia, efficienza, soddisfazione (Norm, 2018). Nello studio del rapporto uomo-macchina (o *human-computer interaction* o *HCI*) si distinguono due branche di ricerca: da una parte una ricerca focalizzata sulla costruzione dell'interfaccia della tecnologia in questione (*technical HCI research*) e una focalizzata sulle priorità di carattere cognitivo (*behavioural HCI research*) (Lazar et al., 2017, p. 10). Dinanzi alle svolte tecnologiche recenti, la facilità nell'uso degli strumenti è diventata una prerogativa indifferibile, data la fruizione delle stesse tecnologie da parte di esperti di qualsiasi settore, pena il mancato uso o l'inutilità della tecnologia stessa (Lazar et al., 2017, p. 2). Nella letteratura accademica, si possono scorgere diverse definizioni di *usabilità* e diverse modalità per misurarla, portando a un concetto fluido e multisettoriale. L'approccio definitorio orientato al prodotto (*product-oriented view*) definisce l'usabilità come l'insieme degli *attributi* che rendono ergonomico un prodotto (Bevan et al., 1991), con una misurazione conseguente che prevede l'analisi quantitativa e qualitativa dell'insieme di questi attributi. Secondo l'approccio orientato all'utente (*user-oriented view*), invece, la misurazione dell'usabilità prende in considerazione lo sforzo mentale e l'attitudine dell'utente (Bevan et al., 1991). Nell'approccio che invece si focalizza sulla *performance* dell'utente (*user performance view*), accettabilità da parte dell'utente e facilità d'uso sono i due parametri chiave (Frittella, 2023, p. 9). Si aggiunge a questi approcci quello orientato al contesto (*contextually-oriented view*), secondo il quale le particolarità del compito da eseguire, la caratterizzazione del bacino utenti e l'ambiente di lavoro sono elementi aggiunti all'equazione (Frittella, 2023, p. 9). La definizione risultante è legata all'*utilità*, l'*accettabilità* e l'*utilizzabilità* di un prodotto: accettabilità pratica (costo adeguato, affidabilità, compatibilità con i sistemi attuali) e sociale (sicurezza ed etica non in collisione) configurano l'accettabilità del prodotto; l'utilizzabilità di un prodotto descrive invece la capacità del prodotto di raggiungere obiettivi specifici; l'utilità descrive se la funzionalità del sistema in linea di principio è in grado di fare ciò che è necessario e, quindi, ciò per cui quel sistema è stato ideato (Nielsen, 2010).

Nielsen (2010) ritiene che l'usabilità di un prodotto sia determinata da cinque elementi: l'apprendibilità (l'utente dovrebbe poter imparare a utilizzare il prodotto in tempi rapidi),

l'efficienza, la memorizzabilità (le funzionalità del sistema dovrebbero essere facili da ricordare), la scarsa quantità di errori e la soddisfazione dell'utente finale. La norma ISO 9251-11 (2018) sull'ergonomia dell'interazione uomo-macchina salda insieme i diversi approcci finora descritti, affermando che l'usabilità può essere definita come la misura con cui un prodotto può essere utilizzato da determinati utenti per raggiungere determinati obiettivi con efficacia, efficienza e soddisfazione in un determinato contesto d'uso. È curioso, dunque, l'inserimento della soddisfazione soggettiva degli utenti all'interno di una definizione che vuole uscire dallo schema utilitaristico (e quindi dalla volontà di sottoporre i prodotti a una prototipazione basata esclusivamente sul fine di soddisfare le esigenze degli sviluppatori e non degli utenti) andando incontro alla complessità degli obiettivi anche di carattere sociale ed emotivo (Bevan et al., 2016). L'inclusione di questa dimensione pragmatica arricchente e più lontana dalla sola tecnicità della configurazione spaziale ha portato all'introduzione del concetto della *user experience* (o UX) (Frittella, 2023, p. 12); con esso si vuole descrivere l'insieme delle percezioni e delle risposte risultanti dall'uso di un prodotto, un sistema o un servizio (ISO, 2018) e alcuni autori inseriscono l'usabilità di un prodotto all'interno della UX, insieme all'utilità, come approfondito da Hassenzahl (2003).

Come approfondiscono Whiteside et al. (1988), l'approccio che maggior consenso genera oggi all'interno della comunità accademica è quello della *usability engineering*, con cui si vuole migliorare il *design* del prodotto già in fase di prototipazione, per evitare un aumento dei costi che scoraggia l'utente finale a un eventuale miglioramento del *design* (Nielsen, 2012, p. 13). Al tempo stesso, esiste la consapevolezza dell'assenza di linee guida e principi generali validi in qualsiasi condizione (Mayhew, 2007, p. 918).

Nielsen (1993, p. 11 - 15) sintetizza la necessità di un approccio che faccia leva sul miglioramento dell'usabilità in fase di prototipazione proponendo sei slogan: (1) *Your Best Guess Is Not Good Enough*, cioè deve essere sempre lasciata aperta la possibilità di modificare l'interfaccia, anche a prodotto ultimato; (2) *The User Is Always Right,* ossia i problemi con l'interfaccia non possono essere imputati all'utente; (3) *The User Is Not Always Right*, cioè l'utente non necessariamente è consapevole di ciò di cui ha bisogno, con la relativa prevedibilità delle conseguenze; (4) *Users Are Not Designers*, cioè gli utenti non sono esperti di personalizzazione del *design*; (5) *Designers Are Not Users*, cioè il divario di conoscenze tra sviluppatori e utenti deve essere sempre preso in considerazione; (6) *Details Matter*, cioè anche gli elementi più piccoli dell'interfaccia potrebbero avere conseguenze sull'usabilità della stessa.

Per poter procedere a una valutazione empirica dell'usabilità di un prodotto che consenta un'ottimizzazione del *design* e la correzione di problemi legati all'usabilità, è fondamentale che il sistema venga usato da utenti reali rappresentativi del gruppo *target* (sia in termini demografici sia in termini di dominio di interesse) (Frittella, 2023, p. 17). Il *setting* per il test potrebbe essere artificiale o realistico, se il prodotto si trova nella fase finale di sviluppo o implementazione (Frittella, 2023, p. 17). Il test potrebbe raccogliere dati sia relativi alla *performance* degli utenti sia relativi alla *percezione* del prodotto da parte degli stessi (Frittella, 2023, p. 18). Il rapporto tra i due fattori non è necessariamente direttamente proporzionale: una maggior accuratezza, ad esempio, potrebbe richiedere un prezzo da pagare in termini di sforzi cognitivi e stress psico-fisico. L'analisi della *performance* può focalizzarsi su principalmente due aspetti: (1) la misurazione quantitativa del successo nello svolgimento degli obiettivi dei singoli *task* e (2) l'identificazione di eventuali segnali di confusione o intralci (Lewis, 2012, p. 1). I dati relativi alla percezione degli utenti possono essere raccolti tramite questionari *post-task* oppure tramite intervista con domande aperte o chiuse. Esempi di questionari utilizzati per avere una descrizione quantitativa della percezione degli utenti sono: *User Experience Questionnaire* (UEQ) (Laugwitz et al., 2008), *Systems Usability Scale* (SUS) (Brooke, 1996) e *NASA Task Load Index* (NASA TLX) (Hart & Staveland, 1988).

La valutazione complessiva dell'usabilità prevede dunque un'osservazione sistematica degli utenti in un ambiente controllato per verificare in che misura il prodotto può essere utilizzato (Seewald & Hassenzahl, 2004): analisi della *performance* e dati sulla percezione degli utenti saranno entrambi indispensabili per valutare l'interazione con il prodotto. Se è vero che i protocolli verbali *think-aloud* e quelli retrospettivi (di carattere osservazionale e qualitativo) aiutano ad approfondire i processi cognitivi inerenti all'esecuzione del *task* (Frittella, 2023, p. 20), è altrettanto vero che la valutazione quantitativa (ad esempio, dei tassi di errore nell'esecuzione dei *task*) può favorire un approccio misto che offre una visione più ampia sull'usabilità del prodotto (Frittella, 2023, p. 20). In fase di valutazione, la comparsa di problemi nell'interazione utente-prodotto è definita *critical incident* (Seewald & Hassenzahl, 2004): si verifica dunque una divergenza tra attività dell'utente e obiettivo. Il *critical incident* può essere esplicito quando, durante l'esecuzione del *task*, l'utente presenta espressioni facciali o reazioni verbali che evidenziano un conflitto rispetto all'obiettivo preposto a un'interazione efficace ed efficiente (Frittella, 2023, p. 21). Se una pluralità di *critical incident* viene ricondotta a un elemento specifico del *design*, ecco che si giunge all'identificazione di problemi d'uso (*usage problems*) (Seewald & Hassenzahl, 2004). L'intero processo di valutazione dell'usabilità di un prodotto prevede quindi: una raccolta dati (*experience*), un'analisi dei dati (*data analysis*),

l'interpretazione mirata all'identificazione di possibili problemi d'uso (*interpretation*), la definizione delle priorità e della frequenza dei problemi (*prioritisation*) e le raccomandazioni sul come migliorare il prodotto (Seewald & Hassenzahl, 2004).

In conclusione, la valutazione dell'usabilità prevede tecniche osservazionali e necessità di dati che permettano di osservare l'interazione utente-prodotto da più punti di vista; questo implica che gli studi richiedano tempi molto lunghi e – pertanto – con un limitato numero di partecipanti disponibili. Hwang & Salvendy (2010) suggeriscono per una valutazione quantitativo-qualitativa un campione ideale minimo di 10 ± 2 partecipanti. Nielsen & Landauer (1993) affermano che un campione di circa cinque partecipanti potrebbe essere sufficiente se l'obiettivo è quello di ottenere un *testing* qualitativo sull'usabilità di un prodotto, tramite protocolli verbali sufficientemente approfonditi. Questo non è altrettanto vero per valutazioni che prevedano l'uso di metriche quantitative, che richiedono in molti casi un campione minimo di 30 partecipanti (Budiu & Moran, 2021).

1.2 L'interpretazione simultanea come attività cognitivamente complessa

Come accennato precedentemente, l'uso delle tecnologie durante un incarico di simultanea si configura quale ulteriore elemento che potrebbe alterare i processi cognitivi *co-occorrenti* che si presentano durante l'incarico. La psicologia dell'apprendimento e la psicologia cognitiva hanno messo in risalto alcuni fattori che potrebbero motivare interferenze significative durante l'utilizzo di un CAI in fase *in-booth*: l'interferenza tra i *task*, l'accesso a un bacino di risorse cognitive limitate, il problema per cui l'attenzione deve essere rivolta a diversi *task* concomitanti in un'estensione di tempo limitata richiedendo un acuto monitoraggio e coordinamento del carico cognitivo (*cognitive load*, o CL) (Prandi, 2023, p. 63). La memoria di lavoro (o *working memory*, o ML) e l'orientamento dell'attenzione (o *attention allocation*) diventano il punto di partenza per un'analisi più ampia delle attività cognitive concomitanti durante l'IS. Di seguito, viene proposta una descrizione della complessità dell'IS in quanto attività cognitivamente complessa, seguendo l'excursus teorico di Prandi (2023).

La ricerca sulla ML mira alla comprensione del ruolo della memoria nell'elaborazione delle informazioni da parte dell'essere umano (ad esempio, approfondendo i meccanismi sottostanti al ragionamento, la comprensione del linguaggio e l'apprendimento) (Baddeley & Hitch, 1974).

Baddeley & Hitch (1974) hanno proposto l'idea di un sistema di ML responsabile sia del trattenimento delle informazioni nel breve termine, così come anche della loro elaborazione, descrivendo la ML come un sistema di controllo con limiti sia sulle capacità di stoccaggio delle informazioni così come anche sulle capacità di elaborazione delle stesse (Baddeley & Hitch, 1974, p. 86). La ML è dunque particolarmente rilevante nell'IS, in quanto osserviamo in essa la necessità di trattenere informazioni che per loro natura sono transitorie così come il bisogno di elaborarle in tempi brevi (Prandi, 2023). I CAI *tool* potrebbero quindi portare gli interpreti a non avere le risorse necessarie o il tempo sufficiente per far fronte ai sotto-processi aggiuntivi comportati dall'interazione con un sistema di ASR (o di altro tipo) (Prandi, 2023, p. 64), in quanto la ML potrebbe non essere in grado di trattenere ed elaborare tutti gli stimoli. L'architettura della ML sembra ancora oggetto di grande dibattito, ma Baddeley & Hitch (1974) offrono un sistema in cui viene descritta come segue: l'informazione verbale viene codificata foneticamente (attraverso un *phonological loop*, PL) per poi essere ripetuta (mentalmente o ad alta voce) in un *buffer fonetico* a capacità limitata (Prandi, 2023, p. 65). Questa ripetizione è fondamentale per il trattenimento dell'informazione. La complessità del modello proposto della ML aumenta considerevolmente a causa della presenza degli elementi visivi-spaziali: la ML ha infatti un sistema di codifica e stoccaggio degli elementi visivo-spaziali (noto come *visual-spatial sketch pad*, o VSSP) che differisce rispetto al PL. Il PL è a sua volta diviso in una componente preposta allo stoccaggio fonologico e un'altra costituita da un sottosistema volto alla ripetizione e alla rievocazione delle tracce mnemoniche per evitarne il decadimento (Prandi, 2023, p. 65). Anche il VSSP è stato diviso in un sottosistema volto al trattenimento dell'input visivo e un altro volto alla sua elaborazione (Logie, 1995). PL e VSSP vengono gestiti e controllati nelle loro funzioni da una Centrale Esecutiva (CE). Successivamente, Baddeley (2000) ha introdotto il concetto di un *Episodic Buffer* che stocca le informazioni integrando il PL e il VSSP con la memoria a lungo termine (o long-term memory, o LTM). Nel modello proposto, trattenimento mnemonico ed elaborazione sono dunque gestiti da due strutture mnemoniche diverse e questo rafforza l'idea dell'IS come un'attività di *multitasking* (Prandi, 2023).

La doppia funzione del PL è stata empiricamente dimostrata più volte; ad esempio, con il *word-length effect*, per cui osserviamo un deterioramento delle tracce mnemoniche quando assistiamo alla presentazione di parole che comportano una soppressione della ripetizione articolatoria (Jacquemot et al., 2011). Nel modello proposto, la Centrale Esecutiva è responsabile della divisione, dello *switching* e dell'orientamento dell'attenzione, nonché dell'interazione con la LTM (Prandi, 2023, p. 65). Altri studiosi propongono modelli di memoria a breve termine (o

short-term memory, STM), in cui questa viene definita quale porzione attivata della memoria a lungo termine (Cowan, 1988), giungendo a una visione unitaria dell'architettura cognitiva.

La Centrale Esecutiva (CE) è cruciale nell'alleggerimento del carico cognitivo ed è volta a evitare un sovraccarico cognitivo laddove la capacità della memoria a breve termine venga superata (Prandi, 2023, p. 67); dunque, la CE interviene ricodificando le informazioni per alleggerire il carico cognitivo, ricombinando i singoli elementi in segmenti informativi (o *chunk*) più grandi e quindi riducendo il carico sulla memoria a breve termine e, conseguentemente, della ML (Prandi, 2023, p. 67).

L'interpretazione è un'attività che richiede l'allocazione dell'attenzione su molteplici flussi informativi (Seeber, 2017), richiedendo processi di orientamento dell'attenzione *bottom-up* e *top-down* (Prandi, 2023). In questo contesto, l'IS assistita da strumenti tecnologici di vario tipo potrebbe imporre un alto carico percettivo sulla ML dell'interprete (Lavie, 1995). Secondo Lavie (1995), l'allocazione dell'attenzione dipende dal carico percettivo posto dal *task*: in condizioni a basso carico, più stimoli possono essere elaborati per poi essere successivamente selezionati; mentre nelle situazioni ad alto carico percettivo, solo una limitata capacità di allocazione dell'attenzione viene lasciata al processo di elaborazione e il materiale informativo viene prima percepito ed elaborato e, solo successivamente, selezionato (Prandi, 2023, p. 70). Pertanto, possiamo ipotizzare che in condizioni di IS con supporto tecnologico, se quest'ultimo comporta un maggior carico percettivo, l'elaborazione di altri stimoli avrà a disposizione una minor quantità di risorse (Prandi, 2023, p. 10).

L'allocazione dell'attenzione deve dunque affrontare la costante condivisione di risorse per *task* concomitanti. Wickens (2002) propone un modello in cui afferma che la condivisione temporale tra due *task* è più efficiente se quei *task* utilizzano strutture separate (e non comuni). Il modello si basa prevalentemente su tre premesse: (1) ogni *task* non automatizzato produce un carico cognitivo; (2) due *task* concomitanti richiedono più risorse rispetto all'esecuzione di un solo *task*; (3) i *task* che richiedono lo stesso tipo di risorse mostrano un maggior livello di interferenza rispetto a *task* che richiedono risorse di strutture separate. Da questo modello strutturale dell'attenzione ne consegue che l'esecuzione simultanea di due (o più) *task* che richiedono le stesse risorse potrebbe essere estremamente difficile; questo spiega il motivo per cui l'esecuzione simultanea di un *task* visivo e uno uditivo è più semplice rispetto all'esecuzione di due *task* uditivi (Prandi, 2023, p. 72). Wickens (2002) identifica dunque quattro dimensioni (stadi dell'elaborazione, modalità percettive, codici per l'elaborazione ed elaborazione visiva) dividendole a loro volta in due livelli. I due livelli delle fasi dell'elaborazione sono la percezione

(*perception/cognition*) e la risposta (*response*); le due modalità percettive sono quella visiva (*visual*) da un lato e quella uditiva (*auditory*) dall'altro; i codici per l'elaborazione delle informazioni sono quello spaziale (*spatial*) e quello verbale (*verbal*); l'elaborazione visiva ha un livello ambientale (*ambient*) e uno focale (*focal*) (Wickens, 2002). Wickens (2002) avanza l'idea di un *conflict matrix* (o CM) per mostrare il livello di sfruttamento delle risorse da parte di ogni sub-*task* e il grado di interferenza dei sub-*task* co-occorrenti, che presenta valori numerici più alti se, ad esempio, due compiti richiedono entrambi lo stesso livello di una determinata dimensione.

L'allocazione dell'attenzione e la memoria di lavoro sono nozioni cruciali nella descrizione dell'interpretazione simultanea come attività cognitivamente complessa, ma diventa adesso necessario un approfondimento del carico cognitivo, per vedere come questo possa essere un fattore fondamentale nel rapporto computer-interprete. Il carico cognitivo intrinseco (o *intrinsic cognitive load*, o CCI) è il carico cognitivo che sperimenta il soggetto durante la fase di elaborazione: esso dipende dalla quantità di elementi elaborati simultaneamente nella memoria di lavoro e dall'interattività dei *task* che devono essere eseguiti (Merrienboer & Sweller, 2005, p. 150); pertanto, CCI e carico cognitivo sono direttamente proporzionali (Prandi, 2023, p. 76). Quando assistiamo a un input multimodale, ad esempio con sovrapposizione di IS e CAI *tool* (o altre tecnologie), il carico intrinseco potrebbe aumentare notevolmente, ma con l'esperienza può diminuire (Prandi, 2023, p. 76). Il *carico cognitivo estraneo* (o *extraneous cognitive load*, o CCE) fa riferimento alla configurazione del *task* e la sua presentazione; esso è dunque direttamente proporzionale alla complessità dell'organizzazione visiva dell'informazione (ad esempio, all'interfaccia di un CAI *tool* o di un'altra tecnologia) (Prandi, 2023, p. 76). Il carico cognitivo pertinente (o *germane load*, o CG) è invece lo sforzo mentale che il soggetto deve fare per eseguire un *task* determinato (Prandi, 2023, p. 77): dipende quindi dal soggetto e può diminuire con l'esperienza.

L'organizzazione visiva dell'informazione ha quindi un impatto sostanziale sul carico cognitivo; pertanto, è necessario procedere all'inquadramento teorico di tre effetti: l'effetto legato alla divisione dell'attenzione (o *split-attention effect*, o SAF); l'effetto legato alla ridondanza (o *redundancy effect*, o RE) e l'effetto legato alla modalità (o *modality effect*, o ME) (Prandi, 2023, p. 77). L'effetto legato alla divisione dell'attenzione dimostra che un *design* integrato è più efficace e funzionale rispetto a un *design* che preveda elementi spazialmente separati (Tarmizi & Sweller, 1988); tuttavia, la contiguità spaziale deve essere accompagnata da una contiguità temporale (e quindi da una bassa latenza dei sistemi di ASR (Montecchio,

2021). Per quanto riguarda l'effetto legato alla ridondanza, Chandler & Sweller (1991) affermano che avere un'informazione integrata e ridondante pone un carico innecessario sulla memoria di lavoro minando l'elaborazione delle informazioni, in quanto parte della capacità di elaborazione viene riservata all'integrazione mentale (ossia, alla comprensione) di fonti informative ridondanti (Prandi, 2023, p. 78). Il mancato controllo diretto dell'ASR da parte dell'interprete porta talvolta alla presentazione su schermo di terminologia già presente nella memoria a lungo termine dell'interprete (e, pertanto, ridondante) (Prandi, 2023, p. 78). L'effetto legato alla modalità è positivo solo se le informazioni presenti nelle diverse modalità non sono ridondanti: ad esempio, se in fase di IS il sistema elabora un tecnicismo per l'interprete che non lo ha sentito (Prandi, 2023, p. 79). Nella fase di comprensione, la presentazione multimodale facilita l'elaborazione (ad esempio, mediante l'integrazione multimodale), anche se l'interferenza tra stimoli visivi e uditivi si deve al fatto che entrambi sono riconducibili a un unico codice: quello verbale (Prandi, 2023, p. 79).

1.2.1 Il modello degli sforzi di Gile

Il modello degli sforzi di Gile mira a spiegare, attraverso i mezzi della psicologia cognitiva, la presenza di errori, omissioni e *infelicities* (o EOI) nella resa di un interprete (Gile, 1997), per poter fornire un potenziale strumento pedagogico che aiutasse nello stabilire il livello della *performance* (Gile, 1999, p. 2). Il processo dell'interpretazione simultanea prevede numerosi sotto-processi (che Gile definisce *sforzi*) non automatici, che richiedono un'allocazione attiva di risorse cognitive limitate (Prandi, 2023, p. 82). Gile (2009) identifica quattro sforzi: (1) lo sforzo di ascolto e analisi (o di ricezione, o L) che abbraccia tutte le operazioni mirate alla comprensione (dalla percezione delle onde sonore alle decisioni relative alla semantica delle stesse); (2) lo sforzo della memoria (o M) che comporta un costante confronto tra elementi percepiti e quelli presenti nella memoria, per procedere a una segmentazione logica del discorso; (3) lo sforzo di produzione (o P) che prevede la formulazione del messaggio in una lingua di arrivo e le operazioni di monitoraggio e autocorrezione; e (4) lo sforzo di coordinamento (o C) che coordina tutti gli altri sotto-processi. L'interpretazione simultanea è definita come la somma di tutti questi sforzi; pertanto: $IS = L + M + P + C$. Gile afferma che la somma di risorse richieste dai diversi sforzi (di seguito, R) non deve superare la somma delle capacità disponibili in termini di allocazione dell'attenzione (di seguito, A); pertanto l'IS è fattibile solo se: $IS = R \leq A$. Gile (2008) definisce come *local cognitive load* il sovraccarico

cognitivo relativo a fonti occasionali di difficoltà. Il modello degli sforzi di Gile viene da lui stesso integrato con la *tightrope hypothesis* (Gile, 1999, p. 198), secondo la quale gli interpreti lavorano costantemente vicini a livelli di saturazione delle risorse cognitive, cercando di mantenere l'equilibrio tra sotto-compiti che richiedono un costante coordinamento. Gli elementi che saturano il sistema e portano l'interprete a sperimentare un sovraccarico cognitivo e una difficoltà nella gestione dello stesso vengono definiti *problem trigger* (Gile, 1999, p. 157). Il modello viene inoltre integrato da un modello gravitazionale della disponibilità linguistica (in inglese, *Gravitational Model of Language Availability*), con cui lo stesso Gile spiega intuitivamente perché la terminologia specialistica potrebbe portare a un sovraccarico o a problemi nell'elaborazione cognitiva. Secondo questo modello, le parole o le espressioni utilizzate spesso dall'interprete (o *Units of Linguistic Knowledge*) gravitano al centro di esso e sono più rapidamente disponibili (Gile, 2009).

1.2.2 Il modello del carico cognitivo di Seeber e la sua applicazione in interpretazione simultanea

Come sintetizza Seeber (2011), ci sono alcuni punti di debolezza di questo modello: (a) *time-sharing*, *multi-tasking* ed elaborazione dell'input visivo non sono presi in considerazione nel modello; (b) il modello seleziona un unico *pool* di risorse che non permetterebbe di spiegare perché è possibile svolgere attività come la ricerca terminologica su CAI *tool* in interpretazione simultanea; (c) la *tightrope hypothesis* non spiega la disponibilità di risorse cognitive in eccesso durante l'IS.

Seeber (2011) osserva infatti che il modello di Gile (1995) basato su un *pool* unico di risorse non prende in considerazione l'interferenza tra due o più *task*. Secondo Seeber, l'interpretazione simultanea è un'attività che richiede principalmente quattro compiti, distribuiti in due macro-fasi: (1) ascolto e comprensione; (2) produzione e monitoraggio. Ascolto e comprensione coinvolgono le risorse auditivo-verbali e cognitivo-verbali nello stadio cognitivo-percettivo, portando a un'analisi del messaggio verbale mirata alla comprensione. Seeber (2017) ha successivamente aggiunto la risorsa visivo-spaziale per far riferimento alle informazioni para-verbali che l'interprete deve elaborare durante il processo interpretativo. Il *task* di produzione e monitoraggio richiede risorse auditivo-verbali e cognitivo-verbali nella fase percettiva, ma non solo: nella successiva fase di risposta (o *response stage*) entrano in gioco le risorse vocali-verbali. Il grado di interferenza dei diversi *task*, ossia la quantità di conflitto generata durante

l'esecuzione dei *task* in parallelo (Seeber, 2017, p. 467), viene calcolato utilizzando il *conflict matrix* di Wickens (2002); in particolare, l'interferenza può essere generata dalla sovrapposizione dei diversi livelli di elaborazione delle informazioni, dei diversi codici e delle diverse modalità (Seeber, 2007). Seguendo il modello di Wickens, in cui il livello di dipendenza di un *task* da una risorsa viene descritto su vettori di domanda (o *demand vectors*) con valori compresi tra 0 (non dipendenza) e 2 (massima dipendenza), Seeber postula un *demand vector* di 1 per ciascuno dei *micro-task* co-occorrenti (risorse), che in IS sono coinvolti in entrambi gli stadi dell'elaborazione (percezione + cognizione, risposta), in entrambe le modalità (uditiva e visiva) e nei due codici preposti all'elaborazione (spaziale e verbale). I compiti scomposti in vettori di domanda sono: P (elaborazione auditivo-percettiva e verbale di input e output); C (elaborazione cognitivo-verbale di input e output); R (elaborazione della risposta verbale in output). Interferenza (I) e Stoccaggio previo all' integrazione e alla produzione (o S) sono altre componenti evidenziate da Seeber (2011). Basandosi sui modelli di Wickens, Seeber (2011) descrive il carico cognitivo come il risultato di elementi caratterizzanti input e output, di interferenze tra *task* co-occorrenti, di elaborazioni parallele e dell'estensione temporale limitata. Questi elementi richiedono la costante applicazione di strategie da parte dell'interprete. Il modello di Seeber ha numerosi vantaggi che motivano la sua applicazione nella valutazione dell'impatto delle tecnologie sul carico cognitivo in IS. Infatti (a) illustra che la maggior parte del tempo le interpreti lavorano sotto il livello di saturazione; (b) considera che l'output è il risultato dell'applicazione di strategie di coordinamento tra *task* co-occorrenti; (c) considera l'interferenza tra più *task* concomitanti (descrivendo il potenziale conflitto); (d) ogni *task* viene diviso in sotto-compiti e (e) mira a quantificare il carico cognitivo con vettori di domanda e coefficienti di conflitto (Prandi, 2023).

Il modello del carico cognitivo di Seeber applicato all'interpretazione simultanea è comunemente conosciuto come *Cognitive Resource Footprint* (o CRF). Il modello aggiornato prevede tre possibili livelli (0, 0.5, 1), che dipendono dal se l'informazione all'interno di uno stadio è complementare o ridondante all'interno di quel livello: quando non vi è la presentazione di quella modalità (ossia, quando un determinato canale non ha un carico informativo) e di quel livello il valore assegnato è pari a 0; quando le informazioni sono presentate solo in una modalità all'interno di un livello, il vettore di domanda corrispondente è pari a 1; e quando le informazioni in una modalità sono complementari alle informazioni in un'altra modalità, il vettore di domanda è pari a 0,5 (Seeber, 2017, p. 468).

La somma dei *demand vector* e dei coefficienti di conflitto assegnati all'IS con input visivo (ossia l'IS nelle condizioni che verranno descritte in questo studio) è uguale al valore totale relativo all'interferenza (o *total interference score*, o TIS) di 11,6, un valore relativamente elevato che descrive un carico cognitivo nettamente superiore in condizioni di IS con input visivo non complementare all'input auditivo.

1.2.2.1 Modello del carico cognitivo per l'IS con testo o trascrizione automatica

Seeber (2017) ha applicato il suo modello anche all'istanza specifica di IS con trascrizione del discorso (o SIMTXT). Questo *setting* interpretativo richiede l'aggiunta di risorse visivo-verbali all'interno delle risorse cognitive necessarie.

Nel *conflict matrix*, l'elaborazione visivo-verbale è aggiunta allo stadio di ascolto e comprensione e ha un vettore dedicato; l'attribuzione di un vettore di domanda dal valore di 1 riflette la duplicazione dell'informazione uditiva in forma scritta (Prandi, 2023, p. 91). Di particolare interesse, sia per l'IS con testo sia per IS con trascrizione automatica è la presenza di informazioni scritte ridondanti. Seeber, et al. (2020, p. 13) hanno osservato che in *setting* di SIMTXT gli interpreti ricorrono al supporto audiovisivo probabilmente per scaricare la memoria a breve termine (Seeber et al., 2020, p.13) e non tanto per sfruttare l'effetto ridondanza per migliorare la comprensione. In sintesi, la ridondanza di stimoli visivo-verbali e auditivo-verbali potrebbe richiedere uno sforzo eccessivo nell'integrazione degli stessi nella fase di ascolto e comprensione, ma può rivelarsi utile nella fase di produzione, per controllare l'elaborazione ed eseguire un buon monitoraggio della resa (Prandi, 2023, p. 91). Seubert (2019) ritiene che la ridondanza informativa possa intervenire non solo nella fase di produzione, ma anche nella fase di pianificazione, ma è d'accordo sul fatto che le informazioni ridondanti replicate su due canali possono portare a un aumento del carico cognitivo (Prandi, 2023, p. 94). Seeber (2017, p. 463) osserva, tuttavia, che sebbene la complementarità e la ridondanza dei segnali sembrino essere componenti importanti del nostro ambiente naturale, il modo in cui li elaboriamo può dipendere dalla composizione del segnale, pertanto l'organizzazione dell'input visivo è sicuramente un elemento da prendere in considerazione. La trascrizione automatica delle funzionalità di *live captioning* (come quella di Google Meet, oggetto di analisi in questo studio) ha evidenti limiti per quanto riguarda la sua composizione: la trascrizione è in costante movimento, la latenza è significativa, la scarsa organizzazione testuale talvolta presenta errori,

la quantità delle risorse visivo-verbali che devono essere elaborate è elevata. Chmiel et al. (2020) hanno analizzato l'incongruenza tra input auditivo e scritto e hanno osservato che, quando assistiamo a stimoli scritti e auditivi co-occorrenti, le interpreti tendono a focalizzarsi sulla modalità visiva; inoltre, quando gli stimoli sono incongruenti, la modalità visiva interferisce con l'input auditivo conducendo l'interprete a commettere più errori.

1.2.2.2 Modello del carico cognitivo per l'IS con CAI e ASR

Un CAI *tool* integrato con ASR, a differenza di qualsiasi altro CAI che richiede una ricerca terminologica manuale anche in condizione *in-booth*, non richiede il coinvolgimento delle risorse motorie; tuttavia, le risorse visivo-spaziali e visivo-verbali sono comunque coinvolte per localizzare ed elaborare il termine mostrato sullo schermo (Prandi, 2023, p. 102). Il riconoscimento vocale agevola non solo la fase di produzione, ma anche la fase di comprensione, in quanto l'interprete può usufruire della capacità del CAI di riconoscere un termine che magari l'interprete non ha capito. Pym (2011) afferma che le tecnologie nel campo dell'interpretazione esternalizzano processi cognitivi, comportandosi come una *memoria esterna* (o *secondo cervello*): l'uso di un ASR offre dunque una maggior facilità nel riconoscimento della terminologia durante la fase di ascolto per ottimizzare la resa di terminologia specialistica (e non solo) in fase di produzione.

Il modello di *cognitive resource footprint* ideale (in condizione di IS con CAI *tool*) è basato sulla premessa di un utilizzo di un riconoscimento vocale che non commette errori. Tuttavia, dinanzi agli strumenti di ASR a disposizione, è evidente che la frequenza di errori del CAI *tool* e il conflitto tra quanto percepito dall'interprete e quanto appare sullo schermo hanno talvolta come risultato uno stress che può ulteriormente appesantire il carico cognitivo cui l'interprete è sottoposto (Prandi, 2023, p. 103), poiché i suggerimenti del *tool* potrebbero passare da essere elementi che agevolano il processo interpretativo a essere elementi che distraggono l'interprete nel coordinamento delle sue risorse (Cauwenberghe, 2020). La latenza dell'ASR spesso porta a una minor fluidità e accuratezza (Montecchio, 2021) e talvolta l'effetto benefico della ridondanza informativa nei diversi canali potrebbe essere inficiato dalla mancanza di un *décalage* adeguato (Mayer & Fiorella, 2014). Dinanzi alle due configurazioni del *cognitive resource footprint* in condizione di SIMTXT e IS con trascrizione automatica, si osserva curiosamente che le risorse coinvolte sembrano essere le stesse; dunque, la differenza risiede

nella misura in cui le risorse possono essere sfruttate al meglio nelle due condizioni. Entra dunque in gioco la capacità di attingere alle risorse visivo-verbali, che dipende – in gran misura – dalla capacità dell'ASR specifico e dall'organizzazione delle stesse in tempi brevi, affinché quelle stesse risorse possano essere utilizzabili non solo in fase di monitoraggio della produzione, ma anche nella fase auditivo-percettiva di comprensione.

1.3 Il riconoscimento vocale automatico e l'interpretazione simultanea

Il riconoscimento vocale automatico permette di convertire dei segnali vocali umani in una sequenza di parole, mediante un programma informatico (Jurafsky & Martin, 2009). I nuovi approcci computazionali (in particolare, il *Deep Learning* e le Reti Neurali Artificiali) e l'interesse commerciale hanno rinnovato l'interesse per l'ASR (Fantinuoli, 2017), anche se la perfezione è ancora lontana. Il linguaggio è, infatti, un sistema complesso in cui ascolto e decodifica dei suoni non sono i soli elementi necessari per giungere a una buona comprensione: le informazioni contestuali (conoscenze preliminari sull'oratore e sul contesto) e la capacità di individuare le ridondanze discorsive (migliorando la prevedibilità) sono infatti complementari ai segnali acustici (Fantinuoli, 2017). Fantinuoli (2017) osserva tre criticità relative all'uso dell'ASR in cabina:

(1) *Utilizzo della lingua parlata*: la trascrizione corretta dei discorsi improvvisati e non organizzati è una grande sfida per l'ASR; nei discorsi spontanei, gli errori degli esseri umani (disfluenze, esitazioni, ripetizioni, frasi aperte, ecc.) pongono un problema per l'ASR, portando a risultati talvolta discutibili;
(2) *Variabilità dell'oratore*: ogni oratore ha uno stile, una voce, un accento regionale e un genere che influenzano il segnale vocale e richiedono un'adattabilità elevata dell'ASR (basti pensare all'uso dell'inglese come *lingua franca*);
(3) *Ambiguità*: è una caratteristica inerente al linguaggio naturale stesso (gli omofoni ne sono un evidente esempio);
(4) *Discorso continuo*: il riconoscimento dei confini delle parole è un problema per l'ASR, in quanto spesso il discorso non ha pause naturali tra le parole (e questo ha effetti anche sull'organizzazione sintattica del discorso);
(5) *Rumore di fondo*: il rumore può inficiare negativamente la qualità dell'output dell'ASR;

(6) *Velocità del discorso*: la maggior velocità del discorso può portare gli oratori a non scandire bene le parole e questo può avere effetti sull'ASR data la sua scarsa prevedibilità sulla semantica del discorso;
(7) *Linguaggio del corpo*: i segnali non verbali (gesti, espressioni facciali, ecc.) fanno parte della comunicazione umana; attualmente, questi elementi comunicativi non vengono presi in considerazione nei sistemi di ASR.

I sistemi di ASR sono tendenzialmente divisi in due classi: quelli che riconoscono parole isolate nel discorso (in cui queste sono precedute e seguite da una pausa) e quelli che riconoscono il discorso continuo (in cui gli enunciati sono pronunciati in modo naturale ed è il programma a dover riconoscere i confini delle singole parole) (Fantinuoli, 2017). Nell'integrazione dell'ASR all'interno dei CAI, osserviamo un processo in cui l'intero discorso deve essere trascritto affinché il CAI possa selezionare i segmenti testuali pertinenti, per poter avviare l'algoritmo di interrogazione del *database*, andando a identificare gli elementi rilevanti (numerali e nomi propri, ad esempio).

1.3.1 Architettura di base di un sistema di ASR

Un sistema di riconoscimento vocale viene sviluppato sulla base di alcune macro-componenti, tra cui un'interfaccia acustica (in inglese, *acoustic front-end*), un modello acustico (o *acoustic model*), un vocabolario (o *lexicon*), un modello di linguaggio (o *language model*) e un decodificatore (o *decoder*) (Karpagavalli & Chandra, 2016, p. 394). L'interfaccia acustica trasforma l'onda sonora in una serie di parametri acustici utili per il riconoscimento e l'onda sonora dell'input acustico viene convertita, pertanto, in una sequenza di vettori acustici a dimensione fissa tramite il processo di estrazione di caratteristiche (in inglese, *feature extraction*). I parametri dei modelli parola/suono sono stimati a partire dai vettori acustici dei dati inseriti nella fase di addestramento dei modelli (*Ibid.*, p. 394). Il decodificatore fa una ricerca tra tutte le possibili sequenze per cercare la sequenza di parole che ha maggior probabilità di essere generata (*Ibid.*, p. 395). La funzionalità chiave di un sistema di riconoscimento automatico del parlato può essere descritta come un processo di estrazione di un certo numero di parametri vocali dal segnale acustico del parlato per ogni parola (o sotto-unità della stessa) (*Ibid*, p. 395). I parametri vocali descrivono la parola (o una sua sotto-unità)

attraverso la variazione nel corso del tempo e insieme vanno a costruire un *pattern* che caratterizza quella parola (o quella sotto-unità) (*Ibid.*, p. 395). Nella fase di addestramento, le parole di tutto il vocabolario devono essere lette e i *pattern* delle singole parole sono stoccati e, quando successivamente una parola deve essere riconosciuta, il suo *pattern* viene paragonato ai *pattern* stoccati e la parola che presenta una maggior corrispondenza di *pattern* viene selezionata (questo processo è conosciuto come processo di *pattern recognition*, mediante il quale è possibile associare un input orale a un output testuale) (*Ibid.*, p. 395).

L'interfaccia acustica è responsabile dell'elaborazione del segnale e dell'estrazione delle sue caratteristiche (o della sua *feature extraction*); l'obiettivo principale nella fase di estrazione delle caratteristiche è quello di calcolare una sequenza minima di vettori di caratteristiche (o *feature vector*) che forniscano una rappresentazione compatta del segnale di ingresso dato (*Ibid.*, p. 395). Il processo di estrazione di caratteristiche prevede tre fasi: l'analisi del discorso (tramite un'analisi dello spettro temporale del segnale acustico e mediante la generazione di caratteristiche grezze che descrivono lo spettro di potenza di brevi intervalli del segnale acustico), la creazione di un *feature vector* esteso composto da caratteristiche sia statiche che dinamiche e la trasformazione del vettore esteso in vettori più robusti e compatti (Karpagavalli & Chandra, 2016, p. 395).

Il modello acustico è una delle principali fonti di informazioni per il sistema di riconoscimento vocale, in quanto presenta le caratteristiche fondamentali per il riconoscimento delle unità foniche. Questo modello serve per calcolare la distanza tra i vettori estratti dall'audio registrato e i vettori campione contenuti nel modello acustico di riferimento. Il modello di linguaggio è una raccolta dei limiti circa le sequenze di parole accettabili in una determinata lingua, questi limiti sono determinati, ad esempio, dalle regole grammaticali o da statistiche relative alla frequenza di una coppia di parole in un dato corpus (*Ibid.*, p. 397). Il modello di linguaggio serve dunque per fornire al sistema di riconoscimento vocale il contesto linguistico delle parole. Nel linguaggio degli esseri umani, i suoni simili sono spesso distinguibili grazie alle conoscenze relative al contesto comunicativo, conoscenze che un sistema di riconoscimento vocale deve ottenere proprio tramite l'uso di un modello di linguaggio. Il modello di linguaggio specifica, dunque, quali sono le parole valide in quella lingua e in quale sequenza si possono presentare (*Ibid.*, p. 397).

La fase di decodificazione ha come obiettivo quello di trovare la sequenza di parole (W) più probabile, data la sequenza osservata (O) e dato il modello acustico-fonetico e di linguaggio di riferimento (*Ibid.,* p. 398). Il problema relativo alla decodificazione può essere risolto tramite

l'uso di algoritmi di programmazione dinamica. Piuttosto che valutare le probabilità di tutti i possibili percorsi del modello che generano W, l'attenzione si concentra sulla ricerca di un singolo percorso attraverso la rete che produce la migliore corrispondenza con O (*Ibid.*, p. 398).

1.3.2 Valutazione dell'ASR

Per valutare la *performance* dei diversi sistemi di ASR, tendenzialmente vengono valutati due parametri: l'accuratezza dell'output del modello proposto e la velocità di elaborazione dello stesso.

La velocità viene tendenzialmente calcolata utilizzando il *real-time factor* (o RTF) (Pratap et al., 2020).

L'RTF può essere calcolato con la seguente formula:

$$RTF = \frac{P}{I}$$

Nella formula, P è il lasso di tempo necessario al sistema per poter elaborare l'input e I fa riferimento alla durata dell'audio in input. Se RTF è pari a 1, allora l'audio in input viene elaborato in tempo reale (*real-time*). Il valore del *real-time factor* dipende fortemente dall'*hardware* e non è limitato al calcolo della velocità di un modello di un riconoscimento vocale, ma può anche essere applicato per il calcolo di qualsiasi modello in grado di elaborare un audio o un video in input (Pratap et al. 2020).

L'accuratezza di un ASR viene tendenzialmente misurata utilizzando due metodi: il *word error rate* (o WER) e il *word recognition rate* (*Ibid.*).

L'accuratezza è difficile da calcolare, in quanto l'output prodotto dall'ASR potrebbe non avere la stessa lunghezza dei dati empirici reali. Il WER è il metodo utilizzato per stimare l'accuratezza della *performance* di un ASR, calcolando l'errore al livello della parola (e non al livello del fonema). La stima del WER viene pertanto calcolata in maniera automatica e si basa sull'analisi di una trascrizione manuale (allineata al segnale) e la relativa trascrizione ottenuta

dal sistema ASR (Palmerini & Savy, 2014, p. 281). Il *word error rate* può essere calcolato utilizzando la seguente formula:

$$WER = \frac{S + D + I}{N}$$

Nella formula, osserviamo che il WER è la combinazione di tre tipi di errori di trascrizione che possono verificarsi: S indica il numero di errori di sostituzione (parole presenti sia nell'ipotesi generata dall'ASR che nei dati empirici reali, ma non trascritte correttamente); I indica il numero di errori di inserzione (parole presenti nella trascrizione dell'ipotesi e non presenti nei dati empirici reali) e D indica il numero di errori di eliminazione (parole non presenti nell'ipotesi, ma presenti nei dati empirici reali) (Google Cloud, 2024). N fa riferimento al numero totale di parole nella trascrizione basata sui dati empirici reali.

Il *word recognition rate* è una declinazione specifica del *word error rate* e anch'esso può essere utilizzato per valutare l'accuratezza della *performance* di un ASR. Per il suo calcolo, è necessario ricorrere alla formula seguente:

$$\begin{aligned} WRR &= 1 - WER \\ &= \frac{N - S - D - I}{N} \\ &= \frac{H - I}{N} \end{aligned}$$

In questa formula, H = N – (S + D) rappresenta il totale delle parole correttamente indovinate (Dietrich & Peters, 2002). Osservando le formule, N corrisponde al numero di parole totali che vengono prese in considerazione, S è uguale al numero delle sostituzioni, D è pari al numero delle omissioni del sistema, I è uguale al numero delle aggiunte e H è uguale al numero di parole corrette.

1.3.3 Integrazione dell'ASR nei CAI *tool*: requisiti minimi

Per poter essere integrato in un CAI *tool*, come sottolinea Fantinuoli (2017) il sistema di ASR deve soddisfare i seguenti requisiti minimi:

(a) Deve essere *speaker-independent* (in questo modo non è necessario che l'oratore invii audio riservati a un fornitore di servizi esterno per abituare il sistema alla sua voce);
(b) Deve essere in grado di gestire un flusso discorsivo continuo;
(c) Deve essere in grado di aiutare nel riconoscimento di un vocabolario ampio;
(d) Deve essere adattabile per il riconoscimento della terminologia specialistica;
(e) Deve avere un'accuratezza elevata (e quindi un basso *word error rate*);
(f) Deve avere una velocità elevata (e quindi un basso *real-time factor*, o RTF).

Fantinuoli (2017) aggiunge che al fine di supportare con successo l'integrazione di un sistema ASR, il CAI deve a sua volta soddisfare i seguenti requisiti:

- Precisione elevata, intendendo per precisione la frazione di occorrenze rilevanti tra le occorrenze recuperate;
- *Recall* elevato, intendendo per *recall* la frazione di occorrenze rilevanti che sono state recuperate rispetto alla quantità totale di occorrenze rilevanti presenti nel discorso;
- La precisione ha priorità rispetto al *recall*, in quanto è meglio evitare direttamente la presenza di errori che potrebbero distrarre l'interprete;
- L'ASR deve essere in grado di individuare le variazioni morfologiche tra trascrizione e voci del *database* senza aumentare il numero dei risultati;
- Il CAI deve disporre di un'interfaccia grafica semplice e priva di distrazioni per la presentazione dei risultati.

1.3.4 Limiti e sfide per il riconoscimento vocale automatico

I problemi relativi alla cattura del suono e alla preelaborazione del discorso (quindi relativi al vocabolario, ai problemi di pronuncia e ai dialetti di una stessa lingua) sono sfide ricorrenti nella letteratura sul riconoscimento vocale. Nella fase di preelaborazione, uno dei dilemmi nel *natural language processing* (o NLP) è la diversità linguistica dovuta alla presenza di dialetti e cadenze regionali che spesso rendono il linguaggio utilizzato nella vita quotidiana molto difficile da elaborare per i sistemi di ASR (Hirayama et al., 2015). Sempre a proposito della variabilità linguistica, un'altra sfida è relativa al discorso spontaneo: i problemi di pronuncia (riconducibili in alcuni casi anche a disturbi, quali la balbuzie) e la minor prevedibilità del discorso (ad esempio, la presenza di frasi aperte) hanno ripercussioni dirette sulla qualità della *performance* dell'ASR. Sempre nella fase di preelaborazione, è necessario osservare la

proporzionalità indiretta tra ampiezza del vocabolario di un sistema ASR e velocità di risposta del sistema, dato che un vocabolario più grande richiede una maggior capacità di calcolo del sistema da parte dell'ASR (Cai et al., 2015). Uno degli ostacoli, ad oggi più significativi, è l'ottenimento di una *performance* accurata in presenza di rumore di fondo: la *performance* dei sistemi di ASR registra un crollo nella qualità in presenza di ambienti rumorosi (Agrawal & Ganapathy, 2019). Sebbene siano stati proposti molti algoritmi, la maggior parte di essi presenta un peggioramento notevole della qualità in ambienti rumorosi (Ming & Crookes, 2017). Inoltre, gli ambienti rumorosi contribuiscono all' assenza spettro-temporale dei segnali di ingresso, il che porta a perdere le correlazioni tempo-frequenza del segnale vocale sottostante (Ganapathy, 2017). Anche se non strettamente correlato al presente studio, un altro limite all'accuratezza degli attuali sistemi di ASR è dato da condizioni quali la sovrapposizione di più parlanti, ovvero la conversazione simultanea (nella letteratura questo problema è comunemente conosciuto come *cocktail party problem*) (Chen et al., 2018). L'interfaccia acustica, inoltre, dipende in gran misura anche dall'*hardware* e – quindi – dal microfono utilizzato: microfoni mediocri portano a una riduzione notevole della qualità della *performance* dei sistemi di ASR (Gosztolya & Grósz, 2016): i microfoni dovrebbero essere tanti quanti le fonti degli input acustici (quindi, tanti quanti gli oratori).

Nonostante le sfide, ci si aspettano numerosi passi avanti nei diversi domini dell'ASR, che vanno dai modelli acustici e semantici all'IA conversazionale e generativa (Rosenwein, 2023). In particolare, a livello dei modelli acustici, Rosenwein (2023) si aspetta un ingrandimento (*data augmentation*, in inglese) dei dati utilizzati per l'addestramento di modelli sempre più grandi; mentre a livello semantico, sempre secondo lo stesso studio, i ricercatori si concentreranno sempre più sul miglioramento dei modelli di ASR integrando contesti acustici e testuali sempre più ampli. Come confermato dallo studio di Rosenwein (2023), si prevede che i progressi nell'apprendimento auto-supervisionato, nell'apprendimento *multi-tasking* e nei modelli multilingue miglioreranno significativamente le prestazioni nelle lingue con risorse linguistiche scarse o non scritte. Questi metodi raggiungeranno prestazioni accettabili utilizzando modelli pre-addestrati e perfezionando la messa a punto su un numero relativamente ridotto di campioni etichettati. Nel campo dell'IA generativa, si prevede che l'uso degli avatar rivoluzionerà l'interazione umana con le risorse digitali. Nel breve termine, l'ASR servirà come uno dei fondamenti dell'IA generativa, poiché questi avatar comunicheranno attraverso un'interfaccia testuale/uditiva. Nel lungo termine, tuttavia, potrebbero verificarsi dei cambiamenti; ad esempio, una tecnologia emergente che probabilmente verrà adottata è la *Textless NLP*, che rappresenta un nuovo approccio di modellazione linguistica per la

generazione di audio (Lakhotia et al., 2021). Questo approccio permette – in modo simile alla generazione automatica di un testo – di utilizzare delle unità audio (con cui il sistema è stato addestrato) per poi generare unità audio successive in modo autonomo (Zeghidour et al., 2021). L'IA conversazionale è stata adottata principalmente attraverso sistemi di dialogo orientati al compito (o *task-oriented dialogue system*), ovvero assistenti personali come *Alexa* di *Amazon* e *Siri* di *Apple*, in grado fornire un accesso rapido a funzioni e informazioni attraverso i comandi vocali. Nel futuro potremo aspettarci un'interoperabilità tra gli assistenti personali, il che significa che inizieranno a comunicare tra loro e questo permetterà di utilizzare qualsiasi dispositivo per connettersi a qualsiasi agente conversazionale in qualsiasi parte del mondo. La segmentazione del parlato sulla base dell'identità del parlante (o *diarizzazione*) ha visto un primo punto di contatto con l'ASR nel 2019 (El Shafey et al., 2019), ma l'arrivo di sistemi *end-to-end* di diarizzazione del parlato migliorerà la prestazione dell'ASR in situazioni di sovrapposizione del parlato, consentendo migliori prestazioni in scenari acustici difficili (come in ambienti lontani e riverberanti) (Rosenwein, 2023). Nel punto di contatto tra necessità dell'interprete in cabina e sviluppi dell'ASR, una sfida inevitabile sarà la riduzione della latenza per velocizzare il tempo di risposta del sistema migliorandone la sua usabilità.

1.4 SmarTerp

SmarTerp è un progetto per l'innovazione finanziato dall'Unione Europea nel quadro della sovvenzione EIT Digital BP2021. Il progetto ha avuto, sin dall'inizio, l'obiettivo di sviluppare una piattaforma di RSI integrata con un CAI *tool* basato su ASR e IA (Frittella & Rodríguez, 2022, p. 141). Il progetto è stato avviato nell'autunno del 2020 dall'interprete di conferenza Susana Rodríguez. Grazie al suo coordinamento, un gruppo interdisciplinare di esperti del settore e un *think tank* costituito da diversi *stakeholder* è stato coinvolto nel progetto (*Ibid.*, p. 141). L'usabilità dello strumento e la valutazione del *design* dell'interfaccia sono stati i principali elementi di indagine scientifica di Frittella, con l'obiettivo di proporre un *design* sempre più *utilizzabile* da parte dell'utente finale (*Ibid.*, p. 141). Il progetto di SmarTerp è stato possibile grazie alla collaborazione tra esperti di NLP/ASR, sviluppatori di *software* e interpreti professionisti che hanno avuto sin dall'inizio l'obiettivo di favorire l'integrazione dei sistemi di ASR in cabina di simultanea. L'ottima esecuzione del sistema di ASR in domini di applicazione specifici è stato uno dei principali obiettivi perseguiti sin dall'inizio dello sviluppo

del CAI *tool* di SmarTerp (Rodríguez et al., 2021, p. 101). Più precisamente, la lingua di partenza può contenere un gran numero di termini tecnici e di variazioni morfologiche che di solito hanno una bassa frequenza nei corpora generici (*Ibid.*, p. 103). Il risultato è che un modello di linguaggio generico (LM) presenta, su dati di dominio, valori elevati nel numero di parole fuori vocabolario (o *out-of-vocabulary*, o OOV) e di allucinazioni, peggiorando il *word error rate* (WER) del sistema ASR che lo utilizza (*Ibid.*, p. 103). Per evitare questo effetto, gli sviluppatori e le sviluppatrici di SmarTerp propongono un sistema che estrae da un dato corpus i testi che sono *più vicini*, in qualche modo, al glossario di termini fornito dall'interprete in fase *pre-booth* (*Ibid.*, p. 103); in questo modo, l'obiettivo è quello di ottenere un sistema di ASR specifico per ogni sessione di interpretazione.

1.4.1 Modelli Acustici di SmarTerp

I modelli acustici di SmartErp sono addestrati su dati di *Common Voice*[8] e le trascrizioni di *Euronews*[9], utilizzando una configurazione standard (o *standard chain recipe*), nella fattispecie, Kaldi, basata sul criterio di ottimizzazione LF-MMI (o *lattice-free maximum mutual information*) (Rodríguez et al., 2021, p. 103). Per garantire una maggior robustezza del sistema dinanzi alle possibili variazioni della velocità d'eloquio dei parlanti, la consueta tecnica di amplificazione dei dati (o di *data augmentation*) per i modelli SmarTerp è stata ulteriormente ampliata, generando versioni allungate nel tempo del *set* di addestramento originale (Rodríguez et al., 2021, p. 103).

1.4.2 Modelli di Linguaggio

Nella configurazione di una sessione con SmarTerp è indispensabile configurare un glossario da cui poter estrarre le *seed word* cruciali sia per aggiornare il dizionario del sistema di ASR, sia per selezionare i testi del modello di linguaggio (di seguito, ML) dai corpora di addestramento disponibili (Rodríguez, et al., 2021, p. 103). Questi testi del ML provengono da

[8] Si tratta di un progetto di *Mozilla Foundation* che mira a raccogliere un gran numero di ore di registrazione di voci su cui poter far allenare i vari *software* per il riconoscimento vocale. https://commonvoice.mozilla.org/it (Ultimo accesso: 24 aprile 2024)
[9] https://www.euronews.com/ (Ultimo accesso: 24 aprile 2024)

notizie raccolte su Internet (nell'arco temporale 2000-2020) e da Wikipedia; per la lingua spagnola (di particolare interesse per il presente studio sperimentale), i corpora di testo utilizzati per l'addestramento dei ML per l'ASR di SmarTerp hanno le seguenti caratteristiche: un *lexicon* di 4,182,225 unità, un numero totale di parole pari a 2246,07 milioni di parole (di cui 1544,51 milioni provenienti da notizie di Internet e 701,56 milioni provenienti dal sito di *Wikipedia* aggiornato al 2018) (Rodríguez et al., 2021, p.104).

Anche se apparentemente questi numeri possono sembrare molto elevati, si noti che i dati provenienti da Internet molto spesso contengono termini errati (refusi); pertanto, per la selezione dei testi, gli sviluppatori e le sviluppatrici hanno seguito tre criteri (Rodríguez et al., 2021, p. 104):

(a) la selezione di *seed words,* ossia termini tecnici che caratterizzano un tema determinato: si tratta delle parole fornite dall'interprete;
(b) la selezione di *adaptation text,* ossia frasi nel corpus di addestramento che contengano almeno una delle *seed word;*
(c) la creazione di un lessico adattato (o *adapted lexicon*) e di un ML adattato (o *adapted LM*).

Poiché si possono utilizzare diversi approcci per ottenere e utilizzare le *seed word*, sono stati stabiliti i seguenti indicatori per misurare la loro efficacia su test di riferimento raccolti e trascritti manualmente nell'ambito del progetto SmarTerp: (1) il tasso di OOV (o *out-of-vocabulary words*), le parole OOV porterebbero a un output dell'ASR con errori, pertanto è fondamentale avere un basso tasso di parole OOV senza aumentare troppo la grandezza del lessico; (2) WER del sistema di ASR; (3) Precisione e *Recall* su campioni di parole tecnicamente significative (Rodríguez et al., 2021, p. 104).

1.4.3 Interpretazione Semantica

Una volta che l'ASR del sistema ha trascritto l'input audio, il ruolo del modulo per l'interpretazione semantica entra in gioco per individuare le entità rilevanti che emergono nelle trascrizioni e che potrebbero essere utili per l'interprete (Rodríguez et al., 2021, p. 104), ad esempio termini molto specifici di un dominio e valori numerici. La sfida principale in questo contesto è che non ci troviamo di fronte a un tipico problema di riconoscimento di entità

denominate, in cui è necessario individuare elementi come nomi di luoghi, persone o organizzazioni (Rodríguez et al., 2021, p. 105) (ad esempio, riconoscere l'entità *Italy* nel testo e offrire la sua potenziale traduzione in italiano *Italia* risulterebbe inutile); tuttavia, potrebbe essere di grande aiuto la conversione del valore numerico ("cinquemila trentacinque") nella sua sequenza numerica araba (5035), in modo da permettere all'interprete il suo immediato riconoscimento (Rodríguez et al., 2021, p. 105). La sfida nello sviluppo di un sistema di riconoscimento delle entità denominate (o di *named entity recognition*) funzionale alla sua integrazione nella cabina di simultanea risiede dunque nella ricerca di un sistema che permetta il riconoscimento di termini rilevanti per l'interprete (o *Interpreter-relevant Term Recognition*) per poter permettere una traduzione rapida ed efficiente nella lingua di arrivo (Rodríguez, et al., 2021, p. 105). Per raggiungere questi obiettivi, il sistema proposto è basato su un insieme stratificato di reti semantiche multilingue (grafi di conoscenza, o *knowledge graph*) non solo generiche, ma anche specifiche per il dominio e specifiche per l'utente, seguendo le migliori pratiche relative ai dati collegati (o *linked data*) multilingue (Rodríguez et al., 2021, p. 105).

1.4.4 Architettura del sistema

Come illustrato in Rodríguez et al. (2021, p. 106), l'integrazione dei diversi moduli (*console* del tecnico, *console* RSI dell'interprete, Cloud ASR, CAI *tool*) può essere descritta come segue:

- *Console* del tecnico: l'interfaccia *web-based* utilizzata dal personale tecnico che gestisce una sessione di interpretariato permette l'inserimento nel sistema dell'audio e del video delle oratrici e degli oratori.
- La *console* RSI dell'interprete permette la visione del video dell'oratore/oratrice e degli eventuali materiali condivisi. Gli interpreti possono utilizzare la loro *console* per la gestione dei canali audio di input e di output, in base alle loro necessità. Nell'interfaccia dell'interprete, è possibile anche visualizzare l'output del CAI *tool*. L'interfaccia dell'interprete prevede, quindi, tre componenti principali: (1) un sistema di RSI con qualità audio e video in conformità con le norme ISO; (2) strumenti per la comunicazione con tutti gli attori coinvolti (tecnico, moderatore, altri interpreti); (3) un CAI *tool* basato su ASR e IA che offre supporto all'interprete nella resa di *problem trigger* frequenti, quali entità denominate, acronimi, termini specialistici e numeri (Frittella & Rodríguez, 2022, p. 142).

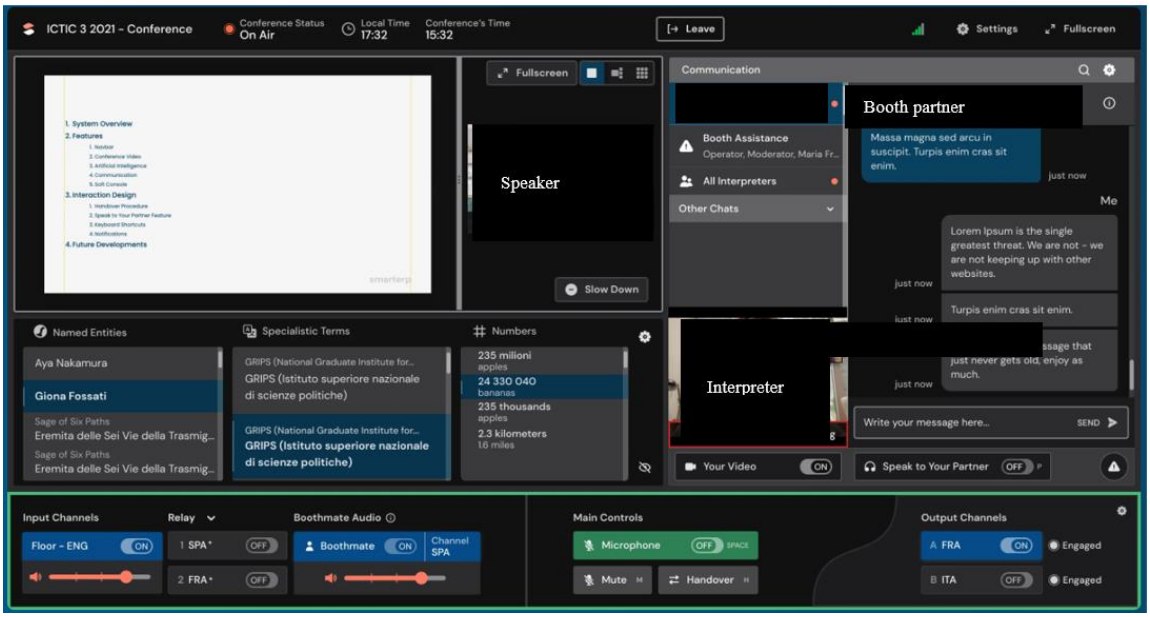


Figura 1 Interfaccia Interprete di SmarTerp in RSI: integrazione del CAI tool nella piattaforma di RSI (Frittella & Rodríguez, 2022, p. 142).

- Il Cloud ASR rappresenta il servizio *cloud on-demand* incaricato di trascrivere in tempo reale un input audio per produrre un output testuale, mediante una tecnologia *WebSocket*.
- Il CAI *tool* riceve la trascrizione dell'audio e genera una serie di termini volti a fornire un supporto all'interprete durante l'interpretazione simultanea.

1.5 La trascrizione automatica in cabina di simultanea

Il distanziamento sociale dovuto alle misure per fronteggiare la pandemia del Covid-19 ha portato a una maggior presenza nel mercato dell'interpretazione della modalità a distanza e l'interazione interprete-macchina è diventata sempre più collaborativa e ricca di novità. L'iniziale difficoltà nella collaborazione con il collega ha portato i professionisti a ricorrere a metodi di *prompting* (ovvero l'attività di supporto a chi in quel momento sta interpretando da parte del collega a riposo) con caratteristiche nuove e la trascrizione automatica (oggi presente nelle maggiori piattaforme, quali Google Meet e Zoom) è stata oggetto di interesse per molti interpreti. Anche se numerosi studi hanno messo in luce i meccanismi sottostanti all'elaborazione dei sottotitoli nel campo della ricerca audio-visiva (Kruger & Steyn, 2013) e altri hanno tentato una spiegazione dei processi di elaborazione del testo durante

l'interpretazione simultanea (Seeber et al., 2020), è fondamentale tracciare il confine tra sottotitoli/simultanea con testo (di seguito, SIMTXT) e simultanea con trascrizione automatica (Yuan & Wang, 2023, p. 1): se da un lato sottotitoli (generati da traduttori professionisti) e SIMTXT hanno una piena congruenza rispetto a quanto afferma l'oratore (in quanto input visivo-verbale e auditivo-verbale hanno un'elevata coincidenza), dall'altro la trascrizione automatica non è immune da errori (Yuan & Wang, 2023, p. 2). Stachowiak-Szymczak & Korpal (2019) e Chmiel et al. (2020) mettono in evidenza l'impatto sui processi cognitivi durante l'attività di simultanea dell'incongruenza tra ciò che viene detto e ciò che viene presentato all'interprete; tuttavia, come sottolineano Yuan & Wang (2023, p. 2), questi studi utilizzano stimoli sperimentali statici (*slide* e testi scritti) che sono inerentemente diversi dalla natura dinamica della trascrizione automatica (o dei sottotitoli codificati). Quando gli interpreti utilizzano la trascrizione automatica, si trovano davanti a una duplice sfida: da una parte devono elaborare la natura dinamica del testo e d'altra parte devono svolgere una valutazione critica della sua accuratezza (Yuan & Wang, 2023, p. 2). Gli stessi autori affermano che è plausibile il miglioramento della completezza e dell'accuratezza della resa mediante l'uso dell'output della trascrizione automatica, ma al contempo sottolineano l'importanza di riconoscere la necessità di una valutazione che metta in luce i benefici e la potenziale distrazione causata dalla trascrizione automatica, dovuta anche alla presenza di un ulteriore input informativo-visivo che richiede un'ulteriore elaborazione cognitiva.

L'input testuale presente nella trascrizione automatica comporta due fattori di particolare interesse: (1) l'attività di interpretazione simultanea prevede l'influenza di un ritmo esterno (è *externally paced*) (Chmiel & Lijewska, 2019), quindi non è l'interprete – ma l'oratore – colui che decide la velocità di elaborazione dell'input; (2) l'input testuale costituisce l'input secondario (e non primario), mentre l'input auditivo costituisce l'input primario (Pöchhacker, 2016). Seeber (2017) suggerisce che quando si lavora con il testo, gli interpreti potrebbero fare affidamento sulla complementarità del segnale per risolvere le ambiguità (quando il segnale fornito su un canale compensa le inadeguatezze del segnale fornito su un altro canale). Allo stesso modo, gli interpreti potrebbero sfruttare la ridondanza del segnale (quando è presente un segnale identico su due canali diversi), anche se gli studi non sono ad oggi conclusivi circa i possibili vantaggi (Lu et al., 2012).

1.5.1 Trascrizione automatica di Google Meet

Google Meet, un servizio di videocomunicazione sviluppato da Google nel 2017, ha una funzione di ASR che permette la generazione di una trascrizione automatica in tempo reale. Ad oggi, è possibile avere le didascalie automatiche (o sottotitoli codificati) in 31 lingue. Ai fini del presente elaborato, verranno presentati alcuni elementi caratterizzanti l'ASR di Google Meet di particolare importanza nel contesto dell'analisi dell'interazione tra la trascrizione automatica di Google Meet e l'interprete in cabina.

Ramadhika et al. (2022) mettono in luce alcuni dei principali elementi:

1. È interessante notare che finché l'oratore pronuncia correttamente le parole, anche a un ritmo veloce, l'ASR (e, quindi, il relativo algoritmo di IA) continuerà a funzionare con precisione nella generazione di una trascrizione automatica a partire dai suoni; tuttavia, l'intelligenza artificiale è molto sensibile ai suoni dell'oratore, tanto che una parola deve essere pronunciata correttamente per poter essere trascritta. Dunque, finché il parlante pronuncia correttamente le parole, l'intelligenza artificiale non sostituirà le parole con altre parole con pronuncia simile;
2. L'IA della trascrizione automatica di Google Meet ha un sistema di correzione automatica, ma esso è limitato esclusivamente al livello delle parole. L'IA trasforma i suoni in parole che hanno una pronuncia simile a quella della parola pronunciata dall'oratore, portando, talvolta, alla generazione arbitraria di parole diverse da quelle pronunciate;
3. Il sistema di correzione automatica di Google Meet non si preoccupa della correttezza grammaticale e non corregge eventuali mancanze di coesione sintattica;
4. La maggior velocità d'eloquio potrebbe comportare una maggior probabilità di omissioni nell'output testuale, in quanto l'algoritmo di IA necessita di un lasso di tempo minimo per elaborare i suoni delle parole;
5. L'ASR e l'IA di Google Meet hanno inoltre una notevole difficoltà nell'uso della punteggiatura in modo corretto. Tendenzialmente vengono usati esclusivamente il punto e la virgola in corrispondenza di pause nell'eloquio dell'oratore e si osserva un'alta frequenza di errori nell'uso delle lettere maiuscole.

1.6 Interpretazione automatica

L'interpretazione automatica (o *machine interpreting*, *speech translation*, traduzione *speech-to-speech*) è la tecnologia che permette la traduzione di testi orali da una lingua ad un'altra per mezzo di un programma informatico (Fantinuoli, 2018a, p. 5). L'obiettivo dell'interpretazione automatica è la sostituzione dell'interprete umano e, come evidenzia Fantinuoli (2018a, p. 5), in questo senso si distingue dalle altre due tecnologie relative al campo dell'interpretazione: i CAI *tool* (che offrono un supporto all'interprete umano) e le piattaforme di RSI (che mirano a offrire un'alternativa alla modalità di erogazione del servizio di simultanea in presenza). L'interpretazione automatica vede la combinazione di almeno tre tecnologie: il riconoscimento vocale (che trascrive il testo orale in forma scritta), la traduzione automatica (o *machine translation*) e la tecnologia di sintesi *text-to-speech*, o TTS (per generare una versione audio del testo tradotto nella lingua di arrivo) (Fantinuoli, 2018a, p. 5). Il rapido progresso relativo ai sistemi di ASR basati su reti neurali e i passi avanti nel campo della traduzione automatica neurale hanno congiuntamente portato a un miglioramento esponenziale della qualità e della precisione dei sistemi di interpretazione automatica. L'interpretazione automatica ha visto il suo avvento con i primi prototipi di traduzione automatica in tempo reale delle lezioni universitarie presso il Karlsruhe Institute of Technology (Müller et al., 2016) che sono stati il risultato di una lunga ricerca nel campo dell'elaborazione del linguaggio naturale (o *natural language processing*). Nei sistemi di *machine interpreting* che sono stati proposti sino ad oggi, molto spesso la latenza e la limitata flessibilità dei sistemi di ASR utilizzati sono stati alcuni dei principali fattori che hanno portato a sistemi di bassa qualità e con un'elevata frequenza di inaccuratezze (Fantinuoli, 2018a, p. 6). Inoltre, la qualità e l'usabilità rimangono due delle principali sfide, non solo a livello tecnico, ma anche a livello comunicativo. Relativamente a quest'ultimo, i sistemi di *machine interpreting* non sono in grado di lavorare prendendo in considerazione il cotesto/contesto e di tradurre tutte le informazioni che non sono verbalmente codificate (Fantinuoli, 2018a, p. 6). In questo panorama, sono tuttavia emersi due sistemi di interpretazione automatica particolarmente promettenti: *Translatotron 2* (di Google AI) e *Seamless* (di Meta AI).

Le attuali tecnologie per la traduzione sono per lo più incentrate sulla traduzione da un testo a un altro (*text-to-text translation*, T2TT). Sebbene l'astrazione del testo sia funzionale per portare il modello a un certo livello, ha il problema di disperdere vari elementi non verbali all'interno della conversazione: questo problema è particolarmente evidente nelle conversazioni. Quando

due persone di lingua diversa utilizzano un traduttore, di solito si assiste a un processo del tipo: *discorso L1 → riconoscimento del testo L1 (testo L1) → L1L2 traduzione (testo L2) → L2 TTS (discorso L2)*. Le informazioni non verbali che scompaiono attraverso il riconoscimento e la traduzione diventano una barriera difficile da superare con i sistemi di traduzione tradizionali (Hee-Jo, 2021). Come soluzione a questo problema, Google ha pubblicato *Translatotron* nel 2019 e *Translatotron 2* nel luglio del 2021. La differenza principale tra *Translatotron* e i modelli di traduzione convenzionali è che si concentra sul parlato (traduzione *speech-to-speech*, S2ST) (Hee-Jo, 2021). Jia et al. (2021) affermano che i sistemi di traduzione *speech-to-speech* presentano vantaggi quali il fatto di poter riflettere le informazioni non verbali del testo orale L1 nell'output vocale, portando anche a una riduzione dei tempi e degli errori (dovuta alla riduzione del processo di elaborazione) (Hee-Jo, 2021).

Meta ha invece annunciato un importante passo avanti nella ricerca sull'intelligenza artificiale che potrebbe rivoluzionare il modo in cui le persone comunicano tra le varie lingue. L'azienda ha sviluppato una nuova *suite* di modelli chiamata *Seamless Communication*[10], in grado di tradurre il parlato tra più di cento lingue preservando la voce, le emozioni e lo stile dell'oratore. I modelli sono stati resi pubblici nel dicembre del 2023 insieme ai relativi articoli scientifici che mettono in luce dati e aspetti tecnici[11].

Il modello di punta della *suite* (chiamato *Seamless*) è il primo sistema disponibile al pubblico che permetta una comunicazione interlinguistica espressiva in tempo reale. Unisce le funzionalità di altri tre modelli (*Seamless Expressive*, *Seamless Streaming* e *SeamlessM4T v2*) in un sistema unificato in grado di gestire input e output sia vocali che testuali. In particolare, i tre sotto-modelli hanno le seguenti caratteristiche:

1. *Seamless Expressive* è il modello che si concentra sul mantenimento dello stile vocale e delle sfumature emotive della voce dell'oratore durante la traduzione tra le lingue. A differenza degli strumenti di traduzione esistenti che si affidano a sistemi di sintesi vocale monotoni e robotici, *Seamless Expressive* è in grado di generare un parlato più naturale e spontaneo che cattura le sfumature dell'espressione umana;
2. *Seamless Streaming* è il modello che consente di tradurre con una latenza minima (di soli due secondi). È il primo modello multilingue a garantire una velocità di traduzione così

[10] https://ai.meta.com/research/seamless-communication/ (Ultimo accesso: 14 giugno 2024)
[11] https://ai.meta.com/research/publications/seamless-multilingual-expressive-and-streaming-speech-translation/ (Ultimo accesso: 19 giugno 2024)

elevata in quasi cento lingue parlate e scritte. *Seamless Streaming* è in grado di gestire frasi lunghe e complesse, pur mantenendo una qualità e una precisione notevoli;

3. *SeamlessM4T v2* è il modello che funge da base per gli altri due modelli. Si tratta di una versione aggiornata del modello originale *SeamlessM4T*, pubblicato nel 2022. La nuova architettura offre una maggiore coerenza tra testo e parlato, garantendo una traduzione coerente e fedele al messaggio originale (Hyscaler, 2023).

1.7 La tecnologia di *eye-tracking* negli *interpreting studies*

Just & Carpenter (1980) hanno elaborato la *Eye-Mind Hypothesis*, secondo la quale ciò che viene fissato dall'occhio corrisponde a ciò che viene elaborato cognitivamente: sulla base di questa premessa, gli interpreti fissano con l'occhio ciò che stanno elaborando; ovviamente, il legame tra movimenti evidenti dell'occhio e processi cognitivi sottostanti non è deterministico e assoluto, in quanto i nostri occhi potrebbero talvolta osservare determinati stimoli mentre la nostra mente pensa ad altro (Hvelplund, 2014). Assumendo, pur con le dovute limitazioni, l'Eye-Mind Hypothesis, la tecnologia di *eye-tracking* permette delle misurazioni in tempo reale (Holmqvist et al., 2015) durante l'attività di interpretazione simultanea, pur garantendo oggettività e un'invasività limitata (Seeber, 2015), mediante un metodo affidabile che permetta la descrizione delle caratteristiche dell'attività cognitiva durante l'IS. Gli *eye tracker* forniscono dati sull'esatta posizione primariamente della pupilla (e, dunque, del coordinamento dello sguardo) e sulla sua dimensione; combinando queste informazioni con gli stimoli visivi e la dimensione temporale dei dati (per quanto un partecipante osserva uno stimolo prestabilito), è possibile ottenere informazioni relative alla distribuzione dell'attenzione (Chmiel, 2021, p. 458).

L'unità d'analisi tipica dei dati di *eye-tracking* è l'area di interesse (di seguito, AOI, dall'inglese, *area of interest*): essa consiste in una parte dell'input visivo che viene analizzato e può essere una parola, una frase, una *slide* o qualsiasi altra forma bidimensionale presente nell'input visivo (*Ibid*, p. 460). Un'altra unità di analisi comunemente utilizzata è il periodo di interesse (o *interest period*): esso consiste nell'unità temporale in cui vengono analizzati i dati di *eye-tracking* (*Ibid.*, p. 460). Le aree di interesse possono essere sia statiche (per testi, grafici, diapositive) che dinamiche (per un input video in movimento). Altre misure di *eye-tracking* frequentemente utilizzate sono il numero di fissazioni e la durata media delle fissazioni,

tendenzialmente quanto maggiore è il valore, più alto sarà lo sforzo dovuto all'elaborazione cognitiva (Holmqvist et al., 2015).

Di seguito, viene proposta una breve rassegna di alcune unità di analisi possibili mediante i dati di *eye-tracking* prevalentemente utilizzate per valutare l'elaborazione cognitiva con input visivi che prevedono testi scritti. La durata della prima fissazione (o *first fixation duration*) è il tempo inizialmente speso dall'occhio per fissare un'AOI (Liversedge et al., 1998). La durata dello sguardo (nota anche come *gaze duration*) è la somma di tutte le fissazioni effettuate in una regione fino a quando il punto di fissazione lascia la regione a sinistra o a destra (Liversedge et al., 1998, p. 58) e riflette l'elaborazione lessicale precoce. Il tempo di *go-past* (noto anche come durata del percorso di regressione o *go-past time*) è la somma di tutte le durate delle fissazioni dal momento in cui l'occhio fissa per la prima volta l'AOI al momento in cui l'occhio fissa una regione a destra dell'AOI. Ciò significa che il tempo di *go-past* include la durata dello sguardo e le fissazioni a sinistra (cioè le rifissazioni su stimoli precedentemente fissati). Questa misura riflette l'integrazione di una parola vista nel contesto della frase (Liversedge et al., 1998). Infine, il tempo totale di lettura (indicato anche come tempo totale di permanenza*, total dwell time,* tempo totale di sguardo o tempo totale di visione) è la somma di tutte le fissazioni sull'AOI, comprese quelle durante la rilettura e indica il carico cognitivo. Per descrivere l'attrattività di un'area di interesse e la *performance* in un'AOI specifica, gli indicatori più comunemente utilizzati sono: il *total dwell time* su quell'AOI (quindi il tempo di permanenza complessivo su quell'AOI), il *fixations count* di quell'AOI (quindi il numero di fissazioni su quell'area di interesse) e l'*average fixation duration* in quell'AOI (quindi la durata media delle fissazioni di quell'AOI) (Kekule et al., 2019, p. 432).

Seguendo l'excursus teorico di Yuan & Wang (2023), è possibile osservare tre principali ambiti di ricerca che sono stati approfonditi e applicati per valutare gli aspetti cognitivi durante l'IS: il primo mira ad approfondire i modelli di attenzione predominanti nel corso dell'elaborazione di canali diversi durante l'IS, il secondo si concentra sulle caratteristiche relative all'elaborazione del canale visivo e il terzo riguarda l'elaborazione dei *problem trigger*.

Il primo filone ha l'obiettivo di esplorare i modelli di attenzione predominanti durante l'elaborazione di canali diversi durante i compiti relativi all'IS. Seeber et al. (2020) hanno preparato una videoregistrazione con la trascrizione (come stimoli sperimentali) e durante l'esperimento gli interpreti partecipanti sono stati invitati a svolgere un'attività di IS in cui dovevano non solo ascoltare la videoregistrazione ma anche leggere la trascrizione (SIMTXT). Questo campione è stato paragonato a un altro campione di interpreti che dovevano svolgere

l'attività di lettura durante l'ascolto (o *reading while listening, RWL*). Gli studiosi hanno stabilito delle AOI attorno agli stimoli prestabiliti: una corrispondente alla frase precedente a quella contente lo stimolo, una contenente il sintagma con lo stimolo, una contenente il resto della frase e un'altra con la frase successiva a quella contenente lo stimolo. Sono stati analizzati i tempi medi di permanenza (o *mean dwell time*) dell'occhio nelle AOI prestabilite e le quantità di fissazioni; lo studio ha dimostrato che gli interpreti tendono a seguire il canale visivo nel compito di RWL, come indicato dalle quantità di fissazioni più alte osservate nel resto della frase, e tendono a guardare avanti rispetto al canale verbale. Tuttavia, quando eseguono l'attività di IS, gli interpreti tendono a seguire i canali verbali più di quelli visivi, poiché è stato osservato un chiaro ritardo visivo. I risultati hanno evidenziato un modello di elaborazione che non avvalora l'ipotesi dell'esistenza di un modello di elaborazione predittiva comune a tutti gli interpreti (Yuan & Wang, 2023, p. 3), a differenza di altri studi (Amos et al., 2022). Recentemente, le conclusioni di Seeber (2020) hanno avuto un più ampio respiro grazie allo studio di Longhui et al. (2022), nel quale è stato chiesto ai partecipanti di svolgere un'attività di IS con il testo del discorso originale mostrato sullo schermo. Per valutare la distribuzione dell'attenzione, gli autori hanno proposto il concetto di *ear-eye span* (o *EIS*) che viene definito come la differenza temporale tra l'input audio-visivo del discorso originale e l'input del testo scritto del discorso originale e hanno osservato un EIS medio di circa 4 secondi. Questo evidenzia che gli interpreti tendono a iniziare l'elaborazione del testo scritto nel canale visivo circa 4 secondi dopo l'ascolto del discorso originale mediante il canale auditivo.

Il secondo filone di ricerca si concentra, invece, sulle caratteristiche relative all'elaborazione del canale visivo e si chiede se gli interpreti seguono un *pattern* determinato nell'elaborazione delle varie forme delle informazioni visive (testi scritti, volto dell'oratore, gesti, ecc.). Seeber (2012) ha creato un esperimento in cui ha creato delle AOI per ogni informazione visiva: volto dell'oratore, gesti e diapositive contenenti numeri. Lo studio ha osservato una minor attenzione sui gesti e una maggior attenzione su faccia dell'oratore e diapositive; inoltre, nell'AOI relativa alle *slide* osserva che gli interpreti cercano i valori numerici quando li sentono mediante il canale auditivo: in questo senso, Seeber (2012) ritiene che le informazioni visive siano complementari alle informazioni provenienti dal canale verbale.

La terza area di ricerca riguarda i *problem trigger*. Numeri e nomi propri – secondo il modello degli sforzi di Gile (2009) – richiedono una maggior capacità di elaborazione e potrebbero causare problemi di gestione dell'attenzione. Per valutare l'elaborazione dei numeri, Korpal and Stachowiak (2018) hanno proposto uno studio in cui hanno preparato delle diapositive

contenenti i numeri presenti nel discorso e le informazioni chiave del discorso in forma sintetica. La durata media delle fissazioni (o *mean fixation duration*) è stata selezionata come indicatore affidabile per confrontare lo sforzo cognitivo necessario per elaborare i numeri e il contesto. La durata della fissazione più lunga (o *longer fixation duration*) è stata riscontrata nell'elaborazione di numeri: ciò indica che interpreti professionisti e in formazione hanno speso uno sforzo maggiore sui numeri rispetto al contesto. Curiosamente, i dati raccolti non avvallano la tesi dello *spillover effect* di Gile (2009): sebbene l'elaborazione dei numeri richieda un maggior sforzo cognitivo, non ha ripercussioni sull'elaborazione del contesto. Stachowiak-Szymczak & Korpal (2019) hanno successivamente spostato l'attenzione sul confronto dei *pattern* di elaborazione dei numeri tra gruppi di professionisti e di interpreti in formazione. I partecipanti sono stati invitati a interpretare simultaneamente un discorso con informazioni visive provenienti da diapositive contenenti numeri. I loro risultati hanno suggerito due conclusioni interessanti: 1) sebbene sia stato riscontrato un maggior numero di fissazioni sui numeri per gli interpreti professionisti rispetto ai non professionisti, la fissazione era relativamente più breve; 2) in generale, i non professionisti hanno ottenuto fissazioni e tempi di sguardo (o *gaze time*) più lunghi, il che indica che gli interpreti in formazione dedicano fissazioni più lunghe sui numeri, il che indica che gli interpreti in formazione dedicano maggiore sforzo all'elaborazione dei numeri. Chmiel et al. (2020) hanno manipolato tre tipi di informazioni (numeri, nomi propri e parole di controllo) per verificare l'incongruenza tra input audio e visivi nell'interpretazione simultanea. I risultati hanno evidenziato una maggiore accuratezza nella condizione di congruenza rispetto a quella di incongruenza.

Come osservato nella letteratura descritta, la descrizione dell'elaborazione cognitiva dei due nuovi strati del canale visivo che vengono proposti nelle due condizioni sperimentali del presente elaborato (la trascrizione automatica di Google Meet e l'interfaccia con i suggerimenti in tempo reale del CAI *tool* di SmarTerp) è interessante per i seguenti aspetti: 1) i due nuovi canali visivi proposti possono essere visti come nuove risorse multimodali che potrebbero comportare sfide cognitive aggiuntive; 2) l'incongruenza tra le informazioni verbali e visive – dovuta all'inaccuratezza dei sistemi di ASR utilizzati nei due rispettivi sistemi – potrebbe avere delle ripercussione significative sull'elaborazione del discorso; 3) la presentazione dei *problem trigger* non è statica, ma dinamica e in movimento.

1.8 Didattica delle tecnologie nella formazione delle interpreti

Carl & Braun (2018) evidenziano che la disponibilità di servizi di videoconferenza basati sul web o sul cloud, così come l'accesso a tali servizi su tablet e altri dispositivi mobili, solleva nuove questioni sulla fattibilità della traduzione attraverso questi sistemi, sostenendo al contempo che gli interpreti dovrebbero essere coinvolti nelle fasi di pianificazione e di implementazione di queste tecnologie. Pym (2011) ritiene che le università possano diversificare i curricula di formazione, mediante un'offerta più ampia che includa non solo le competenze tecniche relative alla traduzione/interpretazione, ma anche quelle informatiche relative all'acquisizione di abilità che mirino a una maggior familiarizzazione con le nuove tecnologie. Enríquez Raído (2011) pone particolare enfasi sullo sviluppo dell'alfabetizzazione informatica nella formazione dei traduttori e sottolinea l'importante ruolo che le competenze informatiche, in particolare quelle relative alla ricerca sul web, dovrebbero svolgere. Nel 2014, viene pubblicata una relazione dell'Unione Europea in cui si osserva che l'interprete di conferenza deve utilizzare sempre di più il computer e le sue funzioni di ricerca; in quanto, quando si accende il microfono, non c'è più tempo per cercare le informazioni su dizionari o enciclopedie (Enríquez Raído, 2011). Riconoscendo la mancanza di materiali sulle moderne tecnologie informatiche per la formazione dell'interprete di conferenza e la necessità di sostenere l'autoformazione degli interpreti, Rütten et al. (2020) hanno avviato un lavoro di ricerca per poter offrire dei materiali di supporto agli interpreti relativi all'uso delle tecnologie di RSI. Una parte fondamentale del lavoro dell'interprete di conferenza di oggi è l'uso dei *software* (specifici per l'interpretazione di conferenza) che permettono una gestione terminologica dei glossari in fase *in-booth*: *InterpretBank, Interplex, LookUp, Intragloss, Interpreter's help*.

Tarasenko et al. (2021) individuano alcuni concetti chiave che devono essere presi in considerazione nella formazione degli interpreti di conferenza:

1. La competenza tecnologica come una componente cruciale nei curricula di formazione degli interpreti. L'interazione tra le abilità umane e quelle del computer è una simbiosi che non ha alternative ed è funzionale al raggiungimento di nuovi livelli qualitativi relativamente all'organizzazione, l'implementazione e l'esecuzione dell'attività interpretativa stessa.
2. Lo sviluppo di competenze informatiche e tecnologiche come base per un cambiamento concettuale nella formazione degli interpreti di simultanea. Dati i compiti che l'interprete

simultaneo deve svolgere, la competenza informatico-tecnologica dell'interprete simultaneo sarà intesa come la sua capacità di ricevere ed elaborare il segnale (proveniente ad esempio dal nuovo canale visivo dovuto alla presenza dell'interfaccia del CAI *tool*) con l'input auditivo, grazie all'utilizzo dei necessari strumenti informatici e tecnologici.

3. Aspetti contenutistici della formazione informatica degli interpreti di conferenza. L'acquisizione di nuove conoscenze, competenze e abilità deve essere finalizzata a garantire modalità efficaci per svolgere le seguenti attività: (a) ricerca mirata di testi specifici di un settore sfruttando le risorse di Internet; (b) uso di strumenti per la creazione di glossari multilingue; (c) strutturazione del materiale terminologico utilizzando software; (d) creazione di glossari e basi terminologiche in formati specifici; (e) estrazione di terminologia specifica di un argomento da testi ad elevato tecnicismo; (f) uso di CAI specifici per imparare i termini che verranno utilizzati durante la conferenza; (g) controllo e trasmissione di *streaming* audio e video attraverso reti cablate e *wireless*; (h) connessione a sistemi di interpretazione simultanea RSI *network-based*; (i) uso di CAI per il supporto terminologico in fase *in-booth*; (l) uso della *console* di cabina simultanea.
4. Forme e modalità di formazione informatica per gli interpreti. La simulazione di *setting* professionali tipici di un interprete simultaneo in un ambiente di laboratorio dovrebbe essere la chiave per formare una competenza tecnologica di alto livello. Per questo, nello studio sopracitato sono stati definiti alcuni requisiti minimi per tutti i laboratori preposti alla formazione di interpreti di simultanea, in particolare: (1) la disponibilità di tecnologie di interpretazione simultanea (tradizionali, CAI, RSI); (2) la possibilità di esercitarsi utilizzando strumenti diversi a seconda della tecnologia; (3) la possibilità di esercitarsi con canali input e output che vengono cambiati durante l'esercizio (ad esempio, con esercitazioni multilingue che prevedano turni in *relay*); (4) la possibilità di interpretare utilizzando CAI *tool*; (5) la possibilità di lavorare in parallelo utilizzando diversi canali output (e quindi diverse lingue di arrivo); (6) la possibilità di interpretare con un unico canale in output da diverse lingue di partenza.

Le partecipanti del presente studio sperimentale provengono tutte dal Dipartimento di Interpretazione e Traduzione (DIT)[12] dell'Università di Bologna, presso il campus di Forlì; pertanto, di seguito vengono descritti i corsi dedicati alle tecnologie presso il DIT. Il Dipartimento ha infatti introdotto nel 2016 il corso di Metodi e Tecnologie per l'Interpretazione, che è distribuito sui due anni del corso di laurea magistrale in Interpretazione. Durante il primo

[12] https://dit.unibo.it/it (Ultimo accesso: 29 aprile 2024)

anno[13], vengono insegnate strategie di documentazione (prevalentemente elettroniche) per la preparazione di un incarico; pertanto, viene insegnata la costruzione di corpora elettronici semi-automatici con l'aiuto di *BootCat* e *AntConc* (per la consultazione semiautomatica). Nello stesso corso del primo anno, gli studenti apprendono gli strumenti per la creazione e gestione di glossari e *database* terminologici (*MultiTerm*) e *software* per la consultazione dei glossari in cabina (*Interplex* e *InterpretBank*). Durante il secondo anno[14], si rafforzano le competenze acquisite precedentemente circa l'uso di *software* per la gestione e consultazione di glossari in cabina (*InterpretBank*) e vengono presentanti nuovi *software* di riconoscimento vocale automatico (*DIT.tafono*[15] e *Dragon Naturally Speaking*), con un esame finale che prevede un *respeaking* (ossia il sottotitolaggio in tempo reale sfruttando l'ASR) utilizzando *Dragon Naturally Speaking*. Durante il secondo anno, viene organizzata una *mock conference*, in cui viene chiesto alle studentesse di utilizzare *InterpretBank* in cabina. Il DIT ha inoltre introdotto il corso di *Interpretazione Dialogica, Multimediale e a Distanza* con una durata trimestrale presso lo stesso corso di laurea magistrale in Interpretazione. Il corso prevede l'utilizzo in laboratorio di un ampio ventaglio di tecnologie per l'interpretazione (simultanea, dialogica e consecutiva): dalle piattaforme tradizionali, come Zoom, a quelle più innovative come le piattaforme di RSI di *SmarTerp* – con il relativo CAI *tool* – e *Interprefy* per l'IS oppure altre tecnologie che offrono un supporto durante l'interpretazione consecutiva (come la *Smart Pen*).

1.9 La particolarità della combinazione linguistica italiano-spagnolo

L'approfondimento della contrastività tra italiano e spagnolo ha portato a numerose pubblicazioni, tra le quali l'analisi di Calvi (2003), gli atti dei convegni dell'Associazione Ispanisti Italiani (AISPI)[16], vari manuali di didattica della lingua spagnola per italofoni (Barbero et al., 2018) e manuali specifici per l'interpretazione tra l'italiano e lo spagnolo, come il manuale curato da Russo (2021) e quello di Russo (2012).

[13] https://www.unibo.it/it/studiare/dottorati-master-specializzazioni-e-altra-formazione/insegnamenti/insegnamento/2023/493148 (Ultimo accesso: 24 aprile 2024)
[14] https://www.unibo.it/it/studiare/dottorati-master-specializzazioni-e-altra-formazione/insegnamenti/insegnamento/2023/433465 (Ultimo accesso: 24 aprile 2024)
[15] https://dittafono.ditlab.it/ (Ultimo accesso: 24 aprile 2024)
[16] http://www.aispi.it/ (Ultimo accesso: 24 aprile 2024)

Data l'affinità tra queste due lingue, in interpretazione simultanea ci può essere la tendenza a riprodurre una versione con struttura superficiale simile o addirittura identica a quella della lingua di partenza (questo è ancora più evidente quando aumenta la velocità d'eloquio), portando a una maggiore difficoltà di comprensione da parte dell'ascoltatore a causa dei vari fenomeni di interferenza linguistica (Bertozzi et al., 2021, p. 292). Le dissimmetrie tra il sistema linguistico della lingua italiana e quello della lingua spagnola possono essere sia a livello morfologico (forma delle parole) che a livello sintattico (ordine e combinazione delle parole) e richiedono inevitabilmente una rielaborazione della lingua di partenza per giungere a una versione nella lingua di arrivo che sia corretta e comprensibile (Bertozzi et al., 2021, p. 292). Come le autrici evidenziano, questa rielaborazione del testo di partenza (e conseguente riproduzione nella lingua di arrivo) può esplicarsi in livelli diversi (che dipendono dalla distanza tra la morfosintassi delle due lingue).

Quando la struttura superficiale dell'italiano può essere quasi interamente ricostruita basandosi su quella dello spagnolo, e viceversa, è possibile affermare che durante l'interpretazione simultanea le dissimmetrie morfosintattiche vengono trattate con *dominanza di elaborazione superficiale*. Questo significa che per comprendere il significato della frase, si opera principalmente a livello di analisi della forma e del significato delle parole (livello linguistico) (Bertozzi et al., 2021, p. 293). Quando la struttura superficiale in spagnolo o italiano (cioè la forma e l'ordine delle parole) non fornisce un accesso immediato alla comprensione e alla traduzione, si verifica la necessità di una *dominanza di elaborazione profonda*. Questo significa che si opera a un livello semantico più profondo, spesso basandosi sul contesto, per accedere al suo contenuto e poi proiettarlo in una struttura superficiale nella lingua di arrivo, che in molte occasioni può essere completamente diversa (Russo, 2012). Si utilizza il concetto di *dominanza* per entrambe queste categorie generali per sottolineare che l'*elaborazione* avviene sia a livello semantico profondo che a livello superficiale contemporaneamente. Inoltre, l'interprete si avvale di procedure intermedie, come la ricerca di parole ed espressioni idiomatiche nei repertori lessicali in spagnolo o italiano (conservati nella sua memoria) (Bertozzi et al., 2021, p. 293).

1.9.1 Dissimmetrie morfosintattiche a dominanza di elaborazione superficiale

1.9.1.1 Strutture funzionalmente equivalenti

Due strutture si definiscono come *funzionalmente equivalenti* quando i loro costituenti sono legati allo stesso modo e – dunque – appartengono alle stesse categorie grammaticali; pertanto, nella trasposizione da una lingua all'altra, la parola viene sostituita con un'altra della medesima

categoria grammaticale che sia conforme alla morfosintassi del sistema linguistico della lingua di arrivo (Russo, 2012). Ne è un esempio l'espressione spagnola *en resumen* (preposizione + sostantivo) che può essere tradotta preservando la preposizione e sostituendo adeguatamente il sostantivo (*in sintesi*, ma anche *quindi*, *in breve*, *per riassumere*, *riassumendo*, ma non *in riassunto*). Oppure l'espressione correlativa *tanto…quanto…* che ha relativa struttura equivalente funzionale in spagnolo: *tan… como*.

Russo (2012, p. 63 - 81) ha riscontrato numerose strutture funzionalmente equivalenti, eccone alcuni esempi:

no ha hecho sino aumentar	non ha fatto che aumentare (non ha fatto altro che aumentare, è sempre/ senz'altro aumentata, ecc.)
con toda modestia	in tutta modestia (a mio modesto avviso, mi permetto umilmente di, ecc.).
igual…que	uguali… rispetto a quelli (allo stesso modo… di quelli, equiparare…e, analogamente a, ecc.)
como no podía ser menos	come non poteva essere altrimenti
aún están sin + infinito	sono ancora da + infinito (non sono state ancora + infinito)
extensa y detalladamente	estesamente e dettagliatamente (in modo esteso e dettagliato)
en lo que se refiere	in ciò che concerne (in merito a, per quanto riguarda, riguardo a)

1.9.1.2 Strutture deficitarie ed eccedentarie

Si definiscono *strutture deficitarie* ed *eccedentarie* quelle strutture che risultano particolarmente problematiche nella combinazione italiano-spagnolo, in quanto – pur mantenendo la maggior parte dei costituenti nello stesso ordine con una corrispondenza biunivoca – almeno uno dei costituenti non ha riscontro nella struttura appartenente all'altro sistema linguistico (Russo, 2012), ecco un esempio (*Ibid.*, p. 85):

‘insistimos’ ‘en’ ‘que’

| | |

‘insistiamo’ ‘ø’ che

In questo caso, l’interprete deve necessariamente optare per una *trasformazione lessicale* (“affermiamo/ sottolineiamo/ insistiamo che”) oppure per una *trasformazione morfosintattica* e *lessicale* (*insistiamo sul fatto*/*nel ribadire che*). Un esempio ben illustrativo è il caso della proposizione relativa con funzione specificativa, come nella frase: “El problema *que* tienen planteado muchos de esos países con sus ingentes deudas externas requiere una solución política”; sebbene la struttura “il problema *che hanno* quei paesi con un ingente debito estero richiede una soluzione politica” sia grammaticalmente e contenutisticamente corretta, risulta più chiara e intellegibile la struttura “il problema *dei* paesi” (Russo, 2012). Altre strutture identificate da Russo (2012, p. 81 - 118) sono:

con solo cambiar	Solo cambiando (cambiando unicamente/ solamente, ecc.)
en el supuesto de que	nel caso in cui (se, qualora, ecc.)
instar a que	richiedere che (sollecitare affinché/ perché, ecc.)
aunque solo sea porque	anche solo perché (se non altro perché, sebbene solo perché, anche se solo perché)
por ser	per il fatto di essere (in quanto è, per il fatto stesso di essere, a causa del fatto che, ecc.)
instar a que	richiedere che (sollecitare affinché)

1.9.1.3 Dissimmetrie della sintassi verbale

I due sistemi linguistici presentano una distribuzione di modi e tempi verbali con scarse differenze e, dunque, tendenzialmente non sono frequenti fenomeni di interferenza significativi in interpretazione simultanea (Russo, 2012). L’uso del congiuntivo, dell’indicativo, del gerundio e delle perifrasi verbali presenta alcune dissimmetrie che talvolta inducono a errori di traduzione. Un esempio è l’uso del congiuntivo nelle subordinate temporali con valore temporale di futuro (ad esempio, “cuando mi madre llegue a casa” → “quando mia mamma arriverà a casa”). Relativamente alle perifrasi verbali (più frequenti in spagnolo che in italiano), esse sono formate da due forme verbali (una esplicita e una implicita) che vengono trattate come un sintagma unico. Russo (2012) propone questi esempi: (a) il futuro perifrastico che esprime

un'azione immediata: *ir a* + infinito; (b) la perifrasi che esprime l'aspetto di durata del verbo: *seguir* / *continuar* + gerundio / participio passato / aggettivo.

Ecco alcuni esempi tratti da Russo (2012, p. 118 - 132):

se ha sometido (pasiva refleja)	è stata sottoposta
ir + *gerundio* (ad es. ir cediendo)	cede gradualmente (cede a poco a poco, ecc.)
seguir + gerundio (ad es. sigue hablando)	continua a parlare (parla ancora, ecc.)
si bien + indicativo	sebbene + congiuntivo
me he estado ocupando	mi sono occupato

1.9.2 Dissimmetrie morfosintattiche a dominanza di elaborazione profonda

Per *dissimmetrie morfosintattiche a dominanza di elaborazione profonda*, si considerano quelle strutture in spagnolo che hanno un senso non immediatamente accessibile, anche se costituite da parole simili a quelle italiane (Russo, 2012). In questo caso, l'elaborazione viene trasposta a un livello più profondo e non è, di conseguenza, possibile una trasposizione nella lingua di arrivo sfruttando la corrispondenza biunivoca tra le strutture superficiali dei sistemi linguistici. Questo livello di elaborazione richiede un maggior carico cognitivo che può essere alleggerito solo se l'interprete ha immediatamente delle traduzioni standard memorizzate (o automatismi) (Russo, 2012). Se l'interprete opera sotto condizioni di sforzo estremo, ad esempio a causa della velocità del discorso o della densità informativa del testo, le strutture spagnole possono causare un calco lessicale e sintattico nella resa italiana. Ciò comporta un aumento della probabilità di errori, specialmente se gli elementi della struttura spagnola sono foneticamente e semanticamente simili alle parole italiane, poiché potrebbero non essere adeguatamente elaborati semanticamente (Russo, 2012). Ne è un esempio la struttura: *desde* + sostantivo (ad esempio, *desde el realismo*). La traduzione letterale *dal realismo* sarebbe non chiara e non fruibile in lingua italiana; pertanto, è necessario rielaborare il concetto di *desde* giungendo a soluzioni come: *da un punto di vista realistico*, *seguendo un approccio realistico*, *in modo realistico*, *volendo essere realisti*, *con senso di realismo*, ecc.

Ecco alcuni esempi tratti da Russo (2012, p. 132 - 191):

se trata de que + sogg.	si tratta di far sì che + sogg.
no está de más hablar	non è superfluo parlare

a la mayor brevedad posible	il prima possibile (il più presto possibile, nel minor tempo possibile, al più presto)

1.9.4 Le sfide lessicali

L'italiano e lo spagnolo hanno elementi problematici che vanno naturalmente oltre le differenze morfosintattiche e che riguardano il livello lessicale; infatti, anche il lessico può portare a una scarsa comprensione, calchi e interferenze (Russo, 2012). In questa combinazione linguistica, ci sono infatti parole che hanno una radice in comune ma con significato diverso, a causa della diversa evoluzione semantica (*falsi amici*). Numerosi studi hanno messo in luce le potenziali insidie per l'interprete a livello lessicale, ad esempio gli studi di Fusco (1990), Spinolo (2018), Russo (2019). Fusco (1990, p. 94) afferma che i *paronimi* (ossia quelle parole che hanno elementi morfologici o fonetici simili) sono fonte di numerosi errori nella combinazione italiano-spagnolo. In particolare, Fusco (1990) suddivide i paronimi in quattro categorie: paronimi con connotazione diversa, paronimi classici, paronimi meno noti e paronimi *double-edged* (ossia con doppio significato). Ecco alcuni esempi tratti da Fusco (1990):

Paronimi con connotazione diversa

existencias	anche: merci invendute, scorte di magazzino
vecino	anche: abitante
pronóstico	anche: previsione
aferrarse	anche: aggrapparsi
justamente	anche: proprio

Paronimi classici

achacar	attribuire
asesor	consigliere, consulente
ilusión (me hace ilusión)	voglia, piacere

Paronimi meno noti

genio ("tiene genio")	irascibile
oposición	concorso
nómina ("estar en nómina")	(far parte dell') organico; busta paga

Paronimi meno noti

abandonado	trasandato
beneficios	utili
compromiso	impegno

1.9.5 Una chiave per la risoluzione delle dissimmetrie linguistiche: l'anticipazione

Lederer (1982) ritiene che l'anticipazione permetta la previsione del senso della frase, consentendo all'interprete un risparmio delle risorse preposte all'attenzione e permettendo un minor sforzo di analisi del testo orale della lingua di partenza. Secondo Ilg (1978), l'anticipazione porta, invece, a un'accelerazione della comprensione mediante la costruzione di significati ipotetici e mediante l'attivazione della rete concettuale pertinente. Se, a causa della saturazione della memoria dovuta a un sovraccarico, l'interprete inizia un'esposizione che si rivela confutata dalle occorrenze successive, deve ricorrere a una correzione che potrebbe portare a una perdita dei segmenti successivi presenti nel discorso (Russo, 2012, p. 43). L'anticipazione è possibile mediante fattori intralinguistici/ contestuali (collocazioni di verbo-complemento, in cui la presenza di una parola determinata ne comporta necessariamente una seconda, ad esempio *redigere un bilancio*) e fattori extralinguistici/ situazionali (ad esempio le formule linguistiche standardizzare che caratterizzano le relative tipologie testuali, come le formule di inizio o chiusura dei lavori) (Bendazzoli, 2010). Gile (2009) propone la stessa distinzione tra anticipazione linguistica e anticipazione extralinguistica: la prima è relativa alla conoscenza della normatività sintattica e lessicale di una lingua (pertanto la comprensione linguistica è basata su inferenze probabilistiche) e la seconda è relativa alla conoscenza del contesto comunicativo in cui ha luogo il discorso (Russo, 2012, p. 44). Considerando che i contenuti presenti in memoria vengono associati a delle espressioni linguistiche, l'IS (che consiste nella trasposizione interlinguistica di contenuti mentali), è agevolata se l'interprete condivide con l'oratore un gran numero di esperienze cognitive, in modo che la continuità cognitiva (quindi, la comprensione) non venga messa alla prova e il reperimento lessicale risulti più immediato (Russo, 2012, p. 44). Il consolidamento di automatismi mnemonici (basati sull'immagazzinamento di conoscenze qualitativamente collegate per facilitare i processi inferenziali e di diffusione dell'attivazione), l'attivazione di associazioni automatiche e preferenziali di termini o sintagmi cristallizzati (ossia di automatismi morfosintattico-lessicali) e l'attivazione di associazioni automatiche tra significati tipici ed espressioni relative (automatismi semantico-cognitivi) permettono di risparmiare la maggior quantità di energie

consentendo un'elaborazione più profonda del testo che, al contempo, lasci spazio a un processo traduttivo che preveda la scelta di un registro e uno stile adeguati alla lingua di arrivo (Russo, 2012, p. 44). La teoria neurolinguistica sull'interpretazione simultanea di Paradis (1994) chiarisce che l'IS prevede due strategie possibili nella trasposizione dalla lingua di partenza alla lingua di arrivo: una basata sulla concettualizzazione (che riguarda la rievocazione del concetto non collegato a una forma linguistica, mediante una rappresentazione mentale *meaning-based* o *top-down* dello stesso) e un'altra basata sull'equivalenza lessicale o strutturale (che implica l'applicazione di regole da un elemento linguistico della lingua di partenza al suo equivalente nella lingua di arrivo a livello fonologico, morfologico, sintattico e lessicale, mediante un processo *bottom-up*) (Paradis, 1994, p. 328 - 329). Come chiarisce Russo (2012, p. 45), gli interpreti professionisti hanno non solo una conoscenza linguistica implicita, ma anche un'abbondante conoscenza metalinguistica esplicita (ossia una forma di conoscenza e di associazione interlinguistica appresa consapevolmente e disponibile al bisogno) e maggiore è l'uso degli equivalenti traduttivi, più veloce sarà la formazione di automatismi che alleggeriscono il carico cognitivo nell'elaborazione del messaggio.

CAPITOLO 2. Metodologia dello studio sperimentale

Questo capitolo illustra la metodologia adottata nel presente studio sperimentale, mettendone in luce prevalentemente gli aspetti pratici e teorici sottostanti. Inizialmente verranno chiariti gli obiettivi dello studio, la scelta dell'approccio di ricerca utilizzato e il *design* dello studio sperimentale. Successivamente, verranno presentati i dettagli dei materiali utilizzati nella fase sperimentale, descrivendone le caratteristiche. Dopo aver descritto il criterio di selezione delle partecipanti, il profilo del campione e il protocollo utilizzato, il capitolo presenta le procedure adottate per la raccolta e la pulizia dei dati e i criteri dell'analisi.

2.1 Obiettivi e quesiti di ricerca

Il presente studio sperimentale desidera mettere in luce le differenze nell'usabilità di due strumenti in cabina di simultanea, in particolare il CAI *tool* di SmarTerp e la trascrizione di Google Meet (generata automaticamente), cercando di evidenziarne la diversa utilità e i possibili impatti sull'accuratezza e la fluidità nelle rese di interpreti in formazione. Per meglio definire gli obiettivi dello studio, sono state definite le seguenti domande di ricerca:

1. In che modo il CAI *tool* di SmarTerp e la trascrizione automatica di Google Meet influiscono sull'accuratezza e la fluidità della resa di interpreti in formazione?
2. Quanto risultano soggettivamente efficienti/soddisfacenti i due strumenti a interpreti in formazione?
3. In che modo vengono utilizzati questi due strumenti da interpreti in formazione durante l'IS?
4. Quali prospettive didattiche emergono dalle opinioni di interpreti in formazione sui due strumenti utilizzati?
5. Qual è l'impatto delle due tecnologie in esame sulla distribuzione dell'attenzione sull'input visivo?

2.2 Approccio sperimentale

Per rispondere ai quesiti di ricerca illustrati precedentemente, alle partecipanti sono stati proposti due testi da interpretare in modalità simultanea: uno con la trascrizione automatica di Google Meet e l'altro con il CAI *tool* di SmarTerp. Nel corso dello studio sono stati raccolti dati sulla percezione, la *performance* e il movimento oculare delle partecipanti mediante metodi sia qualitativi che quantitativi. Lo studio non vuole essere esclusivamente *product-oriented* (Bevan et al., 1991), ma prevalentemente *process-oriented*, e quindi focalizzato sull'analisi dei processi inerenti all'attività di interpretazione simultanea svolta da interpreti in formazione e non sulla valutazione dei prodotti, in quanto il campione utilizzato è di interpreti in formazione e non di interpreti professionisti. Pertanto, il contributo sperimentale non vuole mettere a confronto i due strumenti per decretarne il migliore. La prospettiva adottata è infatti prevalentemente didattica e, quindi, il paragone è funzionale alla comprensione delle attività cognitive relative all'elaborazione del materiale fornito dai due strumenti durante l'IS. Per cercare di catturare la complessità della diversa interazione interprete-strumento, è stato utilizzato un approccio trasversale (o *mixed-method*) in cui sono stati combinati metodi e fonti di dati provenienti da tutto lo spettro quantitativo-qualitativo, come approfondiscono Creswell & Creswell (2017, p. 294 - 324), per andare verso una comprensione più affidabile e olistica possibile grazie a una molteplicità di prospettive utilizzate (che emergono nei dati qualitativi), pur avendo dei valori di riferimento (quantitativi) sistematizzati e paragonabili per potenziali studi futuri. Come già Frittella (2023, p. 78) afferma, questo approccio di ricerca pone sfide significative per il ricercatore, tra le quali la necessità di raccogliere una gran quantità di dati e la gestione di un'attività di ricerca che prevede una complessità di flussi di dati quantitativi e qualitativi che devono essere elaborati in modo organico e preciso (Creswell & Creswell, 2017, p. 298).

2.3 *Design* dello studio sperimentale

Per rispondere ai quesiti di ricerca, è necessario adempiere alcuni requisiti di base, per poter irrobustire la validità dell'esperimento: (a) la selezione di utenti rappresentativi del gruppo *target* dei due strumenti (quindi, studenti e studentesse di interpretazione di conferenza con una familiarità minima con le tecnologie); (b) l'utilizzo dei due strumenti in un contesto di

interpretazione simultanea di due discorsi monologici; (c) l'utilizzo degli strumenti presentati eseguendo compiti rappresentativi di quelli che svolgerebbero normalmente con il supporto dello strumento (pertanto, è stata necessaria una speciale strategia di progettazione del discorso); (d) l'adozione di una procedura standard da proporre ad ogni partecipante (in quanto ogni sessione ha avuto luogo in momenti diversi).

Sulla base di queste esigenze, il protocollo sperimentale prevedeva la somministrazione di: (a) due video da interpretare in modalità simultanea; (b) due attività di analisi retrospettiva (una dopo ogni attività di simultanea); (c) due questionari sull'usabilità degli strumenti (uno dopo l'uso di ciascuno strumento) e (d) un questionario finale funzionale alla profilazione delle partecipanti.

Le attività di simultanea sono state utili a raccogliere dati quantitativo-qualitativi sulle *performance* delle interpreti in formazione nelle due rispettive condizioni sperimentali. I questionari sull'usabilità degli strumenti hanno avuto l'obiettivo di raccogliere dei valori quantitativi sulla percezione dello strumento e della sua usabilità. Le retrospezioni sono state utili ad avere dati qualitativi sulla percezione degli strumenti e sulle difficoltà riscontrate durante l'attività di simultanea, permettendo un approfondimento dell'analisi di *performance* e percezione, con dati qualitativi che mirano ad andare verso una comprensione più ampia, anche in vista di una mirata prospettiva didattica. Oltre a questi strumenti di analisi, è stato utilizzato un *eye tracker* durante entrambe le prove di IS. Nella tabella seguente è stato approfondito l'approccio trasversale utilizzato, specificando per ogni fonte di dati il tipo di dato raccolto.

	IS con trascrizione automatica di Google Meet/CAI *tool* di SmarTerp	*Eye-tracking*	Questionario *post-task*	Retrospezione *post-task*
Tipologia di dato raccolto	Dati quantitativi e qualitativi sulla *performance*	Tracciamento del movimento oculare durante l'attività *in-booth* con dati qualitativi e quantitativi	Dati quantitativi sulla percezione delle partecipanti	Dati qualitativi
Descrizione del dato	Rese delle interpretazioni in modalità simultanea di due video preregistrati con il supporto di SmarTerp/trascrizione automatica di Google Meet	Dati di *eye-tracking*	Questionario	Retrospezione *post-task*
Materiali utilizzati per la raccolta dei dati	Video dei discorsi originali su computer di laboratorio e microfono esterno per la registrazione della *performance*	*Eye tracker* e relativo *software* per la visualizzazione e l'elaborazione dei dati (*EyeLink Data Viewer*)	Questionario cartaceo (questionario in Scala Likert con 7 punti per ogni *item*)	Retrospezione dell'attività di IS svolta ripercorrendo il video originale (con audio originale) e con la sovrapposizione del movimento oculare durante l'intera IS su *EyeLink Data Viewer*
Analisi dei dati	a) Quantitativi: *success rate nei task;* b) Quantitativo-qualitativi: valutazione di intelligibilità e informatività di ogni *cluster*.	a) Quantitativi: *dwell time* nelle AOI preconfigurate nel periodo di interesse corrispondente all'intero video; b) Quali-quantitativi: descrizione del movimento oculare su 51 stimoli prestabiliti nei discorsi.	L'accordo delle partecipanti con affermazioni relative alla soddisfazione generale e ai criteri di usabilità pragmatica.	Descrizione qualitativa di difficoltà ed elementi di interesse sull'esperienza delle partecipanti nel corso dell'attività di IS.

Tabella 1 Organizzazione della tipologia dei dati per attività svolta.

2.4 Materiali sperimentali

2.4.1 Caratteristiche dei discorsi

Ai fini di un'analisi delle rese proficua e ricca di materiale rappresentativo dell'interazione interprete-strumento, è necessaria la creazione di *task* che portino sufficientemente le interpreti

ad utilizzare il CAI *tool* di SmarTerp o la trascrizione automatica di Google Meet. Per questo motivo nel presente studio sperimentale i discorsi sono stati appositamente redatti per creare le condizioni necessarie per delle osservazioni significative e complete e per poter mettere in relazione comparativa i due strumenti con due discorsi paragonabili in termini di difficoltà e di distribuzione degli stimoli.

I discorsi sono stati elaborati seguendo il modello proposto da Frittella (2023, p. 82). Secondo questo modello, le caratteristiche dei discorsi devono soddisfare sei criteri: (1) devono includere tutte le categorie di *problem trigger* per cui è stato ideato e sviluppato il CAI *tool* di SmarTerp; (2) i *problem trigger* devono essere sufficientemente stimolanti per tutte le partecipanti in modo da aumentare la probabilità di consultazione degli strumenti durante l'IS; (3) i *problem trigger* nel discorso originale e la loro distribuzione devono variare, creando – dunque – dei *task* di diversa complessità; (4) il contenuto non deve essere così complesso e poco familiare da interferire con la consultazione degli strumenti; (5) la struttura dei discorsi deve essere logicamente chiara e i *task* contenenti gli stimoli devono alternarsi con dei passaggi discorsivi per evitare che le difficoltà riscontrate in un *task* abbiano ripercussioni sul *task* successivo (portando a una confusione nell'analisi); (6) altre caratteristiche del parlato (ad esempio, qualità audio, complessità sintattica, velocità, ecc.) dovrebbero evitare che fattori diversi dai *problem trigger* confondano l'analisi. Metodologie simili per la redazione dei discorsi sono state adottate da Seeber & Kerzel (2012), Prandi (2017) e Frittella (2023). Sulla base di questi principi, sono stati redatti due discorsi contenenti tutti i *problem trigger* per cui è stato sviluppato il CAI *tool* di SmarTerp (acronimi, numerali, entità denominate, termini specialistici), inserendo anche altri elementi problematici nella combinazione linguistica italiano-spagnolo, quali le dissimmetrie. Ogni *item* inserito è considerato sufficientemente difficile da stimolare l'interprete a guardare lo strumento; pertanto, acronimi comuni (come *UE*), nomi propri molto conosciuti (come *Pedro Sánchez*), numerali a una cifra (come *4*) e termini specialistici comunemente noti (come *algoritmo*) sono stati esclusi. Per ogni categoria di *problem trigger*, sono state create almeno quattro condizioni per creare *task* con complessità variabile. Sulla base dello studio di Frittella (2023), una frase di 20 parole contenente un numerale e nessun altro fattore di complessità viene considerata meno problematica di un passaggio testuale di 60 parole contenente otto numerali e termini specialistici. Sulla base di queste premesse sono stati definiti i *task* contenenti gli stimoli e di seguito ne viene offerta una breve disamina.

Nei *task numerale isolato* (NI), *dissimmetria superficiale isolata* (SI), *entità denominata isolata* (EI), *dissimmetria profonda isolata* (PI), *acronimo isolato* (AI), *sfida lessicale isolata* (LI),

tecnicismo isolato (TI), il *problem trigger* viene presentato in una proposizione semplice, ossia in una proposizione di circa 30 parole e con sintassi e legami logici semplici (quindi, senza nessun altro *problem trigger*). Nel *task referente complesso + numerale* (RN), si osserva un referente complesso (acronimo/ entità denominata/ termine specialistico) che viene accompagnato da un numero a cinque cifre (ordine di grandezza: miliardi). Nel *task tecnicismi + dissimmetria profonda* (TP), due tecnicismi vengono inseriti in una frase semanticamente complessa con una struttura dissimmetrica che richiede un'elaborazione profonda del testo. Nel *task dissimmetria superficiale + dissimmetria profonda* (SP), si osserva un *task* complesso con una struttura che richiede dominanza di elaborazione superficiale e una che richiede dominanza di elaborazione profonda (in proposizione semplice). Nel *task sfida lessicale + dissimmetria superficiale* (LS), il *task* presenta una sfida lessicale e una struttura dissimmetrica che richiede dominanza di elaborazione superficiale in proposizione semplice. Il *task sfida lessicale + dissimmetria profonda* (LP) presenta una sfida lessicale e una struttura che richiede dominanza di elaborazione profonda in proposizione semplice. Il *task numerali + referente complesso + dissimmetria superficiale* (NRS) è un *task* complesso contenente tre numeri complessi, un referente complesso (acronimo/ entità nominale/ termine specialistico) e una struttura che richiede dominanza di elaborazione superficiale. Il *task numerali + sfida lessicale + dissimmetria profonda* (NLP) presenta 2 *cluster* contenenti ciascuno: (1) 1 numerale (ordine di grandezza: milioni); (2) 1 sfida lessicale e (3) 1 struttura che richiede dominanza di elaborazione profonda. Il *task unità informativa con numerali* (UN) è un *task* complesso che prevede un'unità informativa composta da: (1) un referente complesso; (2) un'unità di misura complessa e (3) 5 numerali semplici consecutivi (2 anni, 2 numerali semplici e 1 percentuale). Il *task cluster con numerali ridondanti* (NRI) prevede una parte del discorso numericamente densa costituita da 3 microunità informative e ogni microunità contiene: (1) un deittico spaziale e un deittico temporale (che rimangono gli stessi in tutte e tre le unità e si ripetono); (2) almeno un'unità di misura; (3) un numerale che cambia in ciascuna delle tre unità; (4) un referente che rimane lo stesso in tutti e tre i *cluster*. Il *task cluster con numerali non ridondanti* (NNR) è un *task* complesso e numericamente denso con le seguenti caratteristiche: (1) tre microunità informative; (2) tempo, luogo, referente, unità di misura e numerale cambiano in ogni microunità informativa; (3) o il referente o l'unità di misura devono essere complessi in ogni microunità.

I *task* inseriti nel discorso potrebbero essere classificati sulla base della loro diversa complessità, come di seguito:

1. *Task* a complessità bassa: NI, SI, EI, PI, AI, LI, TI sono i *task* con complessità ridotta e consistono in proposizioni di circa 30 parole caratterizzate da legami logici e sintattici semplici contenenti un solo *problem tirgger*.
2. *Task* a complessità media: RN, TP, SP, LS, LP sono *task* leggermente più difficili di quelli precedenti in quanto due *problem trigger* vi co-occorrono contemporaneamente.
3. *Task* a complessità elevata: NRS, NLP, UN, NRI, NNR sono *task* considerati come altamente complessi in quanto più di due *problem trigger* co-occorrono contemporaneamente in un passaggio testuale.

Sempre seguendo la metodologia di Frittella (2023), dopo aver definito concettualmente i *task*, è stato scelto un tema e un contesto comunicativo, in particolare è stato scelto come filo conduttore di entrambi i discorsi dell'esperimento il *Ministero spagnolo per la Transizione Ecologica e la Sfida Demografica*. Dopo aver circoscritto il tema, sono stati creati i diversi passaggi testuali contenenti i *task* precedentemente descritti. Dopo aver creato tutti i *task* contenenti gli stimoli, le *target sentence* redatte sono state inserite in un discorso in cui ogni *task* si alterna con passaggi testuali discorsivi. Entrambi i discorsi si aprono con un'introduzione di circa 45 parole che non contiene alcun *problem trigger* rilevante ai fini dell'analisi e l'introduzione ha una funzione di riscaldamento (*warm-up time*) per le interpreti. Ogni *task* è seguito da una *frase cuscinetto* (o *control sentence*) che ha la funzione di alleggerire il carico cognitivo in preparazione al *task* successivo (*spill over region*). Alla fine di ogni *control sentence* è stata inserita una pausa di circa 3 secondi (equivalente a un respiro profondo dell'oratrice). Ogni *frase cuscinetto* ha una lunghezza di circa 30 parole e non contiene alcun *problem trigger* o fattore di difficoltà. L'intero discorso (con la rispettiva analisi suddivisa in *task* e *control sentence*) è stato inserito in calce al presente elaborato, ma per esemplificare la struttura del discorso si fornisce di seguito una sequenza *task – control sentence*:

> *Task*: El primer gran reto, relativo a la transición energética, cuenta con un presupuesto de 7397 millones de euros.
>
> *Control sentence*: Con este presupuesto, el Gobierno de España quiere mostrar su firme determinación en su voluntad de asumir un papel de liderazgo en esta transición a nivel europeo. *[pausa di 3 secondi]*

Nella tabella seguente si osservano in forma sintetica i *task* inseriti nei discorsi, con la loro relativa abbreviazione:

Codice	Nome
NI	Numerale isolato
SI	Dissimmetria superficiale isolata
EI	Entità nominale isolata
PI	Dissimmetria profonda isolata
AI	Acronimo isolato
LI	Sfida lessicale isolata
TI	Tecnicismo isolato
RN	Referente complesso + numerale
TP	Tecnicismi + dissimmetria profonda
SP	Dissimmetria superficiale + dissimmetria profonda
LS	Sfida lessicale + dissimmetria superficiale
LP	Sfida lessicale + dissimmetria profonda
NRS	Numerali + referente complesso + dissimmetria superficiale
NLP	Numerali + sfida lessicale + dissimmetria profonda
UN	Unità informativa con numerali
NRI	*Cluster* con numerali ridondanti
NNR	*Cluster* con numerali non ridondanti

Tabella 2 Sintesi dei nomi dei singoli task

Funzione del cluster	Nome *task*	Definizione del *task*	Testo	Stimoli isolati
NRS	Numerali + referente complesso + dissimmetria superficiale	*Task* complesso con 3 numeri complessi, referente complesso (acronimo, entità nominale o termine specialistico) e struttura che richiede dominanza di elaborazione superficiale.	El presupuesto asciende en su conjunto a 10 195 millones de euros, de los que 5817 corresponden a recursos del presupuesto nacional y 4378 corresponden a recursos del plan España Puede, tal y como les explicaré extensa y detalladamente.	10 195 millones, 5817 millones, 4378 millones//plan España Puede // extensa y detalladamente

Tabella 3 Esempio di un task inserito nel discorso A.1.

I discorsi sono stati letti e corretti da una collega con spagnolo come lingua A non a conoscenza della struttura sottostante al discorso, che ha fornito alcuni suggerimenti per migliorare la fluidità e la correttezza grammaticale degli stessi. I discorsi dell'esperimento sono stati letti, seguendo il copione fornito, da una professoressa spagnola del Dipartimento di Interpretazione e Traduzione dell'Università di Bologna. I discorsi sono stati letti con una velocità media in modo da simulare una velocità realistica senza compromettere tuttavia la resa delle interpreti in modo significativo.

Entrando nei dettagli di ciascun discorso, il primo discorso (di seguito, A.1) è un discorso che simula la presentazione del bilancio relativo al *Ministero spagnolo per la Transizione Ecologica e la Sfida Demografica* per l'anno 2023 e risulta avere i seguenti valori relativi alla sua leggibilità:

Leggibilità del testo		
Valore	Indice	Difficoltà
55.43	Fernández Huerta	Abbastanza difficile
37.62	Gutiérrez	Normale
50.88	Szigriszt-Pazos	Normale
50.88	INFLESZ	Abbastanza difficile
47.69	Leggibilità μ	Difficile

Tabella 4 Valori relativi alla leggibilità del discorso A.1

Il discorso A.1 contiene 1122 parole che sono state lette in 10 minuti e 22 secondi, pertanto con una velocità media di 112 parole al minuto. Osservando i valori relativi alla leggibilità del testo, è possibile affermare che si tratta di un discorso con una leggibilità abbastanza difficile.

Il secondo discorso (di seguito, B.1) è un discorso che simula un contesto comunicativo monologico in cui la *Ministra per la Transizione Ecologica e la Sfida Demografica* discute dettagli e obiettivi della legge sui cambiamenti climatici e la transizione energetica durante il dibattito parlamentare per chiederne la sua approvazione. Il discorso B.1 presenta i seguenti valori relativi alla sua leggibilità:

Leggibilità del testo		
Valore	Indice	Difficoltà
58.8	Fernández Huerta	Abbastanza difficile
38.98	Gutiérrez	Normale
54.27	Szigriszt-Pazos	Normale
54.27	INFLESZ	Abbastanza difficile
49.07	Leggibilità μ	Difficile

Tabella 5 Valori relativi alla leggibilità del discorso B.1

Il discorso B.1 contiene 1178 parole che sono state lette in 10 minuti e 59 secondi, pertanto con una velocità media di 108 parole al minuto. Anche in questo caso, osservando i valori relativi alla leggibilità del testo, è possibile affermare che si tratta di un discorso con una leggibilità abbastanza difficile.

2.4.2 Preparazione dei video

I due discorsi sono stati preregistrati circa un mese prima dall'inizio della raccolta dati. I video sono stati registrati con la seguente configurazione spaziale: volto dell'oratrice al centro dello schermo, sfondo bianco e inquadratura tagliata approssimativamente all'altezza del petto. All'oratrice è stato inoltre chiesto di non fare gesti durante la lettura dei discorsi, in quanto avrebbero potuto rappresentare un ulteriore elemento di distrazione per le partecipanti. L'intera configurazione dello spazio dell'esperimento persegue l'obiettivo di garantire un ambiente neutro e privo di ulteriori stimoli (oltre a quelli verbali); ad esempio, un'oratrice con il volto troppo vicino alla videocamera avrebbe potuto comportare una situazione comunicativa diversa rispetto ai *setting* monologici a cui sono abituate le partecipanti, i gesti avrebbero potuto distrarre le partecipanti, oggetti nello sfondo avrebbero comportato ulteriori elementi da dover essere presi in considerazione nell'analisi della distribuzione dell'attenzione. I video sono stati registrati con una videocamera LOGITECH C270 (HD – 3MP) e con un microfono esterno Blue Snowball ICE presso una delle cabine di interpretazione del Dipartimento di Interpretazione e Traduzione dell'Università di Bologna. I video sono stati registrati grazie al registratore online https://webcamera.io/es/ che ha consentito di registrare i video mantenendo

un’ottima qualità audio permettendo il download diretto di un file .mp4 (necessario per gli ulteriori passaggi relativi alla preparazione dei video per l’esperimento). Una volta ottenuti i due video in formato .mp4, è stato necessario procedere alla preparazione dei quattro video dell’esperimento: ogni discorso è stato infatti preparato in entrambe le condizioni sperimentali, in modo da randomizzare sia l’ordine delle condizioni che il discorso presentato in ogni condizione.

Per preparare i video nella condizione sperimentale che prevede la trascrizione automatica di Google Meet, sono stati necessari due computer (di cui uno con doppio schermo). Dal computer con un solo schermo è stata avviata una riunione su Google Meet ed è stato condiviso il video in formato .mp4 nella stessa riunione. Mentre sullo schermo di sinistra del computer con due schermi è stata aperta l’estensione Chrome SCRE.IO (https://scre.io/) per la registrazione dello schermo di destra (la riunione è infatti stata aperta sullo schermo di destra). Per la registrazione del video è stata seguita la seguente procedura: (1) è stato condiviso il video nella riunione di Google Meet dal computer con uno schermo; (2) dallo schermo di sinistra del computer con due schermi è stata avviata la registrazione dello schermo di destra; (3) è stato avviato il video contenente il discorso in formato .mp4 dal computer con uno schermo; (4) è stato registrato l’intero video; (5) a video ultimato, è stata interrotta la registrazione dello schermo; (6) è stato scaricato il video in formato file .WEBM; (7) è stato utilizzato un convertitore *online* per convertire il *file* .WEBM in *file* .mp4. Questa procedura è stata utilizzata per la registrazione di entrambi i video nella condizione sperimentale contenente la trascrizione automatica di Google Meet e in questo modo è stato possibile preparare dei video con un’interfaccia uguale a quella che un’interprete avrebbe nella realtà.

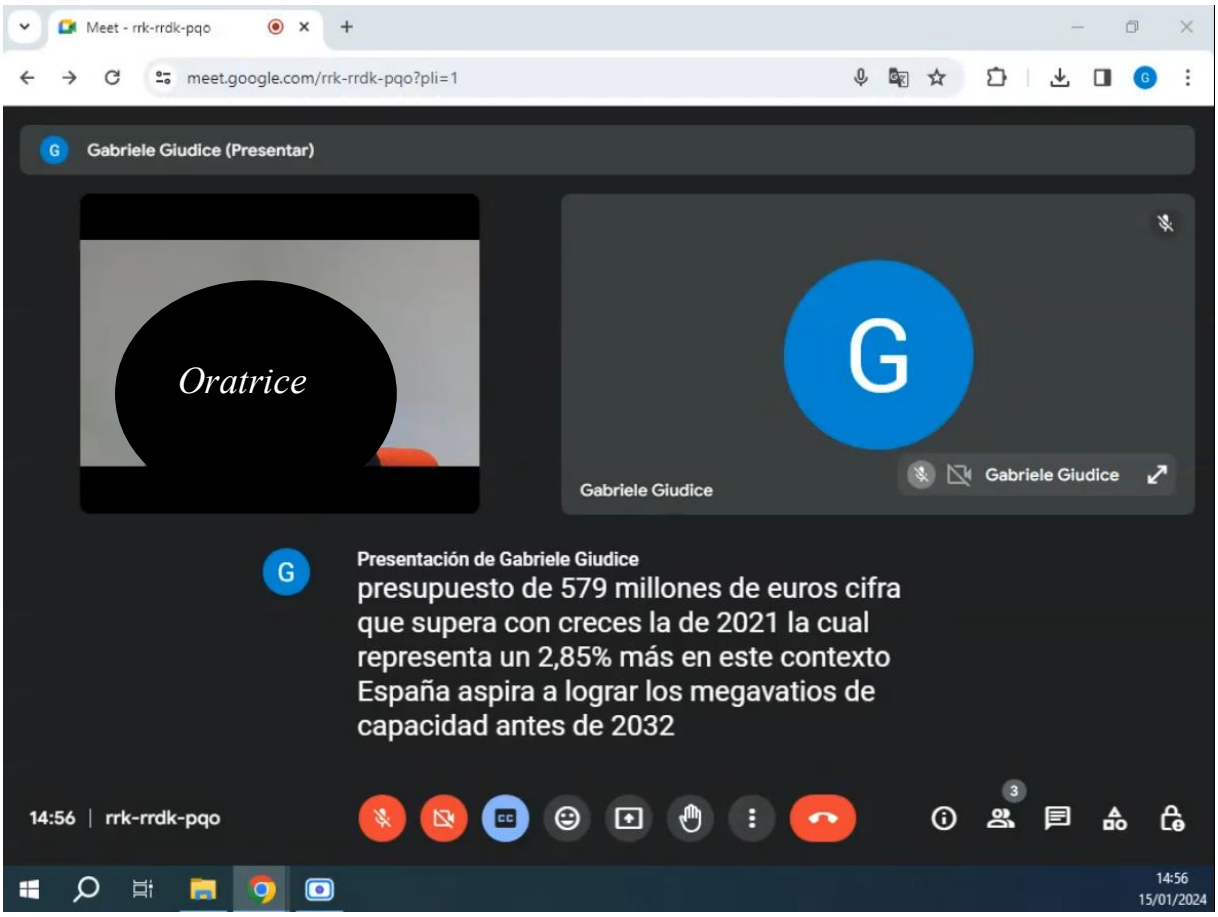


Figura 2 Esempio dell'interfaccia del video del discorso A.1 in formato .mp4 nella condizione sperimentale contenente la trascrizione automatica di Google Meet utilizzato nell'esperimento.

Per preparare i video per la condizione sperimentale che prevede il CAI *tool* di SmarTerp, la preparazione dei video è stata più laboriosa. In termini generali, per la configurazione delle sessioni SmarTerp con CAI è stato necessario preparare un glossario italiano-spagnolo e due corpora (uno in italiano e uno in spagnolo) per ogni video.

Per la configurazione della sessione con CAI per la preparazione del primo video contenente il discorso A.1 è stato preparato un glossario contenente 61 termini in italiano con le corrispettive traduzioni in spagnolo. Per la compilazione del glossario e l'individuazione delle equivalenze semantiche tra le due lingue è stato utilizzato prevalentemente EUR-Lex[17], ossia il portale *online* che permette di accedere al diritto dell'UE fornendo un accesso ufficiale e completo ai documenti giuridici dell'UE. EUR-Lex permette un'interrogazione rapida di una notevole quantità di testi paralleli, consentendo un'agevole ricerca terminologica. Nella compilazione del glossario sono stati inseriti tutti i termini potenzialmente problematici o relativi al tema

[17] https://eur-lex.europa.eu/homepage.html?locale=es (Ultimo accesso: 14 aprile 2024)

circoscritto, senza alcuna distinzione tra termini presenti nelle *control sentence* e quelli presenti nelle porzioni testuali contenenti gli stimoli.

Per la compilazione dei due corpora (uno in spagnolo e uno in italiano) è stato seguito il seguente criterio: ogni termine contenuto all'interno del glossario deve avere almeno 7 occorrenze all'interno del corpus e ogni entità denominata deve presentare anch'essa almeno 7 occorrenze all'interno del corpus. Seguendo questo criterio, per il discorso A.1 è stato compilato un corpus in italiano con la seguente grandezza:

Tokens	51726
Words	44523
Sentences	1364

Tabella 6 Grandezza del corpus in italiano per la preparazione del discorso A.1 nella condizione sperimentale con SmarTerp.

Le fonti prevalentemente utilizzate per la compilazione del corpus in italiano, oltre ad EUR-Lex, sono state: quotidiani online e Wikipedia[18].

Parallelamente, seguendo lo stesso criterio, per il discorso A.1 è stato compilato un corpus in spagnolo con la seguente grandezza:

Tokens	55352
Words	47980
Sentences	1468

Tabella 7 Grandezza del corpus in spagnolo per la preparazione del discorso A.1 nella condizione sperimentale con SmarTerp.

Le fonti prevalentemente utilizzate per la compilazione del corpus in spagnolo, oltre ad EUR-Lex, sono state: quotidiani online e Wikipedia.

Dopo aver preparato i due corpora e il glossario, grazie all'aiuto degli ingegneri del *software* SmarTerp, è stata possibile la registrazione del video definitivo per l'esperimento, in cui è stato condiviso nella sessione CAI il video in formato .mp4 ed è stata registrata l'interfaccia del CAI *tool* mediante registrazione dello schermo.

[18] https://es.wikipedia.org/ (Ultimo accesso: 18 aprile 2024)

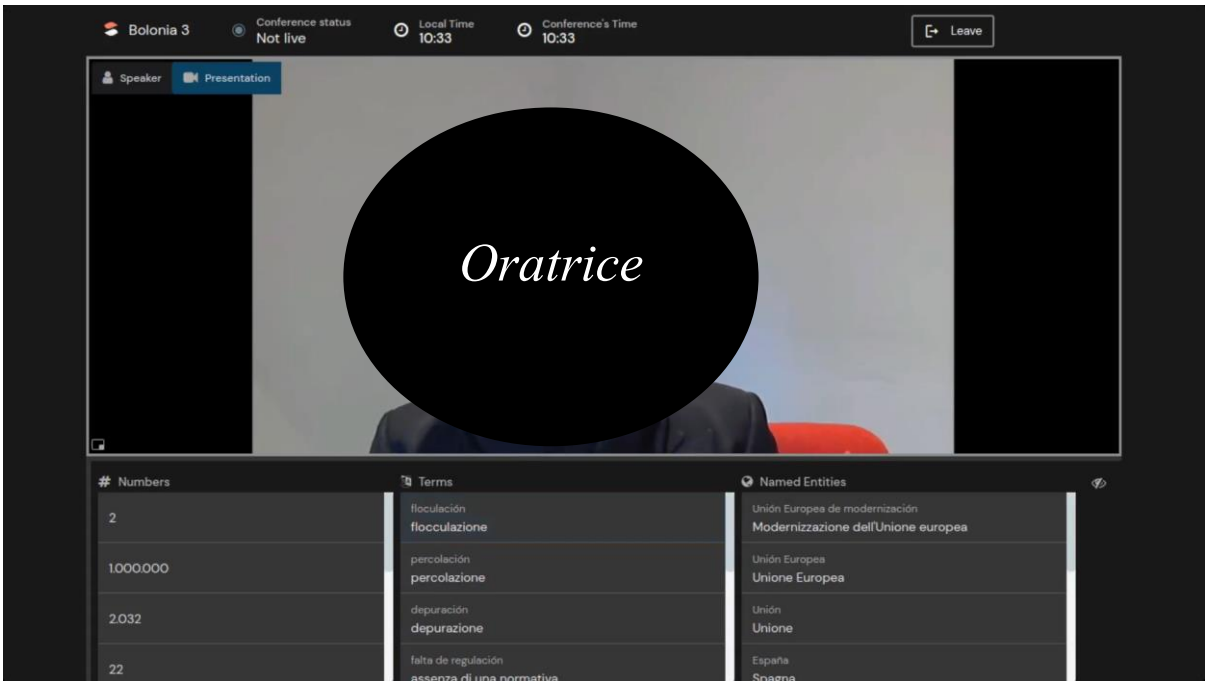


Figura 3 Interfaccia del video del discorso A.1 nella condizione sperimentale con SmarTerp.

Per la configurazione della sessione con CAI per la preparazione del video contenente il discorso B.1 è stata adottata una procedura analoga a quella utilizzata per la preparazione del discorso B.1. È stato infatti compilato un glossario contenente 61 termini in italiano con le corrispettive traduzioni in spagnolo. Le fonti e i criteri utilizzati per la compilazione del glossario e dei corpora sono stati gli stessi utilizzati per la redazione del glossario del discorso A.1. Anche in questo caso, per la compilazione dei due corpora è stato seguito il seguente criterio: ogni termine contenuto all'interno del glossario deve avere almeno 7 occorrenze all'interno del corpus e ogni entità denominata deve presentare anch'essa almeno 7 occorrenze all'interno del corpus.

Seguendo questo principio, è stato compilato un corpus in italiano con la seguente grandezza:

Tokens	47,548
Words	41090
Sentences	1289

Figura 4 Grandezza del corpus in italiano per la preparazione del discorso B.1 nella condizione sperimentale con SmarTerp.

Parallelamente, seguendo lo stesso criterio, per il discorso B.1 è stato compilato un corpus in spagnolo con la seguente grandezza:

Tokens	56232
Words	49054
Sentences	1552

Figura 5 Grandezza del corpus in spagnolo per la preparazione del discorso B.1 nella condizione sperimentale con SmarTerp.

Anche per il discorso B.1, dopo aver preparato i due corpora e il glossario, grazie all'aiuto degli ingegneri del software, è stata possibile la registrazione del video definitivo per l'esperimento, in cui è stato condiviso nella sessione con CAI il video in formato .mp4 ed è stata registrata l'interfaccia del CAI *tool* mediante registrazione dello schermo.

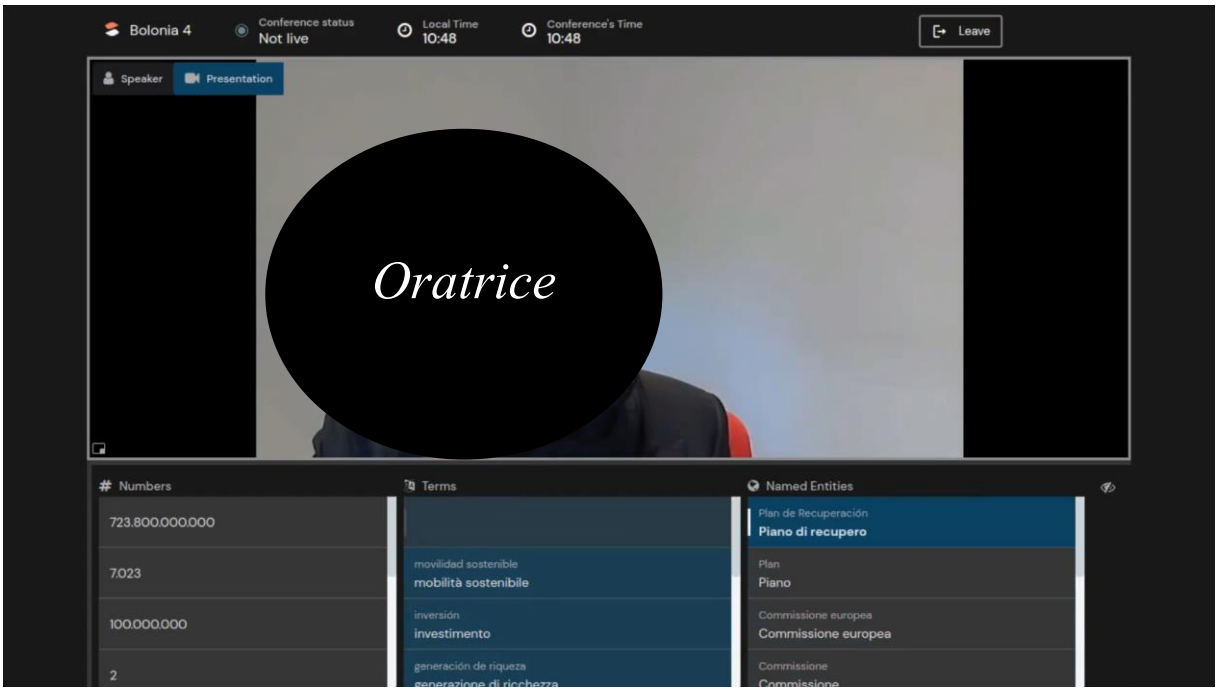


Figura 6 Interfaccia del video del discorso B.1 nella condizione sperimentale con SmarTerp.

2.4.3 *User Experience Questionnaire*

L'usabilità dello strumento non è determinata esclusivamente dalla *performance* delle utenti ma anche dal loro stato mentale durante l'utilizzo dello strumento (Frittella, 2023, p. 86). In particolare, è necessario raccogliere dei dati quantitativi che permettano di catturare non solo la soddisfazione delle interpreti in formazione ma anche la loro percezione di entrambi gli

strumenti con valori di riferimento quantificabili e paragonabili. Uno degli strumenti validati nella ricerca sull'usabilità è lo *User Experience Questionnaire* (di seguito, UEQ) (Schrepp et al., 2017). Il vantaggio di utilizzare strumenti validati è la possibilità di garantire una maggior validità allo studio sperimentale. L'obiettivo principale dell'UEQ è permettere una misurazione immediata e rapida dell'esperienza dell'utente, considerando aspetti relativi alla qualità pragmatica ed edonica (Schrepp et al., 2014, p. 383). Il questionario presenta 6 scale e 26 *item* totali. Ogni *item* prevede due aggettivi con significati opposti che sono agli estremi.

Esempi:

imprevedibile O O O O O O O *prevedibile*

ostruttivo O O O O O O O *di supporto*

Ogni *item* può essere valutato su una scala Likert a 7 punti. Le risposte a un *item* vanno quindi da -3 (completamente d'accordo con il termine negativo) a +3 (completamente d'accordo con il termine positivo); ogni *item* ha una risposta neutra (con valore 0). Metà degli *item* inizia con il termine positivo, il resto con quello negativo (in ordine casuale).

Nell'UEQ a 26 *item*, come evidenziano Schrepp et al. (2014), gli *item* sono così distribuiti nelle categorie di analisi:

a. *Attrattività*: Impressione complessiva del prodotto. Alle utenti piace o non piace? È attraente, piacevole o gradevole?
b. *Apprendibilità*: È facile familiarizzare con il prodotto? È facile da imparare? Il prodotto è facile da capire e chiaro?
c. *Efficienza*: Le utenti possono risolvere i loro compiti senza sforzi inutili? L'interazione è efficiente e veloce?
d. *Controllabilità*: L'utente si sente in controllo dell'interazione? È in grado di prevedere il comportamento del sistema? L'utente si sente sicura quando lavora con il prodotto?
e. *Stimolazione*: L'utilizzo del prodotto è stimolante e motivante? È divertente da usare?
f. *Originalità*: Il prodotto è innovativo e creativo? Cattura l'attenzione delle utenti?

Non si presume che le scale siano indipendenti; infatti, l'impressione generale di un'utente è catturata dalla scala dell'attrattività, che dovrebbe essere influenzata dai valori delle altre 5 scale (Schrepp et al., 2017, p. 41).

2.4.4 Attività di retrospezione

Nel presente studio sperimentale la retrospezione complementa l'analisi delle percezioni delle utenti e la valutazione della *performance* delle partecipanti, fornendo un dato qualitativo che permette: di analizzare i bisogni di interpreti in formazione dinanzi agli strumenti tecnologici che offrono loro supporto, di indagare i motivi per cui la percezione delle partecipanti dinanzi a ciascuno strumento ha determinati valori quantitativi nell'UEQ, di identificare i potenziali problemi significativi nell'uso delle due tecnologie proposte, di integrare i dati quantitativo-qualitativi dell'*eye tracker* con la spiegazione dei motivi sottostanti a determinati comportamenti relativi al come lo strumento viene utilizzato. La retrospezione è infatti un metodo introspettivo che intercetta i processi cognitivi dei soggetti attraverso i loro stessi resoconti; essa si basa sul presupposto che alcune informazioni provenienti dalla memoria di lavoro del soggetto durante un particolare compito siano immagazzinate nella memoria a lungo termine e possano essere recuperate a posteriori (Ericsson & Simon, 1993, p. 149). La retrospezione ha maggior validità quando il compito svolto è recente o di breve durata. Come evidenziano Englund Dimitrova & Tiselius (2014), la retrospezione presenta numerosi limiti; uno di questi è che non permette un *recall* completo di tutte le informazioni relative al *task* svolto in quanto ciò viene ricordato non ha necessariamente la stessa forma di quando è stato elaborato dalla memoria di lavoro (considerata come quella memoria coinvolta quando si assiste a qualsiasi tipo di attivazione dei processi mentali), avendo subito processi di astrazione, generalizzazione ed elaborazione (Englund Dimitrova, 2005). In un compito dalla durata di circa 10 minuti (come quello presentato nello studio sperimentale di questo elaborato), l'accuratezza delle informazioni nelle retrospezioni potrebbe essere ridotta anche a causa della perdita mnemonica di alcune informazioni (Ericsson & Simon, 1993). Tendenzialmente, per agevolare il ricordo delle informazioni e dei processi decisionali messi in campo durante l'esecuzione del compito, la retrospezione viene svolta utilizzando un segnale input (o *retrieval cue*) e la scelta del *cue* è fondamentale, in quanto deve garantire la possibilità di ripercorrere l'attività svolta fornendo quanti più dati possibile. Al contempo, come evidenziano Dimitrova & Tiselius (2009, p. 112), è fondamentale ridurre al minimo il rischio che il *cue* diventi una fonte di nuovi processi cognitivi che non facevano parte del *task* originale; pertanto, la retrospezione è agevolata quanto il segnale input è simile (o uguale) a quello originale (Ericsson & Simon, 1987), pena il rischio di installare delle memorie false (Meade & Roediger, 2002).

Englund Dimitrova & Tiselius (2009, p. 112) sottolineano che è rischioso utilizzare l'audio della versione interpretata da parte delle partecipanti, in quanto la resa rappresenta un *cue* che potrebbe portare a nuovi processi cognitivi che potrebbero distorcere i dati: le partecipanti dinanzi al proprio output commentano il loro output piuttosto che il loro processo; pertanto, è probabile che le partecipanti valutino, reagiscano emotivamente, razionalizzino, spieghino, ecc. In questo studio sperimentale la lunghezza dei compiti (di circa 10 minuti ciascuno) rappresenta un problema relativamente alla necessità dell'immediatezza della retrospezione, che è stato parzialmente risolto proponendo la retrospezione come attività immediatamente successiva all'esecuzione dell'attività di interpretazione simultanea. Il segnale input utilizzato è costituito da due componenti: l'audio del discorso originale in lingua spagnola e il movimento dell'occhio sovrapposto al video del discorso originale. Durante la retrospezione sono stati utilizzati due computer: nel computer di sinistra (quello contenente il video input dell'attività sperimentale di IS) è stato aperto il video contenente il discorso originale con il movimento dell'occhio della partecipante (senza audio) e nel computer di destra è stato aperto il video originale con l'audio del discorso originale. Nel protocollo retrospettivo, prima di iniziare l'attività di retrospezione, è stato chiesto a ogni partecipante di ripercorrere quanto è successo durante l'attività di simultanea senza avere un approccio giudicante della resa, fornendo un'introduzione simile a quella redatta di seguito, che non è stata tuttavia letta ma inserita in una conversazione in cui le partecipanti hanno potuto reagire o fare domande:

> *"Ti chiederei adesso di ripercorrere cosa è successo durante l'interpretazione simultanea. Ti chiederei di non giudicare la tua resa e di non focalizzarti sugli errori, mi interessa cosa è successo e cosa hai pensato durante l'attività. Nel computer di sinistra vedrai il discorso originale così come ti è stato presentato durante la simultanea e vedrai un puntino viola che si muove: è il movimento del tuo occhio durante tutta la simultanea. Il video sul computer di sinistra non ha l'audio, quindi aprirò su quest'altro computer (di destra) il discorso con l'audio. Quando ritieni opportuno, puoi fermare il video sul computer di sinistra e io fermerò il video sul computer di destra. Sentiti libera di dirmi quello che ritieni opportuno e in qualsiasi momento, niente è inutile. Durante tutta l'attività verranno registrati i tuoi commenti, in modo da poterli riascoltare con calma."*

Dopo aver dato l'input, non si è interferito con la retrospezione nel corso di tutta la sua attività. Il ricercatore si è seduto in una posizione che fosse fuori dal campo visivo diretto delle partecipanti: alla destra della partecipante e leggermente dietro. Durante la retrospezione, talvolta le partecipanti hanno cercato approvazione voltandosi ponendo domande circa

l'adeguatezza di alcuni commenti o anche solo sguardi, in questi casi è stato sufficiente dire *"tutto è utile"*, annuire col capo o sorridere per non interferire in modo significativo. Il ricercatore ha talvolta verbalizzato oralmente quanto le partecipanti indicavano con la mano, in modo da verbalizzare l'azione ai fini della trascrizione.

Durante la retrospezione, per ogni partecipante è stato preparato un *file* cartaceo in cui il raccoglitore dei dati ha annotato i minuti in cui il video è stato fermato da ogni partecipante, in modo da garantire una buona sovrapposizione della retrospezione con il video originale.

Minutaggio retrospezione	Codice partecipante: XXXX Numero partecipante: XXXX
Condizione: Google Nome prova: Google_A.1	Minutaggio: 00:50 01:24 02:57 03:59 04:34 05:35 07:22 07:51 10:22
Condizione: SmarTerp Nome prova: SmarTerp_B.1	Minutaggio: 01:18 02:01 02:34 02:55 03:54 04:26 05:35 07:17 08:00 08:57 09:50 10:39 10:56

Tabella 8 Esempio del file cartaceo utilizzato per l'attività di retrospezione.

Il protocollo retrospettivo utilizzato pertanto prevede solo un input iniziale seguito da un'attività autonomamente gestita dalla partecipante. Færch & Kasper (1987, p. 17) distinguono tra protocolli retrospettivi *self-initiated* e *other-initiated*: in quest'ultimi la retrospezione viene gestita dal ricercatore tramite delle richieste esplicite, mentre nei protocolli *self-initiated* è la partecipante a decidere quando parlare e che tema affrontare. Sulla base di queste premesse, il protocollo adottato è *self-initiated*, in quanto l'interferenza del raccoglitore dei dati è minima.

2.4.5 *Eye-tracking*

I dati relativi al movimento oculare sono stati raccolti tramite il dispositivo di *eye-tracking Eyelink 1000 plus*, in configurazione monoculare. Non è stato utilizzato alcun supporto per la testa o per il mento, anche se ciò avrebbe migliorato la qualità dei dati (Holmqvist, 2011, p. 83), in quanto avrebbe costretto le partecipanti a non muovere la bocca durante l'attività

sperimentale. L'*eye tracker* utilizzato garantisce un'invasività minima nei confronti dell'attività di simultanea.

La gestione dell'*eye-tracking* durante tutte le batterie di esperimenti è stata totalmente nelle mani della professoressa relatrice di questo elaborato, che ha gestito il *software* dell'*eye tracker* da un *host computer* vicino al *display computer* utilizzato per svolgere le attività di interpretazione simultanea.

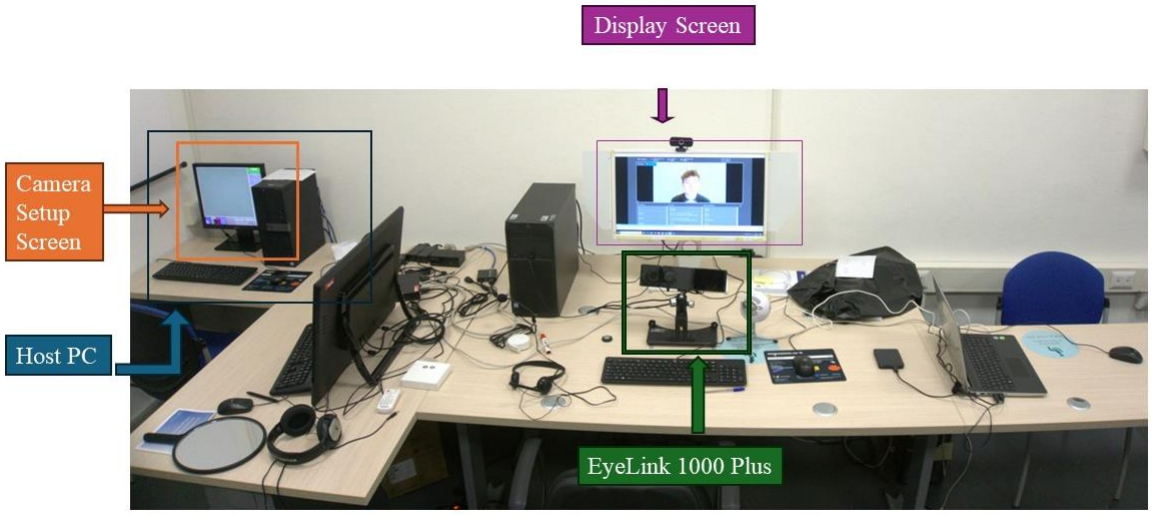


Figura 7 Configurazione spaziale dell'hardware necessario per la raccolta dati tramite eye tracker EyeLink 1000 Plus.

Il *software* ha chiesto a ogni partecipante di inserire un codice all'inizio della procedura. Ogni partecipante ha dunque inserito un codice alfanumerico, secondo un criterio stabilito dal raccoglitore dei dati, che è stato ulteriormente anonimizzato nell'analisi dei dati. Il criterio stabilito è utile per poter risalire ai dati di una partecipante in caso di ritiro dallo studio sperimentale.

Dal *Camera Setup Screen* sull'*host computer* è possibile visualizzare la messa a fuoco dell'occhio e la correttezza della posizione della partecipante. A tutte le partecipanti è stato chiesto di sedersi in una posizione comoda con la zona addominale più vicina possibile al tavolo. L'altezza della sedia è stata regolata per ogni partecipante in modo da far sì che ogni partecipante avesse gli occhi a un'altezza corrispondente (in linea d'aria) ai tre quarti dell'altezza dello schermo.

Prendendo come punto di riferimento le ciglia della partecipante, è stata corretta la messa a fuoco della videocamera dell'*eye-tracker*. Dopo aver corretto la posizione, è stato chiesto alla partecipante di leggere 4 numeri collocati ai quattro angoli dello schermo per verificare la corretta posizione della pupilla all'interno del focus dell'*eye-tracker* in ogni estremità dello

schermo. Successivamente è stata avviata una calibrazione a 9 punti in cui è stato chiesto alla partecipante di fissare lo sguardo sui punti presentati sullo schermo.

Una volta ottenuta una calibrazione buona, è stata validata. La validazione prevede una presentazione dei *target* sullo schermo della partecipante in ordine casuale, simile alla calibrazione. Durante la validazione è possibile osservare sullo schermo dell'*Host Computer* il margine di errore di ogni fissazione. Dopo aver terminato la validazione, la professoressa responsabile dell'*eye-tracker* (in quanto anche docente delle partecipanti all'esperimento) si è allontanata dalla postazione dell'*Host Computer* per non influenzare emotivamente le partecipanti durante la propria resa.

2.5 Partecipanti

2.5.1 Reclutamento delle partecipanti

Lo studio sperimentale ha coinvolto l'intera classe di interpretazione dallo spagnolo all'italiano del secondo anno del corso magistrale in Interpretazione presso il Dipartimento di Interpretazione e Traduzione dell'Università di Bologna. Le partecipanti sono state contattate attraverso canali di comunicazione informali e – una volta espresso il loro desiderio di partecipare allo studio – è stato creato un canale di comunicazione condiviso, in cui sono state condivise tutte le informazioni relative alle attività da svolgere in fase sperimentale: due attività di interpretazione simultanea dallo spagnolo all'italiano di due discorsi della durata di circa 11 minuti con il supporto della trascrizione automatica di Google Meet (in un discorso) e del CAI *tool* di SmarTerp (nell'altro) presso un laboratorio del Dipartimento di Interpretazione e Traduzione e una formazione asincrona suddivisa in due sessioni. Tutte le partecipanti hanno completato il corso di interpretazione dallo spagnolo in italiano del primo anno del corso di laurea magistrale (equivalente a cinque crediti formativi) e le partecipanti con la lingua spagnola come lingua B[19] nella propria combinazione linguistica hanno completato anche il

[19] Nella profilazione delle partecipanti, verrà utilizzato il sistema universale per la classificazione delle lingue attive e passive che è stato stabilito da AIIC, che prevede una distinzione tra: lingua A (equivalente alla lingua madre o a una competenza linguistica equivalente), lingua B (una lingua di cui l'interprete ha un'ottima padronanza e verso la quale l'interprete è in grado di interpretare sia in modalità consecutiva che simultanea) e

corso di interpretazione dall'italiano in spagnolo del primo anno (equivalente a 6 crediti formativi). Inoltre, considerato che l'attività sperimentale ha avuto luogo nel mese di febbraio, tutte le partecipanti hanno completato il primo semestre di formazione del secondo anno del corso di laurea magistrale. Pertanto, tutte le partecipanti hanno complessivamente completato tre trimestri di formazione in aula in interpretazione simultanea e consecutiva e hanno la lingua spagnola come loro lingua A (madre), B (attiva) o C (passiva). Come l'autore, tutte le partecipanti hanno seguito il corso di Metodi e Tecnologie per l'Interpretazione (sia del primo che del secondo anno del corso di laurea). La partecipazione all'esperimento è stata gratuita e non retribuita, ma è stata invece compensata dal fatto che le studentesse hanno potuto approfondire l'uso di uno strumento CAI e della trascrizione automatica grazie alla formazione e al successivo *test* in laboratorio. Ad ogni partecipante è stato chiesto di inserire un codice alfanumerico all'inizio dell'esperimento, in modo da poter risalire alle prove in caso di ritiro volontario dallo studio. Il codice alfanumerico è stato ulteriormente anonimizzato nel presente studio. Nonostante il campione sia già circoscritto data la provenienza da un comune percorso formativo, è stato proposto (in fase sperimentale) a tutte le undici partecipanti allo studio un questionario che prevede la stessa costruzione logica sottostante al LexTALE[20] (*Lexical Test for Advanced Learners of English*). La compilazione del questionario (di seguito, *LexTest*) ha il vantaggio di essere rapida (circa 5 minuti) ed ecologicamente valida. Il questionario utilizzato contiene complessivamente 90 *item*, di cui 63 *word* (parole esistenti in lingua spagnola) e 27 *nonword* (parole con caratteristiche morfologiche simili a quelle della lingua spagnola ma non esistenti nello spagnolo). La ragione di questa proporzione "sbilanciata" è che le *word* hanno una frequenza talmente bassa che è improbabile che qualche partecipante le conosca tutte (trasformando un numero considerevole di *word* in *nonword* soggettive) (Lemhöfer & Broersma, 2012, p. 329).

Il questionario utilizzato non è stato redatto dall'autore di questo elaborato, ma è stato scaricato da internet, ma considerato il numero di *word* e *nonword* diverso da quello utilizzato nel LexTALE disponibile su Internet per la valutazione della lingua inglese, la formula per il calcolo percentuale è stata adattata come segue:

*((number of words correct/63*100) + (number of nonwords correct/27*100)) / 2*

lingua C (una lingua che l'interprete comprende perfettamente ma verso la quale non lavora). https://aiic.org/site/world/about/profession/abc (Ultimo accesso: 30 aprile 2024)

[20] https://www.lextale.com/validity.html (Ultimo accesso: 30 aprile 2024)

Prima della somministrazione del questionario, ad ogni partecipante sono state consegnate delle istruzioni scritte in cui è stato spiegato che nel *test* lessicale ci sarebbero state alcune parole esistenti nella lingua spagnola e altre inesistenti. Successivamente, ogni partecipante ha completato il questionario mettendo un segno di spunta (✓) accanto alle parole che secondo la partecipante erano parole spagnole. L'omogeneità del campione che è stato coinvolto nel presente studio sperimentale è riassumibile nei seguenti valori ottenuti mediante il *test* lessicale:

	Media	Mediana	SD	Minimo	Massimo
Risultati LexTest	82.4	82.5	7.62	70.6	94.2

Tabella 9 Descrizione del campione analizzato sulla base dei risultati del LexTest.

Le partecipanti all'esperimento sono prevalentemente donne (10 donne e 1 uomo) e 10 partecipanti su 11 hanno la lingua italiana come lingua madre, mentre 1 partecipante è di lingua madre spagnola. Le partecipanti hanno un'età media di 24 anni. A ogni partecipante, insieme al LexTest è stato proposto un questionario in cui si chiedevano informazioni relative alla combinazione linguistica e agli anni di studio della lingua spagnola; inoltre, a ogni interprete è stato chiesto di fornire un'autovalutazione circa la propria padronanza della lingua spagnola.

Alla luce del materiale raccolto per una profilazione delle partecipanti più ampia, è possibile osservare che le partecipanti hanno in media studiato la lingua spagnola per 11 anni (media: 11,3): da un minimo di 7 anni a un massimo di 20 anni (SD 3,82); inoltre, l'autovalutazione media delle partecipanti circa la padronanza della lingua spagnola è uguale a 8,23 su 10, con valori che oscillano tra un minimo di 7 e un massimo di 10 (SD 0,984).

Prima di procedere alla descrizione della formazione sottoposta alle partecipanti, è necessario chiarire che, nell'analisi, le prestazioni della partecipante con spagnolo A sono state estratte dal campione, mentre i dati relativi all'UEQ e i dati di *eye-tracking* hanno incluso i dati relativi a questa partecipante. Pertanto, nella valutazione dell'informatività, dell'intelligibilità e del *success rate* ottenuto nei singoli *task*, il campione coinvolto sarà di 10 (e non 11) partecipanti. Questo ha permesso, inoltre, l'analisi su dati ancora più omogenei, in quanto i dati relativi all'accuratezza della resa, in questo modo, sono formati da 10 discorsi A.1 presentati nella

condizione contenente la trascrizione automatica di Google Meet e 10 discorsi B.1 presentati con il CAI *tool* di SmarTerp.

2.5.2 Formazione delle partecipanti

Terminato il reclutamento, tutte le partecipanti hanno completato una formazione erogata in modalità asincrona tramite cartella condivisa su *OneDrive* e suddivisa in due sessioni. L'obiettivo della formazione è stato quello di far loro acquisire una maggior familiarità con la trascrizione automatica di Google Meet e il CAI *tool* di SmarTerp. La formazione ha avuto come filo conduttore lo stesso dell'attività sperimentale (il *Ministero spagnolo per la Transizione Ecologica e la Sfida Demografica*), in quanto ha avuto anche l'obiettivo di simulare una preparazione realistica a un incarico, durante la quale l'interprete vuole raggiungere una maggior familiarizzazione con il lessico e il genere testuale. Ogni sessione di formazione conteneva: (a) 1 video in spagnolo con la trascrizione automatica di Google Meet da utilizzare per un'attività di IS; (b) 1 video con il CAI *tool* di SmarTerp da interpretare in modalità simultanea; (c) un glossario contenente la terminologia specialistica utilizzata nei discorsi proposti; (d) un *file* .doc contenente la trascrizione del discorso (a); (e) un *file* .doc contenente la trascrizione del discorso (b); un *file* .doc contenente un *briefing* sul contesto informativo-comunicativo di entrambi i discorsi.

Tutti i discorsi utilizzati sono stati tratti dal sito ufficiale del *Congreso de los Diputados*[21] e sono tutti interventi parlamentari della Ministra spagnola per la Transizione Ecologica e la Sfida demografica (mandato parlamentare 2019 – 2023). Data la velocità d'eloquio di tutti i discorsi (che non avrebbe permesso un uso adeguato, razionale e produttivo delle tecnologie di supporto proposte), i video utilizzati nell'esperimento sono stati letti in spagnolo dall'autore del presente elaborato (seguendo la configurazione tecnica utilizzata per i video preparati per l'esperimento). In fase di preparazione, è stato chiarito alle partecipanti che l'oratrice sarebbe stata una persona spagnola e non l'autore (italiano).

La prima sessione di formazione prevede: un discorso (discorso 1) in cui la ministra Teresa Ribera presenta la Legge di Bilancio per il 2023, focalizzandosi specificamente sulla Sezione 23, che riguarda il Ministero della Transizione Ecologica e la Sfida Demografica; un discorso (discorso 2) in cui la ministra Teresa Ribera presenta il disegno di legge con cui si adottano

[21] https://www.congreso.es/es/busqueda-de-intervenciones (Ultimo accesso: 14 aprile 2024)

misure urgenti nel settore dell'energia per promuovere la mobilità elettrica, l'autoconsumo e la diffusione delle energie rinnovabili; un video di 14 minuti e 12 secondi contenente il discorso 1 nella condizione con CAI *tool* di SmarTerp; un video di 12 minuti e 52 secondi contenente il discorso 2 nella condizione con trascrizione automatica di Google Meet; 1 *file* contenente i due *briefing*; 1 glossario di 88 termini spagnoli con i rispettivi traducenti in italiano in cui è stata condensata la terminologia specialistica utile per entrambi i video.

Analogamente alla struttura della prima sessione di formazione, la seconda sessione di formazione prevede: un discorso (discorso 1) in cui la ministra Teresa Ribera presenta la Legge di Bilancio per il 2021, focalizzandosi specificamente sulla Sezione 23, che riguarda il Ministero della Transizione Ecologica e la Sfida Demografica; un discorso (discorso 2) in cui la ministra Teresa Ribera presenta il Regio Decreto-Legge 23/2020, del 23 giugno, che approva misure nel campo dell'energia e in altri settori per la riattivazione economica; un video di 14 minuti e 19 secondi contenente il discorso 1 nella condizione con la trascrizione automatica di Google Meet; un video di 17 minuti e 2 secondi contenente il discorso 2 nella condizione con CAI *tool* di SmarTerp; 1 *file* contenente i due *briefing*; 1 glossario di 98 termini spagnoli con i rispettivi traducenti in italiano in cui è stata condensata la terminologia specialistica utile per entrambi i video.

Per quanto riguarda la configurazione delle due sessioni SmarTerp con CAI (indispensabili per registrare i due video della formazione nella condizione con CAI *tool* di SmarTerp), oltre ai due glossari già menzionati sono stati necessari 2 corpora per ogni video (uno in spagnolo e uno in italiano).

Il criterio per la compilazione dei corpora è stato lo stesso adottato per la preparazione dei video per l'attività sperimentale: 7 occorrenze per ogni termine utilizzato nel glossario (non equivalente a quello fornito alle partecipanti, in quanto non contenente i termini relativi al discorso nella condizione Google Meet). I corpora utilizzati per la preparazione dei video nella condizione SmarTerp per la prima e la seconda sessione di formazione hanno le grandezze seguenti:

	Tokens	*Words*	*Sentences*
Corpus in italiano per la prima sessione	39.066	34.200	1.047
Corpus in spagnolo per la prima sessione	39.456	34.564	1.061
Corpus in italiano per la seconda sessione	54.358	46.442	1.437
Corpus in spagnolo per la seconda sessione	64.094	56.210	1.723

Tabella 10 Grandezze dei corpora utilizzati per la configurazione delle sessioni SmarTerp con CAI per la preparazione dei video.

Il glossario utilizzato per la creazione dei corpora per la prima sessione di formazione contiene 59 termini in spagnolo (con i rispettivi traducenti); mentre il glossario utilizzato per la creazione dei corpora per la seconda sessione di formazione contiene 77 termini in spagnolo.

L'interfaccia utente (o IU) di tutti i video utilizzati nella formazione rispecchia l'interfaccia proposta durante l'attività sperimentale.

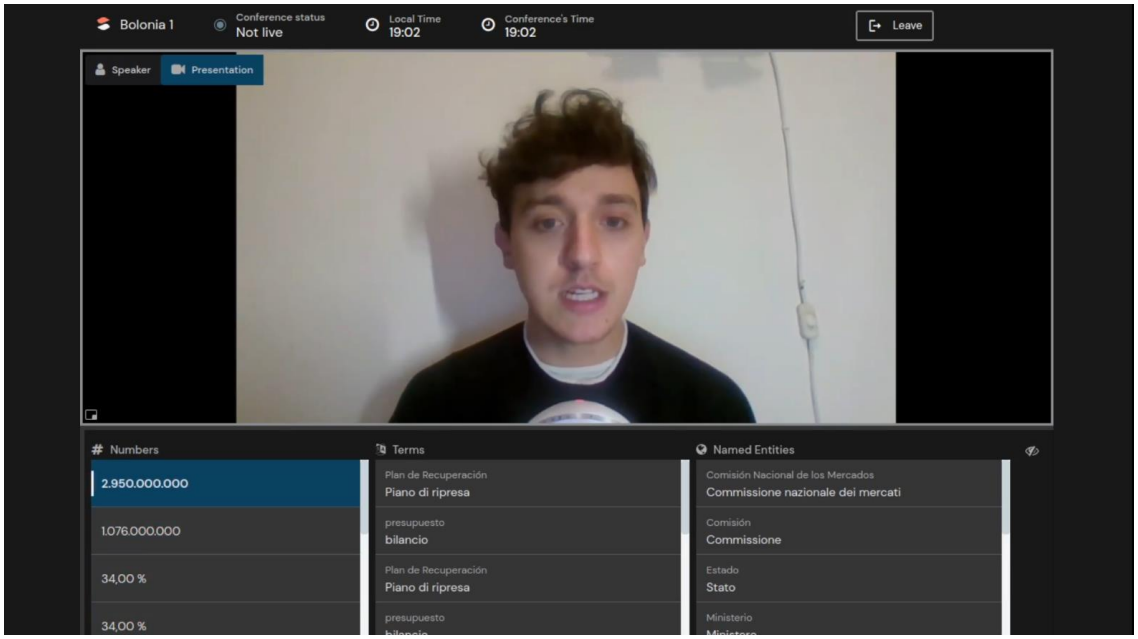


Figura 8 Esempio dell'IU del video utilizzato nella condizione SmarTerp per la prima giornata di formazione.

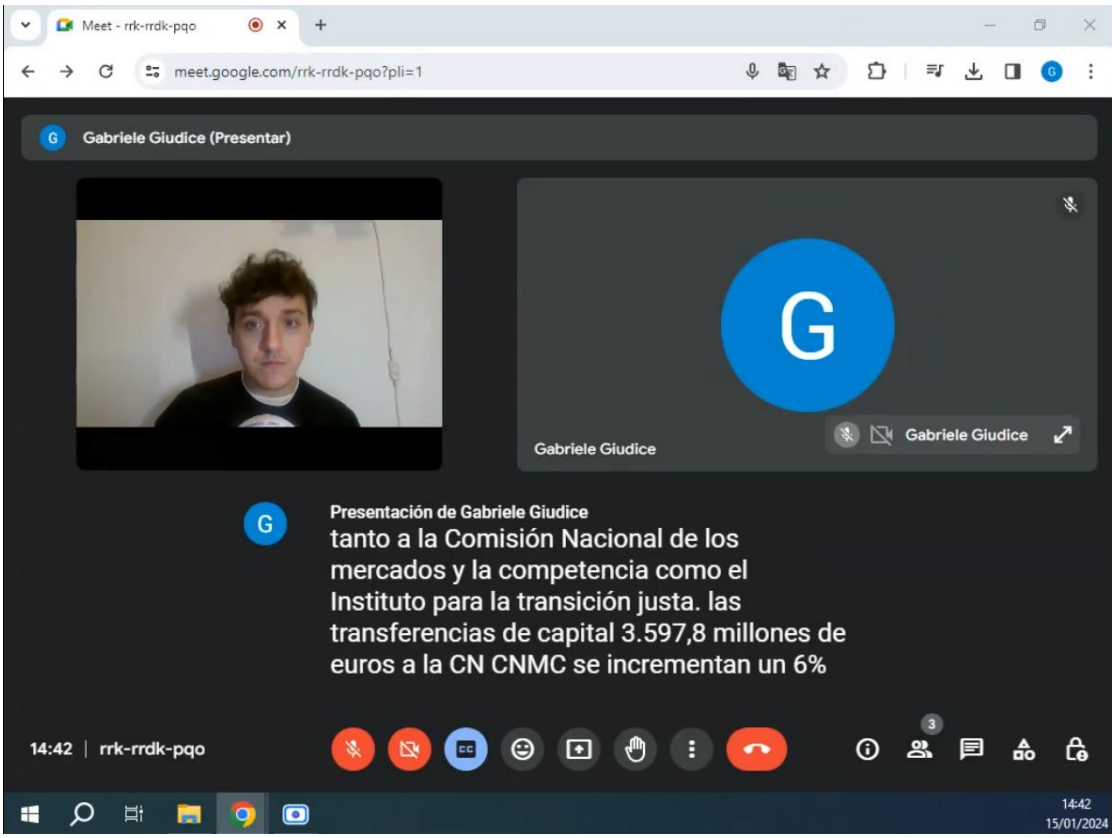


Figura 9 Esempio dell'IU del video utilizzato nella condizione Google Meet per la seconda giornata di formazione.

Considerata la necessità di raccogliere dati da una sola partecipante alla volta, è stato compilato un calendario condiviso con le partecipanti e di comune accordo con la relatrice di questo elaborato (in quanto responsabile della gestione dell'*eye-tracker*). Le 11 attività sperimentali individuali dello studio hanno avuto luogo presso il laboratorio 4 (contenente il dispositivo *eye tracker*) presso il Dipartimento di Interpretazione e Traduzione in un periodo complessivo di 10 giorni, nel febbraio 2024. Ad ogni partecipante è stata inviata la prima sessione di formazione dieci giorni prima della propria prova sperimentale e la seconda sessione quattro giorni prima del *test*. In occasione dell'invio della seconda sessione di formazione, ad ogni partecipante è stato chiarito che i video utilizzati nell'esperimento avrebbero avuto un tema in linea con la formazione. Il giorno precedente all'esperimento, sono stati inviati ad ogni partecipante il modulo informativo sul trattamento dei dati e il modulo relativo al consenso informato per poter dare ad ogni partecipante la possibilità di leggerli con calma prima dell'esperimento; inoltre, congiuntamente ai moduli, è stato ricordato alla partecipante l'orario e il luogo dell'esperimento. Dato l'uso dell'*eye-tracker*, è stato chiesto di evitare un trucco accentuato e, in caso di necessità, di portare sia gli occhiali che le lenti a contatto (nel caso in cui la partecipante avesse a propria disposizione entrambe le possibilità).

2.6 Svolgimento dell'esperimento

Prima di procede alla descrizione del protocollo adottato per la raccolta dati, è necessario chiarire che – come descritto nella sezione §2.4.2 – ogni discorso è stato preparato per essere proposto in entrambe le condizioni sperimentali: in questo modo è stata possibile la randomizzazione sia della corrispondenza dell'ordine delle condizioni che del discorso per evitare l'influenza del fattore stanchezza sui risultati. La randomizzazione è stata eseguita seguendo uno schema *Latin Square* bilanciato.

Partecipante	P1	P2	Nome prova 1	Nome prova 2
P1	BG	AS	Google_B.1	SmarTerp_A.1
P2	BS	AG	SmarTerp_B.1	Google_A.1
P3	AG	BS	Google_A.1	SmarTerp_B.1
P4	AS	BG	SmarTerp_A.1	Google_B.1
P5	BG	AS	Google_B.1	SmarTerp_A.1
P6	BS	AG	SmarTerp_B.1	Google_A.1
P7	AG	BS	Google_A.1	SmarTerp_B.1
P8	AS	BG	SmarTerp_A.1	Google_B.1
P9	BG	AS	Google_B.1	SmarTerp_A.1
P10	BS	AG	SmarTerp_B.1	Google_A.1
P11	AG	BS	Google_A.1	SmarTerp_B.1

Tabella 11 Randomizzazione delle prove per ogni partecipante. A (discorso A.1), B (discorso B.1), S (con SmarTerp) e G (con Google Meet)

Il protocollo sperimentale utilizzato è stato il seguente.

Dopo l'arrivo della partecipante, è stato chiesto di mettere il cellulare in modalità silenziosa e di andare in bagno prima dell'inizio dell'attività, se ne avesse avuto bisogno. Successivamente, è stata proposta la compilazione dei due moduli inviati il giorno precedente (il modulo relativo al trattamento dei dati personali e quello relativo al consenso informato). Dopo essersi assicurati che la porta e le finestre fossero chiuse, la partecipante è stata fatta sedere sulla sedia preposta allo svolgimento dell'attività di IS (quindi di fronte allo schermo utilizzato per la proiezione delle prove di simultanea). Successivamente, il ricercatore ha fornito un breve *briefing* sul contesto informativo-comunicativo dei video, volto a rassicurare la partecipante sulla condizione sperimentale del primo video e sul tema generico in linea con la formazione. Una

volta che è stata perfezionata la posizione della partecipante (altezza e vicinanza rispetto al tavolo) e la messa a fuoco della videocamera dell'*eye-tracker*, è stata eseguita e validata la calibrazione del dispositivo di *eye-tracking*. Durante la validazione dell'*eye-tracker* è stata avviata una registrazione su *Audacity* da un dispositivo portatile presente alla destra dell'interprete, in modo da avere una registrazione di *back-up*. È stata avviata la registrazione dei dati dell'*eye-tracking*. Successivamente si è presentato alla partecipante un breve frammento di uno dei video utilizzati nelle sessioni di formazione per una rapida prova audio. Terminata la regolazione dell'audio input, è stato avviato il video del primo discorso: il ricercatore ha avviato la registrazione in concomitanza con un segnale vocale ("*via*") che permettesse la sovrapposizione dei dati di *eye-tracking* con quelli della registrazione dell'output vocale dell'interprete. Durante tutte le prove di simultanea, la professoressa si è allontanata dalla postazione dell'*Host Computer*, pur mantenendolo visivamente monitorato, per non interferire emotivamente sulla resa delle partecipanti. Terminata l'interpretazione simultanea del primo discorso, è stata interrotta la registrazione di *Audacity* e opportunamente salvata in una cartella precedentemente creata (con un nome *file* che contenesse il codice partecipante e la condizione sperimentale della prova, simile a *p0_smarterp_sim*). Subito dopo il salvataggio della prova, il raccoglitore dei dati ha preparato rapidamente la configurazione necessaria per l'attività di retrospezione: (a) video con l'audio originale sul PC portatile di destra; (b) video con il movimento dell'occhio sul computer centrale; (c) sedia alla destra della partecipante per il raccoglitore dei dati (leggermente dietro alla partecipante in una posizione che non rientrasse nel campo visivo diretto della partecipante); (d) *file* cartaceo su cui appuntarsi il minutaggio della retrospezione. Dopo aver chiesto alla partecipante di ripercorrere cosa è successo e aver avviato la registrazione dell'audio dal *software Audacity* dal PC di destra è iniziata l'attività di retrospezione. Durante la retrospezione il ricercatore ha appuntato nella scheda cartacea il minutaggio dei momenti in cui il video è stato fermato dalla partecipante per commenti o riflessioni. Al termine della retrospezione del primo video, è stata interrotta la registrazione *Audacity* ed è stata salvata con un nome che contenesse codice partecipante + condizione (simile a *p0_smarterp_retro*). Terminata la retrospezione del primo video, è stato proposto alla partecipante l'*User Experience Questionnaire* cartaceo relativo alla tecnologia utilizzata nella prima prova, chiedendo una compilazione istintiva. Una volta riconsegnato, sul questionario è stato scritto il codice della partecipante e lo strumento che è stato valutato. Successivamente, è stato fornito il *briefing* per la seconda prova di simultanea. Dopo aver nuovamente calibrato e validato l'*eye-tracker*, è stata avviata la registrazione da *Audacity* sul PC di destra e è stato avviato il video del secondo discorso (sempre in concomitanza con il segnale acustico "*via*").

Al termine della seconda prova di simultanea, è stata salvata la registrazione (seguendo il criterio di salvataggio adottato per il primo discorso). Successivamente, è stata svolta la retrospezione del secondo video (seguendo configurazione e procedura temporale illustrate per il primo video). Dopo aver salvato la registrazione della retrospezione, è stato sottoposto l'UEQ relativo alla tecnologia utilizzata nella seconda condizione sperimentale. Al termine del questionario, è stato proposto alla partecipante un *file* cartaceo da compilare contenente il *LexTest* a 90 *item* per la profilazione del campione e alcune domande relative agli anni di studio della lingua spagnola, insieme all'autovalutazione circa la padronanza della lingua spagnola e la combinazione linguistica. Terminato l'ultimo questionario, la partecipante è stata congedata.

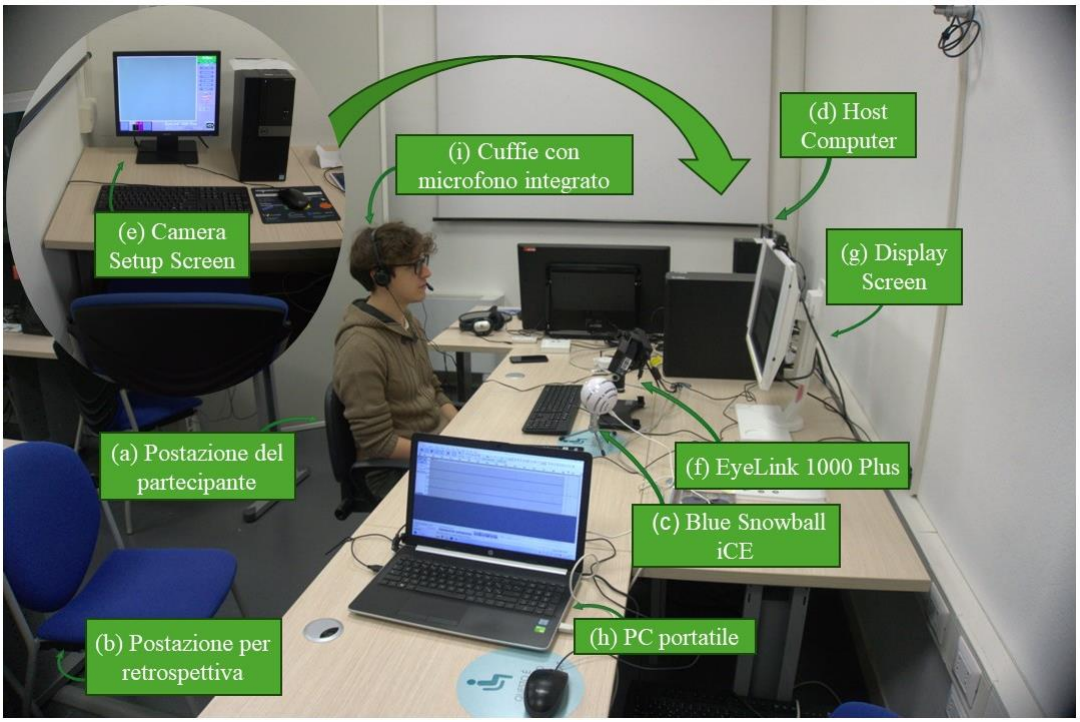


Figura 10 Configurazione spaziale degli strumenti utilizzati durante l'attività sperimentale e posizione della postazione del partecipante.

Durante l'attività sperimentale è stato utilizzato un *host computer* (d) dal quale è stato gestito il *software* dell'*eye-tracker* preposto alla raccolta dati: sull'*host computer* è stato aperto il *Camera Setup Screen* del *software EyeLink Data Viewer* per poter procedere alla calibrazione (e alla validazione della stessa) prima di ogni interpretazione simultanea. Ogni partecipante ha utilizzato le stesse cuffie con microfono integrato (collegate all'*host computer*) per poter ascoltare l'audio in input. Il video input di ogni prova di simultanea è stato mostrato sul *display screen* (g) di fronte alla partecipante. Oltre al microfono delle cuffie, è stato utilizzato un microfono USB esterno, *Blue Snowball Ice* (c), collegato al PC portatile alla destra della

partecipante. Il *PC portatile* (h) alla destra della partecipante è stato utilizzato per registrare gli audio delle prove di simultanea e delle due retrospezioni tramite il programma *Audacity*. Durante le retrospezioni, il raccoglitore dei dati si è seduto alla destra della partecipante (e leggermente dietro) sulla *sedia* (b). Durante le retrospezioni, il PC portatile è stato avvicinato alla partecipante ed è stato utilizzato per proiettare l'audio input del discorso oggetto della retrospezione, mentre il *display screen* è stato utilizzato per proiettare il video originale con la sovrapposizione del movimento oculare durante la prova di IS (senza l'audio). Durante tutte le prove di IS docente (relatrice di questo elaborato) e ricercatore si sono allontanati rispettivamente dall'*host computer* e dalla postazione utilizzata per l'attività di retrospezione. Il dispositivo *di eye-tracking EyeLink 1000 Plus* (f) è stato posizionato davanti alla partecipante. Per soddisfare la necessità di una corretta illuminazione per il dispositivo di *eye-tracking* e il bisogno di una stessa condizione ambientale per tutte le partecipanti, è stata esclusivamente utilizzata un'illuminazione artificiale; pertanto, le finestre sono state chiuse e oscurate.

2.7 Preparazione dei dati

Nei prossimi paragrafi verranno illustrati i metodi e i criteri utilizzati per la preparazione e la sistematizzazione dei dati, in vista dell'analisi (proposta nei capitoli 3 e 4).

2.7.1 Trascrizione delle rese

Le rese delle partecipanti sono anzitutto state ritagliate utilizzando il *software Wondershare Filmora*[22] in modo che contenessero solo il materiale rilevante per l'analisi. Data la struttura dell'esperimento e la quantità delle attività proposte, per evitare delle interferenze accidentali con il *software* di *eye-tracking* che causassero malfunzionamenti, le rese sono state registrate con il microfono esterno collegato al PC portatile attraverso *Audacity*[23], mentre le registrazioni audio del *software* di *eye-tracking* sono state utilizzate come registrazioni di *back-up*. Dopo aver eliminato le parti degli audio non rilevanti ai fini dell'analisi, il *file* .mp3 ottenuto è stato importato sul *software* basato su intelligenza artificiale *Whisper*[24] che ha permesso una prima

[22] https://filmora.wondershare.es/ (Ultimo accesso: 13 aprile 2024)
[23] https://www.audacityteam.org/download/ (Ultimo accesso: 13 aprile 2024)
[24] https://openai.com/research/whisper (Ultimo accesso: 13 aprile 2024)

trascrizione automatica degli output orali delle interpreti. La trascrizione generata automaticamente è stata corretta manualmente. Per la trascrizione delle singole rese sono state utilizzate le convenzioni di trascrizione HIAT[25] (Ehlich & Rehbein, 1976), in quanto permettono la trasposizione nel testo scritto di pause, riformulazioni e altri fenomeni utili alla valutazione dell'intelligibilità degli stessi. Le convenzioni sono state adattate, con piccole integrazioni, come segue:

Elemento del discorso	*Convenzioni di trascrizione*	*Esempio*
Parole	Trascrizione letterale per le deviazioni	*il triunfo del pluralismo*
Interiezioni e pause piene	Trascrizione abituale	*eh; ehm*
Numeri e date	Trascrizione letterale	*duecento*
Abbreviazioni	Trascrizione estesa	*C O due*
Elementi inintelligibili	Scrivere ((inintelligibile)) oppure inserire parole incerte tra parentesi	*((inintelligibile))* oppure *sembra una scelta (azzardata)*
Fenomeni non fonologici	Descrizione del fenomeno non fonologico tra parentesi doppie	*((ride))*
Pause	Usare da uno a tre puntini per pause fino a 1 secondo	•••
	Usare un valore numerico per pause superiori a 1 secondo	*((2,4s)*
Termine di un enunciato	Usare un punto interrogativo per le domande	*È arrivata Paola?*
	Usare un punto esclamativo per le esclamazioni	*Che bella giornata!*
	Usare un simbolo di ellissi per enunciati non terminati (frasi aperte)	*Sembrava una bella giornata ma…*
Riparazioni	Segnare una riparazione con uno /	*Che non permetta vivir/ che ci permetta di vivere*
Allungamenti vocalici	Utilizzare due punti dopo la vocale oggetto dell'allungamento vocalico	*sono sta:ti spesi molti soldi*

[25] https://www.exmaralda.org/pdf/HIAT_EN.pdf (Ultimo accesso: 13 aprile 2024)

Velocità d'eloquio	Utilizzare >*questa struttura*< o <*questa struttura*> per segnalare rispettivamente un'accelerazione o un rallentamento nella velocità d'eloquio dell'interprete	*Nel >mese di agosto< hanno sviluppato <un modo per produrre energia>*
Spelling	Trascrizione mediante sequenze di lettere maiuscole e separate da uno spazio	*il P N R R è importante*

Tabella 12 Convenzioni utilizzate per la trascrizione delle rese delle partecipanti, adattate da Ehlich & Rehbein (1976).

2.7.2 Valutazione della *performance*

L'accuratezza terminologica nella resa dei numeri e nella risoluzione delle singole dissimmetrie morfosintattiche potrebbe non essere una misura sufficiente per la valutazione dell'accuratezza e della resa complessiva dell'interprete, in quanto un numero o un termine correttamente interpretato ma contestualizzato nel modo sbagliato comporta comunque un'interpretazione sbagliata. Pertanto, è stato utilizzato un approccio di valutazione misto quantitativo-qualitativo che da un lato mira a verificare l'accuratezza quantitativa mediante la misurazione del tasso di successo (o *task success rate*), come in Frittella (2023), raggiunto nei singoli stimoli predefiniti e dall'altro lato mira a ottenere una valutazione dell'intelligibilità e dell'informatività delle rese mediante una valutazione globale di ciascuna resa, suddivisa nei singoli *cluster* (sia quelli contenenti gli stimoli che quelli contenenti le *control sentence*). Seguendo la metodologia proposta da Mellinger & Hanson (2017, p. 87 - 109), per valutare la significatività statistica delle differenze relative alle medie ottenute dallo stesso soggetto nelle due condizioni sperimentali per quel che riguarda la media totale per ogni partecipante relativa a *task success rate*, informatività e intelligibilità è stato condotto il *Wilcoxon signed-rank test*, un test statistico di tipo non parametrico. È stato utilizzato un test di tipo non parametrico in quanto il campione coinvolto è esiguo e quindi i dati non hanno una distribuzione normale. Inoltre, dato che abbiamo eseguito test multipli (*task success rate*, informatività e intelligibilità) sullo stesso test di dati, abbiamo applicato la correzione di Bonferroni (Dunn, 1961) per ridurre il rischio di errori di tipo I, ossia per evitare falsi positivi con test ripetuti sullo stesso campione. Abbiamo

stabilito il nostro livello di significatività complessiva a 0,05; applicando la correzione di Bonferroni, il livello di significatività di ogni test è stato dunque aggiustato a 0,0167.

2.7.2.1 *Task Success Rate*

Per calcolare il *success rate* di ogni *cluster* contenente gli stimoli, la resa di ogni partecipante è stata trasferita su un *file* .xlsx in cui ogni *task* è stato inserito in una riga del *file*. Ogni riga è stata suddivisa in 6 colonne: una contenente l'abbreviazione del *task* presentato nel *cluster*; una contenente la definizione degli stimoli; una contenente il discorso originale così come è stato letto (durante la lettura ci sono state leggere modifiche dell'oratrice); una contenente gli stimoli isolati che sono seguiti da due parentesi *(x)* vuote in cui verranno inseriti i codici relativi alla valutazione; una contenente la resa dell'interprete e una contenente la media ottenuta tra tutti gli stimoli presenti in un determinato *task*. Prima di illustrare un esempio di valutazione, è necessario chiarire quali convenzioni di codificazione sono state utilizzate per la valutazione dei singoli stimoli:

- Resa accurata: viene inserita una *k* tra le due parentesi.
- Errori (a livello sintattico/semantico/ morfologico): vengono codificati con la lettera *e* e fanno riferimento a errori semantici gravi in cui il significato del discorso originale viene cambiato in modo significativo nella resa.
- Omissioni: codificate con la lettera *o* e identificano l'omissione totale di uno stimolo.
- Omissioni parziali: sono codificate dalla lettera *x* seguita dalla percentuale di contenuto che non è stato rimosso (ad es. x40%); il valore numerico ottenuto è una valutazione approssimata del valutatore.
- Strategie quali la sintesi e la generalizzazione sono state codificate con la lettera *n*.
- Errori meno rilevanti ai fini della presente analisi (quali errori di genere e pronuncia) sono stati codificati con la lettera *x*.

Seguendo il modello proposto in Frittella (2023, p. 96), qui leggermente modificato, a ogni stimolo è stato attribuito un valore percentuale sulla base dei criteri seguenti:

- Resa corretta (100%, codificata con *k*): la resa è corretta e completa.
- Strategia (100%): l'interprete utilizza una strategia che non cambia il messaggio e non causa una perdita di informazioni, ad esempio: *nel 2024 → quest'anno*.
- Errori meno significativi (95% e codificati con la lettera *x*): la resa è accurata e completa, nonostante la perdita di qualche dettaglio o nonostante piccoli errori di pronuncia.

- Resa parziale (proporzionale al contenuto dello stimolo, codificata con x + la percentuale del contenuto interpretato). Ad esempio, il *Ministero spagnolo per la Transizione Ecologica e la Sfida Demografica* → il *Ministero spagnolo per la Transizione Ecologica* (70%).
- Generalizzazione o sintesi (30%, codificata con *n*): l'interprete omette alcuni elementi e sintetizza l'informazione.
- Omissione (0%, codificata con *o*): l'intero stimolo viene omesso senza l'adozione di strategie causando una perdita informativa nel messaggio.
- Errore semantico (0%, codificato con *e*): la resa non è plausibile e coerente.

Successivamente, per ogni *task* è stato calcolato il valore medio, dato dalla media dei *success rate* raggiunti in ogni singolo stimolo. Di seguito, viene proposta una breve sintesi dei criteri utilizzati per la valutazione del *success rate*[26]:

Errore / Fenomeno	Codice	*Success Rate*
Omissione	o	0%
Errore semantico o calco	e	0%
Resa parziale	x	Proporzionale al contenuto del *task*
Errore di pronuncia	x	95%
Errore di Genere	x	95%
Errata attribuzione	e	0%
Frase aperta	e	0%
Errore (relativo alla sintassi)	e	0%
Incoerenza	e	0%
Distorsione	e	0%
Errore di plausibilità	e	0%
Nonsense	e	0%
Omissione di *item* ridondante	k	100%
Sostituzione lessicale	k	100%
Generalizzazione	n	30%
Sintesi	n	30%
Resa senza errori	k	100%

Tabella 13 Task success rate - criteri di valutazione (adattati da Frittella, 2023).

[26] Nella tabella relativa alla tipologia di errori e alla loro rispettiva valutazione quantitativa, rispetto alla versione proposta da Frittella (2023), sono stati integrati due elementi: il calco e la resa senza errori.

Di seguito, viene mostrata una riga del *file* .xlsx contenente l'analisi del *success rate* di un *task*:

Abbreviazione Task	Definizione degli stimoli	Discorso A.1 così come è stato letto	Stimoli isolati, accanto a ogni stimolo/subtask inserire la valutazione tra ()	Resa	Media dei success rate dei subtask = Task Success Rate
UN	Unità informativa con numerali	La Secretaría de Estado liderada por Morán Fernández cuenta con un presupuesto de 579 millones de euros, cifra que supera con creces la de 2021, la cual representa un 2,85% más. En este contexto, España aspira a lograr los 22 megavatios de capacidad antes de 2032.	Morán Fernández(n) // megavatios(k) // 2032(k) // 2021(k) // 22(k) // 579 millones(e) // 2,85%(k)	e:h • la segreteria dello Stato • e:h con e:h il ministro eh **Fernández** eh ha un bilancio di e:h **cinquecento novantasette milioni** di euro che supera quella del **duemila e ventuno** •• e rappresenta un **due virgola ottantacinque per cento** in più • in questo eh progetto >dobbiamo sottolineare< i **ventidue eh megawatt** entro il **duemila e trentadue**	75,70%

Tabella 14 Esempio della riga del file .xslx utilizzata per l'analisi quantitativa del success rate di un task.

2.7.2.2 Valutazione globale

Per una valutazione globale dell'accuratezza e dell'informatività dell'output testuale delle interpreti, sono state utilizzate le griglie di valutazione che Tiselius (2009) ha realizzato riadattando le scale di Carroll (1966). Nell'adattamento delle scale, Tiselius ha: (a) eliminato i riferimenti e gli elementi relativi ai testi scritti e alla traduzione scritta; (2) aggiunto dei riferimenti alla lingua orale e all'interpretazione; (3) cambiato alcune formulazioni. Come affermano Cohen et al. (1996, p. 224), nelle griglie di valutazione ci possono essere dimensioni sottostanti alla valutazione effettuata; in questo caso, ad esempio, l'intelligibilità può avere la fluidità, la chiarezza e l'adeguatezza come dimensioni sottostanti. Griglie di valutazione con parametri multidimensionali, come in questo caso, permettono che il valutatore possa essere influenzato – con maggior probabilità – da più di una dimensione. Nell'adattamento delle scale di Carroll, Tiselius propone un minor numero di gradi all'interno della scala (sei), che permette – nonostante la minor variabilità – una rapida valutazione del testo orale. Ogni grado all'interno di ogni scala corrisponde a una descrizione verbale sintetica (ad esempio, i poli della scala relativa all'intelligibilità sono *totally intelligible* e *totally unintelligible*). Tiselius ritiene inoltre che siano più produttive delle scale senza un valore medio (e quindi con un numero di gradi pari). Prima di descrivere brevemente l'organizzazione delle scale e come sono state utilizzate

nel presente studio, è necessario chiarire che la scala per l'intelligibilità ha *6* come punteggio più alto e *1* come punteggio più basso; mentre nella scala relativa all'informatività *0* denota una corrispondenza elevata rispetto al testo originale (e quindi è il punteggio più alto), mentre *6* denota una bassa corrispondenza rispetto al testo originale (e quindi è il punteggio più basso). La scala relativa all'intelligibilità ha sei punteggi possibili; mentre la scala relativa all'informatività presenta sette valori possibili.

Nel presente studio sperimentale, ogni resa è stata trasferita su un *file* .xlsx contenente la griglia di valutazione proposta da Tiselius (2009). In tutti i file, sia intelligibilità che informatività sono state valutate suddividendo la resa in sequenze testuali, corrispondenti a *target sentence* contenenti gli stimoli e *control sentence* senza stimoli. Nella griglia relativa alla valutazione dell'intelligibilità è stato inserito il numero del *cluster* nella colonna di sinistra, seguito dalla sequenza testuale contenente esclusivamente la resa della partecipante e poi 6 colonne contenenti i punteggi possibili. È stata inoltre inserita un'ulteriore colonna a destra per inserire il risultato raggiunto in ogni *cluster*. Dopo aver valutato ogni *cluster*, per ogni *file* è stata calcolata la media totale relativa all'intelligibilità della resa. Prima di procedere alla valutazione dell'informatività delle rese, è stata completata la valutazione dell'intelligibilità delle rese, per evitare un'influenza del testo originale (presente nella griglia di valutazione dell'informatività) sulla valutazione.

Numero del cluster	Resa	Totally unintelligible - 1	Generally unintelligible - 2	Seems intelligible - 3	General idea intelligible - 4	Generally intelligible - 5	Completely intelligible - 6	Risultato per cluster
Cluster 19	non possiamo separare le decisioni e:h che hanno ripercussioni sulla nostra salute e sul benessere e che hanno e:h impatti sull'ambiente • non possiamo eh far vincere eh opzioni che ignorino i costi ambientali (1,4s)			3				3

Figura 11 Esempio di griglia di valutazione per la valutazione dell'intelligibilità.

Dopo aver completato la valutazione dell'intelligibilità di tutte le prove, nei rispettivi *file* .xlxs, è iniziata la valutazione dell'informatività delle rese, mediante un *file* .xlsx che, oltre a resa e valori possibili nella griglia, contiene il discorso originale così come è stato letto dall'oratrice. Anche in questo caso, dopo aver valutato ogni *cluster*, per ogni *file* è stata calcolata la media totale relativa all'intelligibilità della resa.

Numero del cluster	Discorso_B.1 così come è stato letto	Resa per Valutazione Informativeness	Original contains less information than rendition - 0	Without any new information - 1	No new information, strengthens the intended meaning - 2	Minor changes in meaning - 3	Gives some new information - 4	Original explains and improves - 5	Only new information - 6	Risultato per cluster
Cluster 19	No podemos separar las decisiones que repercuten en nuestra salud y bienestar de las que afectan al medio ambiente, no podemos dejarnos vencer por opciones que ignoren costes medioambientales. *[pause]*	non possiamo separare le decisioni e:h che hanno ripercussioni sulla nostra salute e sul benessere e che hanno e:h impatti sull'ambiente • non possiamo eh far vincere eh opzioni che ignorino i costi ambientali (1,4s)				3				3

Figura 12 Esempio di griglia di valutazione per la valutazione dell'informatività.

2.7.3 *User Experience Questionnaire* (UEQ)

Per quanto riguarda l'analisi dei questionari riguardanti l'usabilità dal punto di vista delle utenti del *software*, sono stati scaricati dal sito ufficiale dell'UEQ[27] due *file* .xlsx: uno che permette di effettuare un'analisi statistica dei questionari riguardanti entrambe le tecnologie (*UEQ_Compare_Products*) e uno che permette di approfondire i dati inerenti a una sola tecnologia (*UEQ_Data_Analysis_Tool*). Il secondo *file* .xlsx è stato duplicato in modo da permettere l'analisi dei questionari riguardanti Google Meet e SmarTerp in due *file* .xlsx separati. Per utilizzare in modo corretto i *file* .xlsx, sono state seguite le istruzioni presenti nel manuale di Schrepp (2023).

2.7.3 Attività di retrospezione

Dopo aver trascritto le retrospezioni con il software *Whisper*, sono state corrette manualmente mantenendo una punteggiatura semplice (punto e virgola); pertanto, le retrospezioni non sono state trascritte utilizzando le convenzioni utilizzate per la trascrizione delle rese. Le trascrizioni sono state successivamente organizzate inserendo il minutaggio dei commenti delle partecipanti utilizzando il foglio cartaceo su cui è stato appuntato il momento in cui le partecipanti hanno fermato il video. Ogni retrospezione è stata trascritta su un *file* .doc che contenesse solo l'output testuale di una partecipante. Le trascrizioni sono state successivamente importate sul *software Taguette*[28] (Rampin et al., 2021), che è stato utilizzato per la codifica dei dati qualitativi di tutte le retrospezioni. Tutte le rese sono state oggetto di una lettura approfondita e ripetuta nel corso

[27] https://www.ueq-online.org/ (Ultimo accesso: 28 aprile 2024)
[28] https://app.taguette.org/ (Ultimo accesso: 28 aprile 2024)

di più giorni. Successivamente, ogni *file* contenente la retrospezione di ogni partecipante è stato aperto su *Taguette* ed è iniziata l'analisi tematica del *file* sottolineando gli elementi rilevanti in funzione degli obiettivi e delle domande dello studio. Ogni sequenza testuale rilevante è stata sottolineata ed è stata codificata e, al termine della codifica di tutti i *file*, alcune categorie codificate sono state raggruppate in categorie più ampie con altre dallo stesso filo tematico. In particolare, le categorie emerse dalle retrospezioni sono le seguenti: (1) numeri; (2) entità denominate; (3) terminologia specialistica; (4) percezione della latenza dell'ASR; (5) impatto dello strumento sulle strategie dell'interprete; (6) omissioni; (7) calchi; (8) allontanamento dello sguardo dagli stimoli; (9) impatto dello strumento sulla prosodia dell'interprete; (10) percezione della prosodia del discorso; (11) gestione input visivo e auditivo; (12) monitoraggio della resa; (13) recupero delle informazioni; (14) organizzazione dell'input visivo di SmarTerp; (15) ridondanza informativa; (16) punteggiatura della trascrizione di Google Meet; (17) traduzione a vista; (18) paragone tra Google Meet e SmarTerp; (19) percezione dell'usabilità dello strumento; (20) prospettiva didattica e (21) aspettative. Come verrà illustrato ampiamente nella fase di discussione dei dati, le categorie (14) e (15) sono specifiche per il CAI *tool* di SmarTerp, mentre le categorie (16) e (17) sono specifiche per la trascrizione automatica di Google Meet. Tutte le altre categorie sono emerse nelle retrospezioni relative a entrambi gli strumenti.

2.7.4 *Eye-tracking*

2.7.4.1 Dati qualitativi

Per un'analisi che miri all'osservazione descrittiva del movimento oculare nel corso della presentazione degli stimoli, i dati di *eye-tracking* sono stati analizzati utilizzando un *software* per l'annotazione di *file* multimediali: *ELAN*. Ogni video in formato .mp4 contenente il movimento oculare durante l'IS è stato scaricato dal *software* di analisi dei dati dell'*eye-tracker* utilizzato: *EyeLink Data Viewer*. Successivamente il video è stato importato su un *software* per il *video-editing* (*Wondershare Filmora*) in cui è stato ritagliato il video facendo coincidere l'inizio con il momento in cui è stato premuto il pulsante PLAY. Una volta ritagliato il video, è stato opportunamente esportato in formato .mp4 in una cartella apposita utilizzata per i dati grezzi. Successivamente è stato recuperato il *file* audio (.mp3) contenente la resa della partecipante. La resa è stata ritagliata, sempre utilizzando *Filmora*, allineando audio e video.

Anche in questo caso, una volta che è stato ritagliato il *file* multimediale, è stato opportunamente esportato in una cartella apposita (in formato .mp3).

Successivamente il *file* .mp3 (contenente l'audio della resa) e il *file* .mp4 (contenente il movimento oculare durante l'IS) sono stati importati sul *software ELAN*. Sul *software* è stato creato un unico *type: qualitative eye-tracking*. Successivamente è stato creato un livello (*tier*) per ogni *target sentence* contenente gli stimoli (utilizzando le abbreviazioni definite in fase di analisi del discorso). Dopodiché, sono state segmentate tutte le *target sentence*, con dei segmenti lunghi tanto quanto la resa di tutta la *target sentence*. Dopo aver segmentato tutto il discorso, in ogni segmento è stata inserita la descrizione del movimento oculare durante tutto il *task* (non solo durante, ma anche prima e dopo la comparsa degli stimoli). Successivamente, le annotazioni sono state esportate in *file* .txt e sono state copiate e incollate in un *file* .doc. Questo procedimento è stato ripetuto per tutte le 22 rese dell'attività sperimentale.

Sia per il discorso A.1 che per il discorso B.1 sono stati identificati 51 stimoli e per ogni stimolo è stato determinato se: (1) la partecipante guarda lo stimolo prima di iniziare l'output; (2) la partecipante guarda lo stimolo dopo aver iniziato la traduzione e non si riformula; (3) la partecipante guarda lo stimolo dopo aver iniziato la traduzione e si riformula migliorando l'output; (4) la partecipante guarda lo stimolo dopo aver iniziato la traduzione e si riformula peggiorando l'output; (5) la partecipante non guarda lo stimolo; (6) il dato non è disponibile. È necessario chiarire che gli stimoli selezionati sono presenti sia nell'input visivo-verbale del CAI *tool* di SmarTerp che in quello della trascrizione automatica di Google Meet.

Per il discorso A.1, gli stimoli selezionati sono: *10.195 millones // 5.817 // 4.378 // plan España Puede // 7.397 millones // 442 millones // 53 millones // proyectos generadores de empleo // bono social térmico // 1.233 millones // FEDER // Programa DUS 5000 // Secretaría de Estado // 579 millones // 2021 // 2,85 por ciento // 22 // megavatios // 2032 // termostato // Protección Civil // falta de regulación // depuración // percolación // floculación // suministro // plantas potabilizadoras // gobierno del agua // cohesión territorial y social // legislatura // 102.500 // megavatios // 7 // 14.650 // Gobierno socialista // 23 por ciento // siete años // Comunidad Valenciana // gases de efecto invernadero // 2030 // potencia total instalada // 157 // gigavatios // sector de movilidad y transporte // 28 millones // 2027 // 2019 // región de Murcia // ahorro de energía acumulado // 1.300 // ciudadanía.*

Per il discorso B.1, gli stimoli selezionati sono: *Bastilla // señorías // igualdad de oportunidades // 48.570 millones // sociedad civil organizada // sindicatos // electrificación // Congreso de los*

diputados // descarbonización // CMNUCC // mitigación // proyecto de ley // toma de decisión // PSOE // 2 años // 25.000 // vuelos cortos // 200.000 // 38 millones // legislatura // 58,5 millones // 100 millones // modelo productivo // mitigación // organismos internacionales // Estados miembros de la UE // 723.800 millones // empleo cualificado // ecologización // trabajo decente // despoblación // gestión del agua // movilidad sostenible // plan de recuperación y resiliencia // 163.200 millones // 83.000 millones // 80.200 millones // subvenciones // lucha contra el cambio climático // 2030 // 20 por ciento // 1990 // 40 // hidrógeno verde // referente // 2028 // 20 por ciento // 21.000 millones // Andalucía // 10 // iridio.

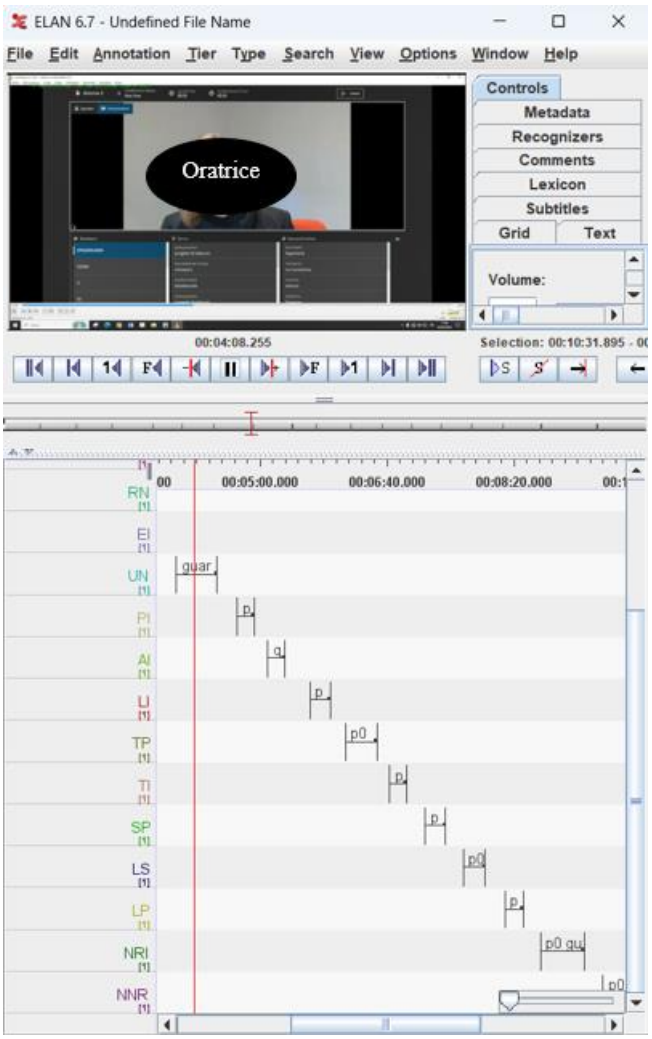


Figura 13 Esempio dell'Interfaccia Utente del software ELAN.

Per ogni discorso è stato compilato un *file* .xlsx contenente tutti i dati per quel discorso. Di seguito, viene proposto un esempio:

Stimolo	Guarda lo stimolo prima di iniziare l'output	Guarda lo stimolo dopo aver iniziato la traduzione e non si riformula	Guarda lo stimolo dopo aver iniziato la traduzione e si riformula migliorando l'output	Guarda lo stimolo dopo aver iniziato la traduzione e si riformula peggiorando l'output	P non guarda lo stimolo	Dato non disponibile
cohesión territorial y social	0					
legislatura		0				
102.500				0		
megavatios	0					
7					0	
14.650					0	
Gobierno socialista		0				
23 por ciento					0	
siete años					0	
Comunidad Valenciana						0
gases de efecto invernadero		0				

Tabella 15 Esempio del file .xlsx contenente i dati quantitativi relativi al comportamento oculare della partecipante durante la comparsa dello stimolo nell'input visivo.

2.7.4.2 Dati quantitativi

I dati quantitativi di *eye-tracking* utilizzati fanno riferimento al *dwell time* in ogni *area di interesse* (o *area of interest*, di seguito AOI). Complessivamente, sono state analizzate 22 registrazioni con *eye-tracking*, delle quali 11 nella condizione sperimentale contenente la trascrizione automatica di Google Meet e 11 nella condizione sperimentale contenente i suggerimenti del CAI *tool* di SmarTerp. L'analisi quantitativa del *dwell time* sulle AOI permette di quantificare la distribuzione dell'attenzione visiva sui diversi input visivi. Le aree di interesse configurate sono AOI statiche e con un *periodo di interesse* (o *period of interest*) corrispondente alla durata del video input. Per pulire i dati e garantirne l'utilizzabilità, in ogni registrazione è stato inserito un *period of interest* nel quale sono state applicate le AOI utilizzando il *software EyeLink Data Viewer* per non applicare l'area di interesse anche nel momento iniziale di prova. Successivamente, sono state configurate delle aree di interesse che sono state applicate a tutti i video con la stessa condizione sperimentale (data la stessa distribuzione degli elementi di interesse nell'interfaccia utente con la stessa condizione).

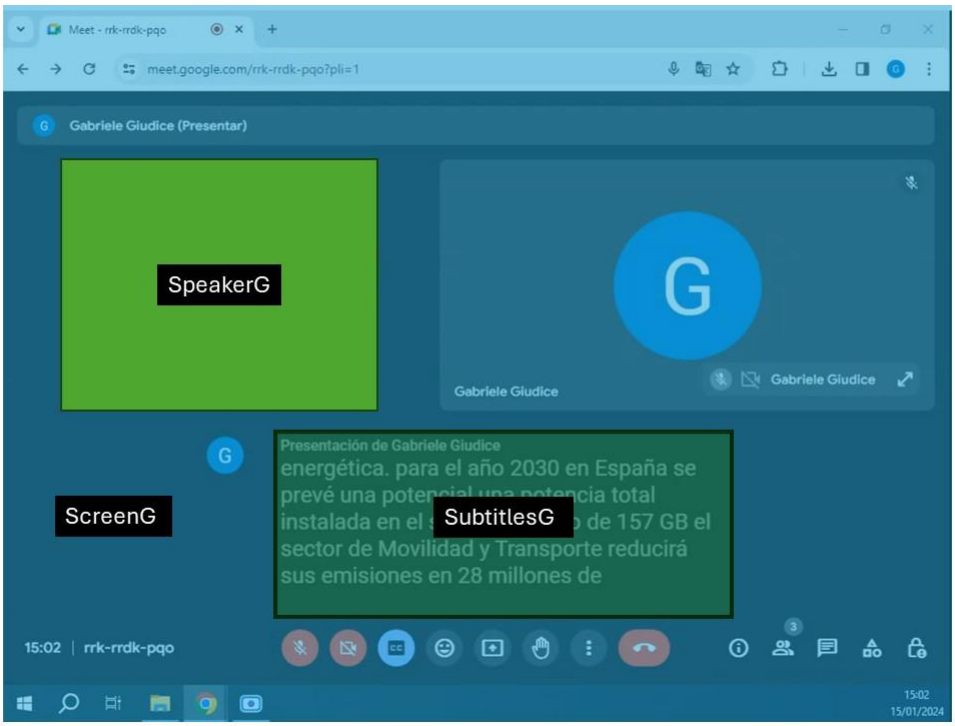


Figura 14 AOI impostate per la condizione sperimentale contenente la trascrizione automatica di Google Meet.

Le AOI impostate per la condizione sperimentale con la trascrizione automatica di Google Meet sono le seguenti:

- *SpeakerG*: contenente il video originale dell'oratrice;
- *SubtitlesG*: contenente la trascrizione generata automaticamente da Google Meet;
- *ScreenG*: contenente lo schermo con tutti gli elementi del *Display Screen* così come è stato mostrato a ogni partecipante.

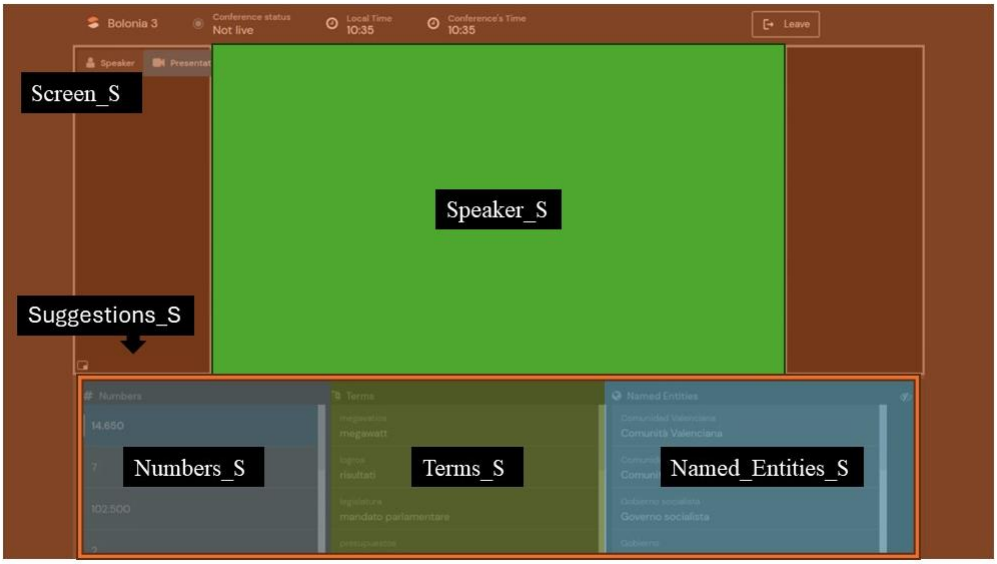


Figura 15 AOI impostate per la condizione sperimentale contenente il CAI tool di SmarTerp.

Le AOI impostate per la condizione sperimentale con il CAI *tool* di SmarTerp sono le seguenti:

- *Speaker_S*: contenente il video dell'oratrice;
- *Numbers_S*: contenente la colonna del CAI *tool* relativa ai numeri;
- *Terms_S*: contenente la colonna del CAI di SmarTerp relativa alla terminologia specialistica;
- *Named_Entities_S*: contenente la colonna del CAI dedicata alle entità denominate:
- *Suggestions_S*: contenente tutte e tre le colonne del CAI *tool* (numeri, termini specialistici e entità denominate);
- *Screen_S*: contenente lo schermo con tutti gli elementi del *Display Screen* così come è stato mostrato a ogni partecipante.

Le AOI impostate non prendono in considerazione le singole parole, ma aree tematiche di interesse; per questo motivo, queste non forniscono dati quantitativi sulla distribuzione dell'attenzione tra parole spagnole e rispettivi traducenti proposti dal CAI *tool* di SmarTerp. La scelta di aree di interesse più grandi, oltre che fornire parametri statistici più ampi, risponde al bisogno di limitare il margine di errore fornendo al contempo dei termini di paragone per le due tecnologie. Infatti, in entrambi i casi è presente un'AOI contenente il riquadro con l'oratrice e in entrambe le configurazioni c'è un'AOI dedicata al materiale testuale generato dalla tecnologia (nel caso di Google Meet c'è un'AOI dedicata alla trascrizione automatica e nel caso di SmarTerp è stata configurata un'AOI comprensiva delle tre colonne del CAI). Oltre alle tre AOI paragonabili (*ScreenG* e *Speaker_S*; *ScreenG* e *Screen_S*; *SubtitlesG* e *Suggestions*_S), sono state impostate tre ulteriori AOI per la condizione sperimentale con il CAI *tool* di SmarTerp (*Numbers_S*; *Terms_S* e *Named_Entities_S*) per approfondire la distribuzione dell'attenzione sui tre elementi costitutivi del CAI *tool*.

2.8 Esperimento Pilota

Tutti i materiali e i metodi sperimentali (formazione asincrona, metodologia proposta, analisi dei dati) sono stati pilotati circa 10 giorni prima dall'inizio della raccolta dati. Lo studio pilota è stato condotto con una dodicesima partecipante che ha le stesse caratteristiche del campione coinvolto nello studio principale e che ha raggiunto un punteggio pari a 83,59 nel LexTale. La partecipante – in media con il campione dello studio principale – studia spagnolo (al momento

dello studio) da circa 13 anni e quantifica in una scala da 1 a 10 la sua padronanza dello spagnolo attorno a un valore di 8 punti; ha inoltre la lingua spagnola come lingua B nella propria combinazione linguistica. Prima dello studio pilota, sono stati preparati i messaggi da inviare rispettivamente in concomitanza con l'invio della prima sessione di formazione, con l'invio della seconda sessione di formazione e il giorno prima dell'esperimento (che sarebbero poi stati utilizzati anche per tutte le partecipanti dello studio principale). Durante lo studio pilota è stato seguito il protocollo predefinito, in una versione contenente al fianco di ogni passaggio uno spazio bianco per inserire eventuali commenti. Durante lo studio pilota, sono stati inseriti nel protocollo gli elementi riguardanti la calibrazione dell'*eye-tracker* e sono state inserite piccole modifiche (ad esempio, il *briefing* non era stato inserito nel protocollo ed è stato dimenticato nel pilota). Alla fine dell'esperimento pilota, è stato chiesto alla partecipante di quantificare il tempo necessario per la formazione asincrona (che ha quantificato su due ore complessive). Dopo l'attività sperimentale, tutti i dati dello studio pilota sono stati analizzati e paragonati, in modo da poter validare tutta la metodologia; tuttavia, dato il coinvolgimento di una sola partecipante nell'esperimento pilota, non è stata possibile la validazione della metodologia relativa all'UEQ, in quanto richiede almeno due partecipanti per poter generare dei valori statistici. Durante l'analisi dei dati dell'esperimento pilota, sono stati creati dei *file* con nomi generici da poter riutilizzare per tutti le partecipanti (ad esempio, il file *valutazione_informativeness_A.1_CodicePartecipante_NomeProva* che è stato riutilizzato per la valutazione dell'informatività di tutte le prove contenenti il discorso input A.1).

CAPITOLO 3. Analisi e discussione dei dati relativi all'accuratezza delle rese e all'*User Experience Questionnaire* (UEQ)

In questo capitolo vengono presentati e discussi i dati raccolti relativamente all'accuratezza delle rese e ai risultati dell'*User Experience Questionnaire*. Verranno presentati in primo luogo i dati quantitativi relativi alla *performance* nelle singole rese mediante l'analisi e la discussione dei risultati dell'analisi statistica dei *success rate* raggiunti dalle partecipanti nella resa dei singoli *task* proposti (§3.1). Successivamente, vengono presentati i dati quali-quantitativi che mirano a stabilire una valutazione dell'intelligibilità (§3.2.1) e dell'informatività (§3.2.2) delle rese in entrambe le condizioni sperimentali. Terminata la valutazione delle rese e l'analisi comparativa dei dati, verrà proposta l'analisi quantitativa dei dati provenienti dagli *User Experience Questionnaire* (§3.3), con cui verrà discussa l'usabilità degli strumenti proposti, così come viene percepita da parte di interpreti in formazione. Per l'analisi statistica dei dati quantitativi è stato utilizzato il *software* Jamovi[29] (The jamovi project, 2024).

3.1 *Task Success Rate*

Questo paragrafo mira a illustrare i risultati dell'analisi statistica dei *success rate* ottenuti dalle partecipanti mediante la resa dei *task* proposti. Per procedere a un'analisi approfondita dei dati disponibili, il paragrafo prevede inizialmente due sezioni: la prima è dedicata interamente alle statistiche descrittive dei dati quantitativi relativi ai *success rate* raggiunti nei *task* proposti nella condizione sperimentale contenente la trascrizione automatica di Google Meet; la seconda prevede invece l'analisi dei dati quantitativi relativi ai *success rate* ottenuti nella condizione sperimentale con il CAI *tool* di SmarTerp. In ogni sezione, l'analisi dei dati verrà suddivisa in tre categorie tematiche corrispondenti al grado di complessità dei *task* proposti: *task a complessità bassa* (NI, SI, EI, PI, AI, LI, TI), *task a complessità media* (RN, TP, SP, LS, LP) e *task a complessità elevata* (NRS, NLP, UN, NRI, NNR). In seguito, verrà proposta un'analisi comparativa tra i *success rate* ottenuti nella condizione sperimentale contenente la trascrizione automatica di Google Meet e quella contenente il CAI *tool* di SmarTerp. Prima di procedere alla descrizione dei risultati dell'analisi statistica dei dati quantitativi dei *success rate*, è opportuno ricordare che questi non prendono in considerazione eventuali omissioni e

[29] https://www.jamovi.org/ (Ultimo accesso: 6 giugno 2024)

inaccuratezze dell'ASR specifico del *tool* proposto nella condizione sperimentale; questi fattori non sono stati considerati nel presente studio, in quanto i due strumenti sono molto diversi e, quindi, difficilmente comparabili da questo punto di vista.

3.1.1 Task Success rate con la trascrizione automatica di Google Meet

La Tabella 16 contiene la descrizione dei dati quantitativi relativi ai *success rate* ottenuti per i *task* NI, SI, EI, PI, AI, LI, TI (a bassa complessità) nella condizione sperimentale contenente la trascrizione automatica di Google Meet.

	NI	SI	EI	PI	AI	LI	TI
Media	35.0	50.0	50.0	40.0	50.0	30.0	0.00
Mediana	0.00	50.0	50.0	0.00	50.0	0.00	0.00
Deviazione standard	47.4	52.7	52.7	51.6	52.7	48.3	0.00
Minimo	0	0	0	0	0	0	0
Massimo	100	100	100	100	100	100	0

Tabella 16 Descrizione dei valori di statistica descrittiva dei task success rate raggiunti nella condizione sperimentale contenente la trascrizione automatica di Google Meet per i task a bassa complessità.

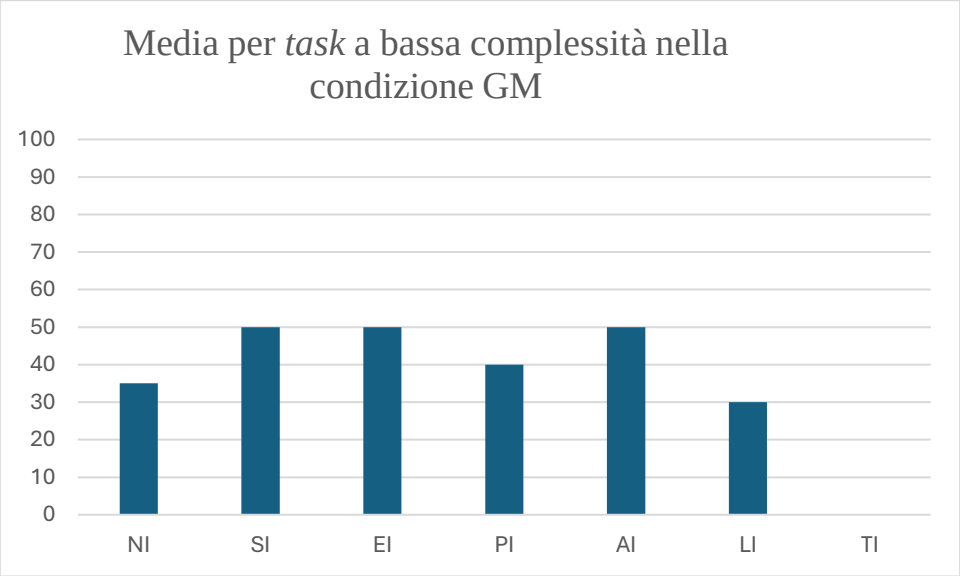


Figura 16 Grafico relativo ai valori medi raggiunti nei task success rate per i task a bassa complessità per la condizione GM.

Osservando i dati, è necessario ricordare che i *task* a bassa complessità contengono un solo stimolo; nella maggior parte dei casi, la presenza di un solo stimolo ha portato a valori pari a 0 (non resa) o 100 (resa corretta), senza valori intermedi; inoltre, questo spiega anche la presenza

di valori minimi e massimi uguali a 0 e 100 per tutti i *task* proposti. Osservando i valori medi dei *success rate* ottenuti nei singoli *task*, spicca la media pari a 0 per il *task* TI (tecnicismo isolato): nessuna partecipante ha infatti interpretato correttamente il tecnicismo proposto in quel *task*. Infatti, nonostante i tecnicismi presenti nei discorsi A.1 e B.1 siano stati inseriti nei glossari proposti durante la formazione, nessuna partecipante è riuscita a recuperare il termine in modo sufficientemente rapido da poter interpretare correttamente lo stimolo presente nel *task* TI. In particolare, nei discorsi A.1 e B.1 gli stimoli del *task* TI sono rispettivamente *plantas potabilizadoras* ed *ecologización*. Nel caso di *ecologización* (in italiano, *inverdimento*) nessuna partecipante ha recuperato dalla memoria a lungo termine il traducente corretto (o ipoteticamente il traducente non era presente nella memoria a lungo termine) e nella maggior parte dei casi le partecipanti hanno optato per il calco *ecologizzazione* (in questo caso, inadeguato, in quanto nei corpora interrogati questo termine è decisamente meno frequente nei contesti istituzionali, come quello simulato nell'esperimento). Nel caso di *plantas potabilizadoras*, le partecipanti hanno invece spesso adottato strategie di sintesi e di ampliamento descrittivo che tuttavia hanno spesso portato a omissioni e inaccuratezze, di seguito vengono proposti due esempi:

> *Plantas potabilizadoras*
>
> P10
>
> <le: inst/installazioni: urbane> ••• comprendono un sistema di ehm/••• mh >**sistemi di pot/<• di acqua potabile** ••• che necessitano ovviamente di investimenti ••
>
> P2
>
> le installazioni urbane (1,7s) >comprendono< **un sistema complesso di impianti eh che rendono l'acqua potabili** • ma che hanno bisogno • di investimenti e di manutenzione •

Nel caso del *task* NI (numerale isolato) solo 3 partecipanti su 10 hanno reso correttamente il numero, mentre la maggior parte delle partecipanti non hanno reso correttamente il numerale (6 su 10); inoltre, P2 ha raggiunto un *success rate* pari a 50 (in quanto ha tradotto correttamente due cifre e le altre due sono state omesse). Nell'accuratezza del numerale ha particolarmente influito la struttura dissimmetrica *x mil x millones* che in italiano corrisponde a *x miliardi*, nella maggior parte dei casi la struttura è stata calcata, con conseguenti errori; mentre in altri casi, la

struttura dissimmetrica ha causato incomprensioni del numerale. Si propone di seguito un esempio della resa del numerale *7 mil 397 millones*:

> P7
>
> la: • transizione energetica •• in quanto a quest'ultima (3,1s) sette/•• **oltre sette milioni di euro** • vengono • e:h dedicati dal budget •

Per quanto riguarda il *task* SI (dissimmetria superficiale isolata), 5 partecipanti su 10 hanno reso correttamente la dissimmetria superficiale, mentre le altre 5 hanno ottenuto valori pari a 0. Anche nel caso del *task* EI (entità nominale isolata) solo 5 partecipanti su 10 hanno reso correttamente lo stimolo.

Il *task* PI (dissimmetria profonda isolata) presenta valori sotto la soglia del 50%, in quanto solo 4 partecipanti su 10 hanno reso correttamente la dissimmetria presente nel discorso. Ad esempio, nel caso della dissimmetria *"se trata de que…"*, questa è stata prevalentemente ricollegata alla dimensione dell'obbligatorietà (e, quindi, è stata spesso tradotta con il verbo *dovere*), perdendo quindi la dimensione dell'incoraggiamento e della sollecitazione.

> *Se trata de que…*
>
> P9
>
> l'elettrificazione • è • la chiave della: sostenibilità nel tempo (2,1s) <il congresso dei deputati **deve** collaborare> • in questa trasformazione • approvando le misure inevitabili • favorevoli a una <de:carbonizzazione> (1,2s)
>
> P8
>
> <l'elettrificazione è la chiave per la sostenibilità> • nel tempo ••• il Congresso dei deputati **deve** cooperare durante questa trasformazione • approvando • i mezzi • inevitabili ••• verso la decarbonizzazione ••

Il *task* AI (acronimo isolato) è stato reso correttamente solo nel 50% delle rese, mentre è stato reso in modo sbagliato o omesso nelle restanti rese. Il *task* LI (sfida lessicale isolata) ha invece riportato un maggior numero di valori negativi che positivi, è infatti stato reso correttamente solo in 3 rese su 10. In molti casi, nel caso di questo stimolo, si osserva un possibile *spillover*

effect (Seeber & Kerzel, 2012) significativo sull'intera resa della *target sentence*, a titolo illustrativo si propongono questi esempi:

> Discorso originale:
>
> Los expertos **achacan** a la falta de regulación, el descontrol que se ha producido en los servicios de extinción de los incendios forestales.
>
> Resa P3:
>
> gli esperti (2,1s) mh ehm dicono che spe/per mancanza di eh ehm regola/mh >di regolamenti< e mh di mancanza di controllo • nei servizi si producono tutti una serie di incendi forestali •••
>
> Resa P2:
>
> gli esperti (1s) sottolineano la mancanza di un regolamento • il de/descontrollo/il/la mancanza di controllo prodotto nel >servizi di estinzione< degli incendi boschivi è • evidente ••

La Tabella 17 contiene la descrizione dei dati quantitativi relativi ai *success rate* ottenuti per i *task* RN, TP, SP, LS, LP (a media complessità) nella condizione sperimentale contenente la trascrizione automatica di Google Meet.

	RN	TP	SP	LS	LP
Media	10.0	40.0	70.0	55.0	65.0
Mediana	0.00	33.3	75.0	50.0	50.0
Deviazione standard	17.5	30.6	35.0	28.4	33.7
Minimo	0	0.00	0	0	0
Massimo	50	100	100	100	100

Tabella 17 Descrizione dei valori di statistica descrittiva dei task success rate raggiunti nella condizione sperimentale contenente la trascrizione automatica di Google Meet per i task a media complessità.

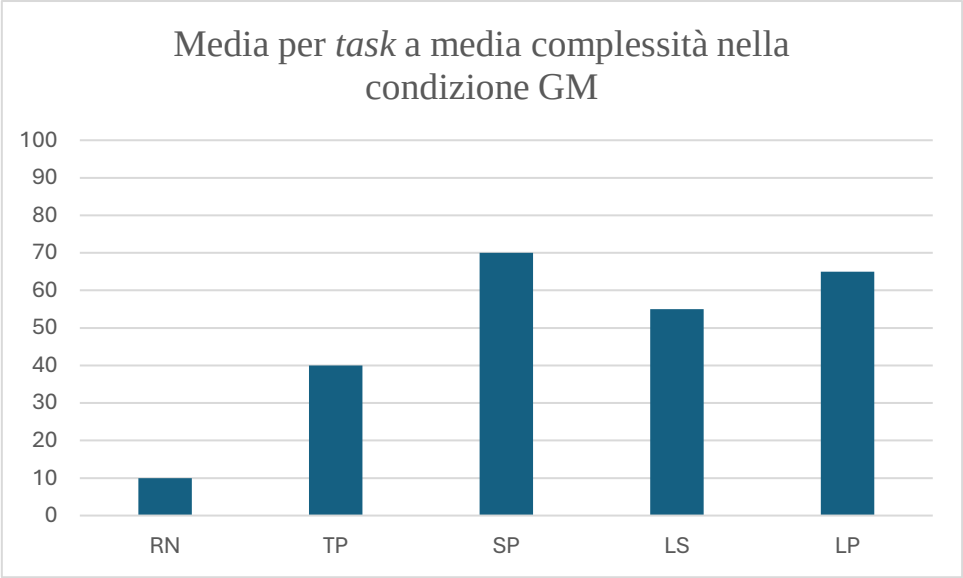


Figura 17 Grafico relativo ai valori medi raggiunti nei task success rate per i task a media complessità nella condizione contenente la trascrizione automatica di Google Meet.

Tra i *success rate* ottenuti nella resa dei *task* a media complessità, spicca la media pari a 10 del *task* RN (referente complesso + numerale) che in 7 rese su 10 ha visto un punteggio pari a 0, in 2 casi pari a 25 e in un solo caso pari a 50. In questo *task*, la tendenza al calco della struttura spagnola per il numerale con ordine di grandezza dei miliardi si è aggiunta alla tendenza a calcare gli acronimi, senza strategie di sintesi o di ampliamento, portando a calchi di acronimi semanticamente vuoti in italiano; di seguito se ne offrono alcuni esempi:

> Stimoli in *task* RN del discorso B.1: MRR [Mecanismo de Recuperación y Resiliencia] //723.800 millones
>
> P8
>
> il fun/il fon/il fondo del M R R • è uno strumento fondamentale • per • riforme • e • investimenti negli stati membri dell'Unione Europea • per un valore totale di ••• settemila (2,1s) settecentoventitré mil/a/ >ottocent<••• /settecento ventitré miliardi ottocento mila euro •••
>
> Stimoli in *task* RN del discorso A.1: 1233 millones //FEDER
>
> P10
>
> per questo •• queste politiche (1,1s) avranno • >mille e duecento trentatré milioni< di euro ••• che possono essere finanziati con fondi F E D E R (1,1s)

Come si osserva nella Tabella 17, un altro *task* che presenta una media relativa ai *success rate* inferiore a 50 è il *task* TP (tecnicismi + dissimmetria profonda). Contrariamente alle aspettative, nel *task*, la dissimmetria profonda è stata resa correttamente in molti casi, mentre i tecnicismi hanno portato molto spesso a errori, nonostante i tecnicismi fossero presenti nei glossari forniti nel corso della formazione (e quindi potenzialmente recuperabili dalla memoria a lungo termine). Di seguito si osservano alcuni esempi di rese del *task* TP contenuto nel discorso B.1:

> Stimoli del *task* TP in B.1 (in grassetto gli stimoli resi correttamente, anche negli esempi successivi): vehículos compartidos// circulación rodadada// a la mayor brevedad
>
> Resa P1 (*Success Rate*: 33,33):
>
> l'investimento nelle infrastrutture che•/che • richiede la mobilità sostenibile (2,2s) e anche la riduzione della circolazione sono • delle questioni che dobbiamo trattare **nel a:h più breve tempo possibile** •
>
> Resa P8 (*Success Rate*: 33,33):
>
> gli investimenti nelle infrastrutture • necessari • per la mobilità sostenibile • gli investimenti in **veicoli condivisi** e la riduzione della circolazione sui mezzi sono questioni •• che/e:h >delle quali bisogna< parlare • a: breve ••

Nel caso del *task* SP, 9 partecipanti su 10 hanno reso correttamente almeno uno dei due stimoli, senza cadere in calchi morfologici. In molti casi, *success rate* più bassi non sono stati causati da calchi sintattici o da incomprensioni, ma da omissioni:

> Testo originale:
>
> **Insistimos en que** es necesario renovar las infraestructuras y desplegar la tecnología necesaria para calcular de manera fiable el consumo de agua, **sin dejar nada por tocar**.
>
> Resa P3:
>
> >sp/per noi< è veramente fondamentale rinnovare le infrastrutture e usare le tecnologie necessarie •• per calcolare in modo efficace il consumo di acqua ••• senza lasciar <niente ehm in disparte> (1,3s)

Nel caso del *task* LS (sfida lessicale + dissimmetria superficiale), 7 partecipanti su 10 hanno ottenuto un *success rate* pari a 50, in quanto frequentemente solo uno dei due stimoli è stato reso correttamente, prevalentemente la dissimmetria morfosintattica:

> Discorso originale:
>
> **Con toda modestia**, pienso que es fundamental impulsar políticas para la cohesión territorial y social: desde la rebaja de las facturas eléctricas hasta el aumento de las **nóminas**.
>
> Resa P11 (*Success Rate*: 50):
>
> modestamente penso che sia • necessario • promuovere politiche per la coesione sociale e territoriale •• sia da un ribasso delle: bollette • dell'energia (1,4s)

Il *task* LP (sfida lessicale + dissimmetria profonda) è stato reso correttamente in 4 casi, mentre in 5 casi è stato reso correttamente con un *success rate* pari a 50 (quindi, un solo stimolo è stato reso correttamente) e in un solo caso nessuno stimolo è stato reso correttamente. Tendenzialmente, quando è stato reso un solo stimolo, è stata sbagliata la struttura dissimmetrica e non la sfida lessicale, di seguito se ne offrono due esempi:

> Discorso originale:
>
> Señorías, **con independencia de** nuestra orientación política, ¡qué **ilusión** sería ver a nuestros sobrinos, a nuestros hijos y a nuestros nietos vivir en un mundo más respirable!
>
> Resa P2 (*Success Rate*: 50):
>
> onorevoli (1,8s) con l'indipendenza della nostra/>del nostro orientamento politico< sarebbe • fantastico poter vedere i nostri nipoti i nostri figli vivere in un mondo più respirabile
>
> Resa P11 (*Success Rate*: 50):
>
> onorevoli (1,8s) con • >indipendenza< del nostro orientamento politico • sarebbe stupendo vedere i nostri figli • i nipoti ••• vivere in un mondo in cui si respira meglio •

La Tabella 18 contiene una descrizione dei dati quantitativi relativi ai *success rate* ottenuti per i *task* NRS, NLP, UN, NRI, NNR (a elevata complessità) nella condizione sperimentale contenente la trascrizione automatica di Google Meet.

	NRS	NLP	UN	NRI	NNR
Media	45.9	64.9	71.4	94.9	77.5
Mediana	49.5	58.3	71.4	100	78.1
Deviazione standard	22.6	25.3	16.5	6.86	14.2
Minimo	20	33.3	42.9	83.3	49.6
Massimo	80	100	85.7	100	100

Tabella 18 Descrizione dei dati relativi ai success rate raggiunti nei singoli task ad elevata complessità nella condizione sperimentale con la trascrizione automatica di Google Meet.

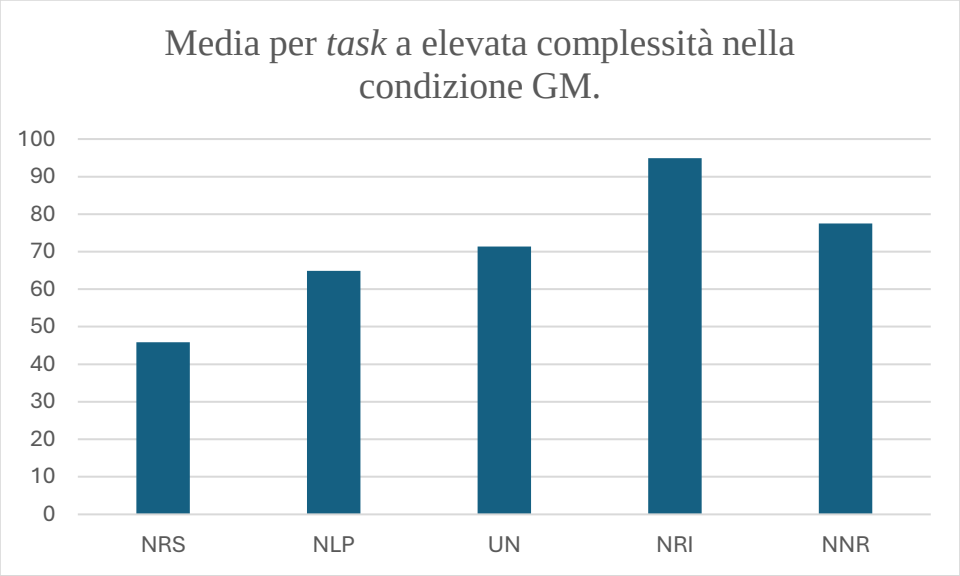


Figura 18 Grafico relativo ai valori medi raggiunti nei task success rate per i task a elevata complessità nella condizione contenente la trascrizione automatica di Google Meet.

Osservando i dati disponibili, sembra particolarmente rilevante il *task* NRS (numerali + referente complesso + dissimmetria superficiale), nel quale cinque partecipanti (corrispondenti alla metà del campione) hanno ottenuto dei valori medi inferiori a 50. In molti casi, hanno influito sulla valutazione del *task* l'omissione della dissimmetria superficiale (che non ha mai causato controsensi) e il mantenimento della struttura dissimmetria *x mil x millones* nella resa in lingua italiana, di seguito vengono proposte due rese che ben esemplificano la tendenza generale:

Task NRS del discorso A.1:

El presupuesto asciende en su conjunto a **10 195 millones** de euros, de los que **5817** corresponden a recursos del presupuesto nacional y **4378** corresponden a recursos del **plan España Puede**, tal y como les explicaré **extensa y detalladamente**.

Resa P2:

il bilancio •• vale all'incirca • dieci miliardi di euro dei quali • cinque miliardi •• all'incirca •• corrispondono alle risorse del bilancio nazionale e i restanti quattro eh quattro corrispondono ai eh (1,7s) alle risorse del Piano Spagna >come vi spiegherò< •

Valutazione del task NRS di P2:

10 195 millones (x50%)// 5817 millones(x50%)// 4378 millones(x50%)//plan España Puede (e)// extensa y detalladamente(o)

Resa P11:

il bilancio • ammonta >totalmente< • a • dieci virgola >cento novanta cinque< milioni di euro •• di cui >cinque mila ottocento diciassette< corrispondono a risorse del bilancio nazionale • e quattromila >trecento< settantotto • corrispondono a risorse di • España Puede •• così come spie:gherò in modo dettagliato (1,1s)

Valutazione del task NRS di P11:

10 195 millones (e)// 5817 millones(e)// 4378 millones(e)//plan España Puede (e)// extensa y detalladamente(k)

In generale, la densità informativa dei *task* non ha avuto ripercussioni significative sulla resa dell'informatività dei *task* a elevata complessità presenti nei due discorsi, così come emerso dai valori medi dei *task* NLP, UN, NRI e NNR, che sono tutti superiori a 50. In generale, i numerali con l'ordine di grandezza dei milioni – presenti nel *task* NLP (numerali + sfide lessicali + dissimmetrie profonde) – non hanno causato difficoltà generalizzabili, mentre hanno spesso influito l'omissione della sfida lessicale e la mancata chiarezza nella resa della struttura dissimmetrica a dominanza di elaborazione profonda. Tuttavia, ci sono stati rari casi in cui la

concomitanza degli stimoli ha portato alla perdita informativa oltre che a errori semantici anche al livello dei numeri con l'ordine di grandezza dei milioni, ne è un esempio la resa di P4:

> *Task* NLP in B.1:
>
> Señorías, les voy a dar un ejemplo: **justamente** a lo largo de la pasada legislatura se emitieron más de **58,5 millones** de toneladas de CO2 al año, y **no se** adoptaron medidas adecuadas **más que** en algunos sectores. Además, **debido a** la importación de combustibles fósiles, se gastaron más de **100 millones** de euros diarios. Ahora tenemos que **aferrarnos** a nuestros objetivos sostenibles.
>
> Resa di P4:
>
> adesso vi faccio un esempio (1,7s) durante • >lo scorso mandato< • sono stati emessi • cinquantotto •• mh miliardi • cinquecento milioni di tonnellate di C O due all'anno •• e non sono state adottate delle misure adeguate • >come in altri settori<• inoltre (1,5s) ahm a causa del/dell'importazione dei combustibili fossili • sono stati sp ••/usati • più di cento milioni di euro • >ogni giorno< (1,7s) adesso dobbiamo concentrarci sul nost/sui nostri obiettivi sostenibili •
>
> Valutazione della resa:
>
> Cluster 1: justamente(o) // 58,5 millones(e) // no se…más que…(e); cluster 2: debido a(k) // 100 millones(k) // aferrarnos(e)

Nel caso del *task* UN (unità informativa con numerali), 8 partecipanti su 10 hanno ottenuto *task success rate* superiori a 50. In termini generali, gli anni presenti nel *task*, l'unità di misura complessa e i numeri con ordine di grandezza inferiore alle centinaia sono stati resi correttamente. La maggior parte delle difficoltà si sono concentrate sulla resa degli acronimi, che sono stati spesso calcati nella resa in italiano, giungendo quindi a soluzioni semanticamente vuote in lingua italiana, ecco alcuni esempi:

> *Task* UN in B.1:
>
> En **2030** proponemos que España haya reducido sus emisiones al menos un **20 %** con respecto a las de **1990**, alcanzando los **40 gramos de dióxido de carbono equivalentes por megajulio** [gCO2eq/MJ], tal y como nos recomienda la **AIE** [Agencia Internacional de la Energía]. Tenemos **7** años.

Resa P9:

nel duemila trenta (3,1s) speriamo che la Spagna abbia ridotto le proprie emissioni almeno del venti per cento • rispetto ai livelli del mille novecento novanta ••• e speriamo che abbia raggiunto • i quaranta grammi di ••• C O due ••• <equivalenti> •• per • megajoule (1,7s) esattamente come ci è stato ••• raccomandato •

Valutazione della resa di P9:

AIE(o) // gCO2eq/MJ(k) // 1990(k) // 2030(k) // 40(k) // 20%(k) // 7(o)

Resa p5:

nel duemila e trenta •• >proponiamo che la Spagna< abbia ridotto le sue emissioni di • almeno ventici per cento • rispetto a quelle del >millenovecento novanta< (1,2s) raggiungendo così i quaranta grammi di diossido di carbonio equivalenti per megajoule • così come è raccomandato dall'A E E ••• abbiamo sette anni ••

Valutazione della resa di P5:

AIE(e) // gCO2eq/MJ(k) // 1990(k) // 2030(k) // 40(k) // 20%(e) // 7(k)

Nel caso del *task* NRI (*cluster* con numerali ridondanti), nonostante la sua complessità, ben 6 partecipanti su 10 hanno ottenuto un *success rate* pari a 100 e tutte le altre hanno ottenuto valori superiori a 50. L'identificazione del materiale informativo ridondante da parte delle partecipanti ha permesso in tutti i casi una resa corretta dei numerali e dei referenti ridondanti. Nei pochi casi in cui sono stati commessi errori, questi si sono concentrati sul numerale con ordine di grandezza superiore alle centinaia e diverso in ogni *cluster*, ecco un esempio:

Task NRI in A.1:

El **Gobierno socialista** propone instalar **102.500 megavatios** nuevos de energías renovables **en 7 años** solo en la **Comunidad Valenciana**. Solo en **esta Comunidad** el **Gobierno socialista** propone instalar **14.650 megavatios** nuevos al año **durante siete años**. Es decir, el **Gobierno socialista** propone

aumentar un **23%** la producción de energías renovables **en periodo de solo siete años** en la **Comunidad Valenciana**.

Resa P3:

il governo socialista propone ••• installare (1,2s) ehm centodue • ehm >milioni cinquecento mila megawatt< •• di • energia rinnovabile in solo sette anni nella comunità valenziana ••• e (2,1s) il governo vuole •• ehm • >installare quattordici mila seicentocinquanta mila nuovi megawatt< • nei prossimi sette anni • quindi si vuole aumentare del ventitré porcen/per cento la produzione di energia rinnovabile • in questa comunità •

Valutazione della resa:

Cluster 1: Gobierno socialista(k) // 102.500 megavatios(e) // en 7 años(k) // Comunidad Valenciana(k); *cluster* 2: esta Comunidad(k) // Gobierno socialista(k) // 14.650 megavatios(e) // durante siete años(k); *cluster* 3: Gobierno socialista(k) // 23%(k) // en un periodo de solo 7 años(k) // Comunidad Valenciana(k).

Per quanto riguarda il *task* NNR (*task* con numerali non ridondanti), in un solo caso è stato raggiunto un valore medio del *success rate* inferiore alla soglia di 50 e in nessun caso è stato raggiunto un risultato pari a 100. Tendenzialmente, i referenti semplici, gli anni e i tecnicismi sono stati resi correttamente; la maggior parte degli errori si sono concentrati sui numerali e la loro relativa unità di misura in quanto, talvolta, sono state adottate strategie di sintesi che hanno avuto esiti non positivi. A titolo illustrativo, ecco un esempio:

Task NNR in A.1:

Para el año 2030, en **España**, se prevé una **potencia total instalada** en el sector eléctrico de **157 gigavatios**. El **sector de movilidad y transporte** reducirá sus emisiones en **28 millones de toneladas de CO2 equivalente** [MtCO2-eq] **antes de 2027**, **en toda nuestra península**. **Antes de que acabe 2019**, solo en la región de **Murcia**, habrá un **ahorro de energía acumulado** en **una década** de más de **1.300 kilotoneladas equivalentes de petróleo** [ktep].

Resa di P2:

per l'anno duemila e trenta in Spagna • si prevede (2,1s) una: (1,1s) capacità totale del settore elettrico di centocinquanta • ehm megawatt nel settore della • >mobilità e dei trasporti< • che ridurrà le sue emissioni die/di ventotto milioni di tonnellate • equivalenti >prima del duemila e ventisette< • eh a tutta la nostra penisola ••• >prima della fine del duemila e diciannove< • solo nella regione di Mursia • >del duemila e ventinove< ci sarà un risparmio di energia accumulato •• eh <in un decennio> (2,9s) che equivarranno a una grande quantità di tonnellate di petrolio •

Valutazione della resa:

Cluster 1: para el año 2030(k) // España(k) // potencia total instalada(e) // 157 gigavatios(e); *cluster* 2: sector de movilidad y transporte(k) // 28 MtCO2-eq(e) // antes de 2027(k) // en toda nuestra península(k); *cluster* 3: antes de que acabe 2019(k) // Murcia(k) // ahorro de energía acumulado(k) // una década(k) // 1.300 ktep(e)

3.1.2 Task Success rate con CAI *tool* di SmarTerp

La Tabella 19 contiene una descrizione dei dati quantitativi relativi ai *success rate* ottenuti per i *task* NI, SI, EI, PI, AI, LI, TI (a bassa complessità) nella condizione sperimentale contenente i suggerimenti generati automaticamente dal CAI *tool* di SmarTerp.

	NI	SI	EI	PI	AI	LI	TI
Media	54.5	70.0	37.0	60.0	53.0	30.0	80.0
Mediana	72.5	100	12.5	100	65.0	0.00	100
Deviazione standard	49.2	48.3	43.7	51.6	50.3	48.3	42.2
Minimo	0	0	0	0	0	0	0
Massimo	100	100	100	100	100	100	100

Tabella 19 Descrizione dei task success rate raggiunti nella condizione sperimentale contenente il CAI tool di SmarTerp per i task a bassa complessità mediante i valori di statistica descrittiva.

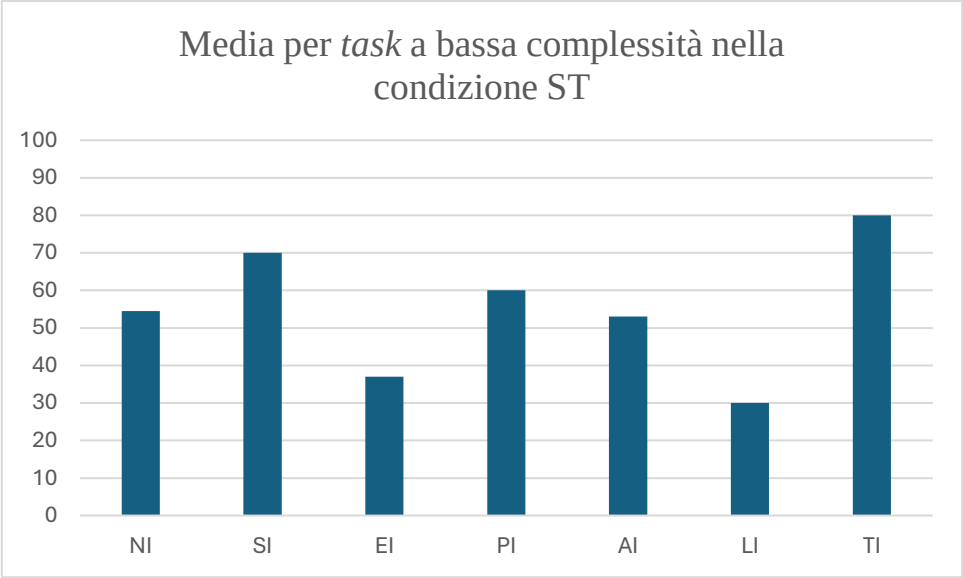


Figura 19 Grafico relativo ai valori medi raggiunti nei task success rate per i task a bassa complessità nella condizione contenente il CAI tool di SmarTerp.

Osservando i valori medi dei *success rate* ottenuti nei singoli *task* dalle dieci partecipanti, spicca la media pari a 30 (con valori minimi di 0 e massimi di 100) ottenuta nel *task* LI (sfida lessicale isolata); infatti, solo 3 partecipanti su 10 hanno reso correttamente la sfida lessicale presentata, mentre nelle restanti sette rese il lessema proposto è stato omesso o tradotto con alterazioni semantiche significative. In molti casi, la non corretta comprensione della sfida lessicale presente nel discorso ha portato a uno *spillover effect* sull'intero *cluster* contenente lo stimolo. Per esemplificare quanto accaduto in termini generali, si osservino i seguenti esempi in cui sono evidenti le distorsioni di significato:

> *Task* LI in A.1:
>
> Los expertos **achacan** a la falta de regulación, el descontrol que se ha producido en los servicios de extinción de los incendios forestales.
>
> Resa P5:
>
> gli esperti (1,3s) danno la mancanza di una normativa (1,3s) la >mancanza di controllo< nella regolazione e gestione dei/gli incendi boschivi ••
>
> Resa P4:
>
> pe/ secondo gli esperti • è una: •/l'assenza di normativa • è p/la causa d/de:gli incendi forestali (1,1s)

Un altro *task* particolarmente interessante, data la media dei *success rate* inferiore a 50, è il *task* EI (entità nominale isolata), si osservano infatti 5 rese del *cluster* con *success rate* pari a 0 e 5 con valori variabili compresi tra 25 e 100 (di cui uno inferiore a 50). Tutti i *success rate* con valore pari a 0 sono riconducibili allo stimolo inserito nel discorso A.1 per l'esecuzione del *task*: *programa DUS 5000*. In particolare, il CAI *tool* di SmarTerp non ha riconosciuto l'entità denominata che era stata inserita durante la fase di addestramento del CAI in preparazione alla sessione e *DUS* è stato trascritto dal CAI con il valore numerale foneticamente simile *dos*, portando quindi tutte le partecipanti a errori nella resa dello stimolo. A fronte di potenziali critiche circa la pronuncia del lessema, il *file* video è stato riascoltato più volte ed è da escludersi un errore di pronuncia dell'oratrice. Nel caso dello stimolo per il *task* EI presente in B.1, trattandosi di un nome proprio di un deputato del *Congreso de los Diputados*, non ci sono stati casi di omissioni totali del nome proprio, ma nella maggior parte dei casi sono state omesse alcune parti del nome, portando quindi a omissioni parziali, di seguito si propongono due esempi:

> *Task* EI in B.1:
>
> Señor **Francisco José Contreras**, no podemos ignorar los datos fruto del conocimiento científico, como usted hace con su negacionismo; al contrario, debemos integrarlas en la toma de decisión a tiempo y de manera eficaz.
>
> Resa P2 (*Success Rate*: 25):
>
> segnor Franci/Fransisco Contrega ••• non possiamo negare i dati • scientifici • come • lei fa col suo negazionismo • ma al contrario dobbiamo integrare • questi dati alla/>nelle prese di decisioni< •
>
> Resa P11 (*Success Rate*: 75):
>
> signor Francisco José On/Ortrega •• non possiamo tralasciare i dati frutti del ••• /del/ della scienza •• come vuole fare lei con il suo negazionismo • dobbiamo considerare (1,6s) >nelle nostre prese< di decisioni

Nel caso del *task* AI (acronimo isolato), si osserva una media (53) estremamente vicina alla soglia di 50, pertanto con un valore relativamente basso: 4 rese hanno ottenuto un *success rate* nella resa del *task* pari a 0, 5 rese hanno ottenuto un valore pari a 100 e una pari a 30. L'acronimo *CMNUCC* non è stato correttamente riconosciuto dal CAI *tool* di SmarTerp e non è mai stato

reso correttamente; in un solo caso è stata adottata una strategia di sintesi plausibile, di seguito l'unica resa con un valore superiore a 0, in quanto è stata adottata una strategia di sintesi:

> *Task* AI in B.1:
>
> No, señorías, las previsiones de la **CMNUCC** [Convención Marco de las Naciones Unidas sobre el Cambio Climático] ni son alarmistas ni son mentirosas. Al contrario, lamentablemente se ven confirmadas por los hechos.
>
> Resa P2 (*Success Rate*: 30):
>
> Onorevoli (2,3s) le previsioni (1,9s) fatte • da organizzazioni (1,3s) ahm addette sono vere e sono confermate dai fatti ••

Per quanto riguarda il *task* NI (numerale isolato), anch'esso presenta una media vicina al valore-soglia di 50; infatti, 4 partecipanti hanno reso correttamente il numerale, in 1 caso il numerale è stato reso correttamente ma con un lieve errore di pronuncia, in 1 caso è stato reso con un'approssimazione (con omissione parziale del 50%) e nelle restanti 4 rese il numerale non è stato reso correttamente. La prevalenza di errori dovuti al mantenimento della struttura dissimmetrica *x mil x millones* (osservata nella condizione sperimentale con la trascrizione automatica di Google Meet) è venuta meno nel *task* NI nella condizione sperimentale con il CAI di SmarTerp; infatti, gli errori prevalenti sono stati dovuti a una mancata comprensione del numerale e talvolta a errori quali il mancato ricollegamento del numerale con l'elemento a cui faceva riferimento, di seguito vengono illustrati due esempi in cui il *success rate* è stato pari a 0:

> *Task* NI in B.1:
>
> Señorías, España ha importado combustibles fósiles por un valor de **48.570 millones de euros** solo el año pasado, una cantidad demasiado elevada.
>
> Resa P3 (*Success Rate*: 0):
>
> onorevoli • la Spagna •• ha importato combustibili fossili • per un valore di quaranto (1,1s) ehm >miliardi di euro< solo l'anno passato • una quantità troppo elevata ••
>
> Resa P7 (*Success Rate*: 0):
>
> la Spagna ha importato ••• >combustiboli</combustibili fossili • per quarantotto/eh/oltre quarantotto miliardi •• <di eh tonnellate> (1,2s)

Nel caso del *task* PI (dissimmetria profonda isolata), si osserva che 6 partecipanti su 10 hanno ottenuto *success rate* pari a 100, mentre le restanti 4 hanno commesso errori o omissioni che hanno portato a *success rate* pari a 0. In linea generale, tutti i *task* PI del discorso A.1 (*son más de… que de…*) sono stati resi correttamente e tutti gli errori commessi si sono concentrati nel *task* PI del discorso B.1 (*se trata de que*), in quanto nella maggior parte dei casi la dissimmetria è stato tradotta con il verbo *dovere,* nonostante in questa dissimmetria morfosintattica sia assente una sfumatura di obbligatorietà (e quindi un tono di comando). Nonostante queste caratteristiche generali circa la tipologia degli errori, osserviamo anche casi di calchi morfosintattici con conseguente distorsione semantica, come nel caso seguente:

> *Task* PI in B.1:
>
> La electrificación es la clave para la sostenibilidad en el tiempo. **Se trata de que** el Congreso de los diputados coopere en esta transformación aprobando las medidas inevitables hacia la descarbonización.
>
> Resa P11 (*Success Rate*: 0):
>
> l'elettrificazione è chiave per la sostenibilità nel tempo (1,3s) si tratta ••• che: • il congresso dei >deputati< • cooperi in questo processo • attuando le misure inevitabili per •• raggiungere la decarb/>decarbonizzazione< ••

Nel caso del *task* SI (dissimmetria superficiale isolata), si osserva una media di 70 e in 7 rese del *task* è stato raggiunto un *success rate* pari a 100, mentre nelle 3 restanti la resa è stata non corretta (*success rate* = 0). La dissimmetria superficiale presente nel *task* SI del discorso B.1 (*les insto a que*) è stata resa correttamente in tutte le 5 interpretazioni contenenti il discorso B.1, mentre la dissimmetria del *task* SI del discorso A.1 (*se ha sometido*) è stata resa correttamente in 3 interpretazioni su 5, come mostrato di seguito:

> *Task* SI in A.1:
>
> Nuestro compromiso frente a la pobreza energética queremos reforzarlo en este año complicado. Quiero poner en valor el bono social térmico, el cual **se ha sometido** a una comisión de expertos.
>
> Resa P8 (*Success Rate*: 0):

il nostro impegno •• nell'ambito della povertà energetica • lo vogliamo • rafforzare in questo anno difficile (1,1s) vo:glio • >sottolineare l'importanza del bonus< riscaldamento ••• eh che è eh stato studiato da esperti ••

Resa P9 (*Success Rate*:0):

il nostro impegno •• in questa situazione di • povertà energetica •• deve essere rafforzato • in questo: anno complicato •• >vorrei sottolineare l'importanza< del bonus riscaldamento (1,3s)

Il *task* in cui è stato ottenuto il *success rate* con un valore medio più elevato è il *task* TI (tecnicismo isolato); infatti, 8 partecipanti su 10 hanno reso correttamente il tecnicismo presente nel *task*, talvolta con lievi riformulazioni dovute ad autocorrezioni. In soli due casi il tecnicismo è stato reso con soluzioni non corrette, ricorrendo a calchi lessicali:

Task TI in B.1:

Es fundamental, promover la **ecologización** de la industria, para que haya desarrollo sostenible y trabajo decente a la vez.

Resa P3 (*Success Rate*: 0):

è fondamentale • promuovere (1,4s) <l'ecologizzazione> • dell'industria • affinché ci sia uno sviluppo sostenibile e • >allo stesso tempo< • un lavoro decente

Resa P7 (*Success Rate*: 0):

è (1,2s) essenziale ••• promuovere <l'ecologizzazione> dell'industria affinché • vi sia sviluppo sostenibile ed un lavo/>e lavori< •• dignitosi •••

La Tabella 20 contiene una descrizione dei dati quantitativi relativi ai *success rate* ottenuti per i *task* RN, TP, SP, LS, LP (di media complessità) nella condizione sperimentale contenente il CAI *tool* di SmarTerp.

	RN	TP	SP	LS	LP
Media	58.8	76.7	80.0	65.0	60.0
Mediana	56.3	66.7	100	50.0	50.0
Deviazione standard	40.8	22.5	25.8	33.7	39.4
Minimo	0.00	33.3	50	0	0
Massimo	100	100	100	100	100

Tabella 20 Descrizione dei dati relativi ai success rate raggiunti nei task di media complessità con CAI tool di SmarTerp mediante i valori di statistica descrittiva.

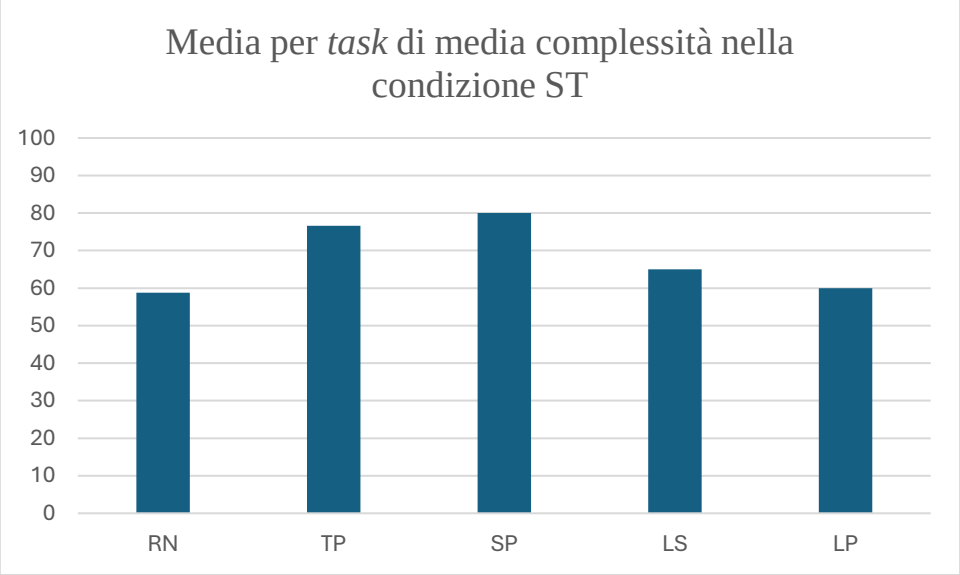


Figura 20 Grafico relativo ai valori medi raggiunti nei task success rate per i task di media complessità nella condizione contenente il CAI di SmarTerp.

Prima di procedere con l'analisi descrittiva di quanto emerso nei singoli *task*, è necessario osservare la presenza di valori medi superiori a 50 per tutti i *task* di media complessità proposti, anche se 5 *task* su 5 presentano valori minimi al di sotto di tale soglia.

Per quanto riguarda il *task* RN (referente complesso + numerale), solo in 2 rese l'intero *task* non è stato reso correttamente (quindi con *success rate* = 0), mentre in 4 rese è stato reso correttamente l'intero *task* (*success rate* = 100), in 3 rese è stato reso con valori medi uguali o superiori a 50 (ma inferiori a 70) e in 1 resa è stato ottenuto un valore medio inferiore a 50 (ma superiore a 0). Curiosamente, in molti casi, l'acronimo proposto è stato reso correttamente, mentre le difficoltà si sono concentrate sulla resa del numerale, di seguito si propongono due esempi:

Task RN in B.1:

El fondo del **MRR** [Mecanismo de Recuperación y Resiliencia] es un instrumento fundamental para fomentar reformas y efectuar inversiones en los Estados miembros de la UE por un valor total de **723 800 millones** de euros.

Resa P2 (*Success Rate*: 62,5):

il fondo • per il piano di ripresa è fondamentale •• per aumentare • e incoraggiare le: riforme per gli investimenti negli stati membri dell'U E • per un valore di (5,1s) settecento e:h ventitré miliardi di euro in totale •

Resa P11 (*Success Rate*: 50):

il fondo •• per la ripresa e la resilienza è fondamentale pe:r • promuovere riforme • e • pe:r >effettuare investimenti< negli stati membri dell'Unione Europea • per un valore totale ••• di: (1,8s) settecento ventritré mila • ottocento milioni di euro (2,1s)

Il *task* LS (sfida lessicale + dissimmetria superficiale) presenta un valore medio pari a 65: in 4 casi l'intero *task* è stato reso correttamente, in 5 casi è stato reso correttamente al 50% (quindi uno solo dei due stimoli) e in una resa la soluzione proposta risulta non corretta. Nei casi in cui solo uno dei due stimoli è stato reso correttamente, tendenzialmente è stata resa correttamente la sfida lessicale e le difficoltà si sono concentrate sulla struttura dissimmetrica, come nelle rese seguenti:

Task LS in A.1

Con toda modestia, pienso que es fundamental impulsar políticas para la cohesión territorial y social: desde la rebaja de las facturas eléctricas hasta el aumento de las **nóminas**.

Resa P4 (*Success Rate*: 50):

ehm penso che sia fondamentale (1,8s) a/avviare delle:/>delle politiche per< la coesione: te/territoriale e sociale •• fino al/>all'aumento delle buste paga< •••

Resa P5 (*Success Rate*: 50):

>secondo noi< (4,3s) <penso che sia fondamentale> • incoraggiare politiche per la coesione territoriale e sociale (1,1s) da:lla diminuzione delle bollette elettriche all'aumento dei salari •••

Per quanto riguarda il *task* LP (sfida lessicale + dissimmetria profonda), simile al *task* precedente, 4 partecipanti hanno reso correttamente l'intero *task*, 4 hanno reso correttamente solo la metà degli stimoli (*success rate* = 50) e 2 hanno ottenuto valori pari a 0 in quanto nessuno dei due stimoli è stato reso correttamente. Quando uno solo dei due stimoli è stato reso correttamente, tendenzialmente le difficoltà (e quindi i valori pari a 0) si sono concentrate sulla struttura dissimmetrica, come nell'esempio seguente:

Task LP in A.1:

Señorías, **con independencia de** nuestra orientación política, ¡qué ilusión sería ver a nuestros sobrinos, a nuestros hijos y a nuestros nietos vivir en un mundo más respirable!

Resa P4 (*Success Rate*: 50):

onorevoli ••• sarebbe davvero • bellissimo • poter vedere i nostri figli i nostri nipoti • che vivono in un mondo (1,6s) in/più respirabile

Nonostante la tendenza generale, sono presenti anche esempi di errori dovuti a una resa non corretta della sfida lessicale, come nell'esempio seguente:

Resa P5 (*Success Rate*: 50):

onorevoli • indipendentemente dalla nostra ideologia politica •• tutti noi • >desidereremmo< che • i nostri figli • i nostri nipoti • vivessero in un mondo migliore •

Nel caso del *task* TP (tecnicismi + dissimmetria profonda), si osserva un valore medio pari a 76,6: dato da 9 *success rate* per il *task* superiori alla soglia di 50 e 1 solo *success rate* inferiore al valore-soglia. Nella maggior parte dei casi, almeno uno dei due tecnicismi presenti nel *task* è stato reso correttamente; per quanto riguarda la struttura dissimmetrica, tendenzialmente è stata resa correttamente oppure omessa, ma mai interpretata in modo non corretto.

Task TP in A.1:

Sin embargo, **desde el realismo**, somos conscientes de las limitaciones de los procesos de depuración de nuestro país. Por ello estamos trabajando con científicos de gran nivel para ir modernizando las fases de **percolación** y **floculación**.

Resa P1 (*Success Rate*: 66,67):

eh della/in/in quest'ambito e quindi •• siamo coscienti del fatto • che i processi di depurazione del nostro paese sono limitati ••• quindi stiamo lavorando con scienziati di grande livello per ••• modernizzare • le fasi di percolazione e • flocculazione •

Resa P9 (*Success Rate*: 66,67):

ciò nonostante siamo • consapevoli •• dei limiti ••• <dei processi di depurazione del nostro Paese> ••• per questo •• stiamo >collaborando< co:n scienziati >prominenti< (3,1s) per/• >nell'ambito della< percolazione • e della flocculazione ••

Tra i *task* a media complessità, il *task* SP (dissimmetria superficiale + dissimmetria profonda) è quello con il valore medio più elevato, infatti in tutte le rese è stato correttamente almeno uno dei due stimoli (infatti, il valore minimo è pari a 50) e in 6 rese entrambi gli stimoli sono stati resi correttamente (con *success rate* pari a 100). Nei casi in cui uno dei due stimoli non è stato reso correttamente, gli errori si sono concentrati sulla resa della struttura dissimmetrica a dominanza di elaborazione profonda, si osservino i seguenti esempi:

Task SP in B.1:

En los últimos años las desigualdades **no han hecho sino aumentar**; ahora, queremos más igualdad de oportunidades y en el acceso a beneficios ambientales, a través de una política **todo lo arriesgada que** se quiera, pero necesaria.

Resa P2 (*Success Rate*: 50):

negli ultimi anni •• le disuguaglianze • non hanno fatto che aumentare •• ora vogliamo più uguaglianze e più opportunità nell'accesso • ai servizi ambientali mediante una politica •che sia/• che è necessaria ••

Resa P10 (*Success Rate*: 50):

negli ultimi anni • le disuguaglianze • >non hanno fatto altro che< aumentare (1,4s) ora vogliamo più uguaglianza di: opportunità (1,2s) nell'accesso ••• a:l/le/••• alle >varie politiche< ••

La Tabella 21 contiene una descrizione dei dati quantitativi relativi ai *success rate* ottenuti per i *task* NRS, NLP, UN, NRI, NNR (a elevata complessità) nella condizione sperimentale contenente il CAI *tool* di SmarTerp.

	NRS	NLP	UN	NRI	NNR
Media	46.1	68.7	61.7	91.2	76.3
Mediana	40.0	70.8	60.7	91.6	77.9
Deviazione standard	31.8	25.6	23.0	6.21	13.5
Minimo	0	16.7	28.6	83.3	46.1
Massimo	100	100	96.4	100	90.4

Tabella 21 Descrizione dei dati relativi ai success rate raggiunti nei task di alta complessità con CAI tool di SmarTerp mediante i valori di statistica descrittiva.

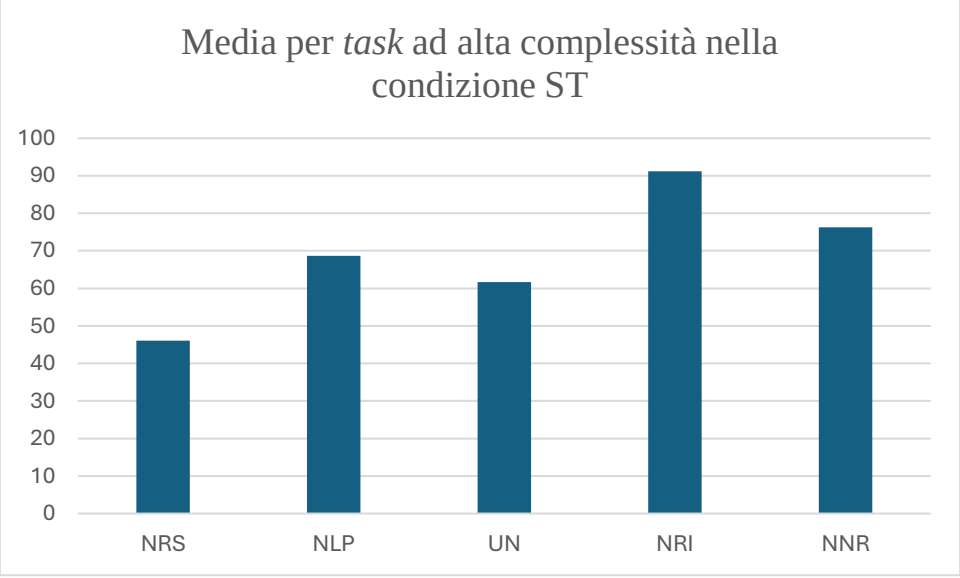


Figura 21 Grafico relativo ai valori medi raggiunti nei task success rate per i task ad alta complessità nella condizione contenente il CAI di SmarTerp.

Tra i *task* a elevata complessità nella condizione sperimentale con il CAI *tool* di SmarTerp spicca il valore medio del *task* NRS (numerali + referente complesso + dissimmetria superficiale) inferiore al valore-soglia di 50 (46,1); infatti, 6 rese su 10 hanno ottenuto un

success rate inferiore a 50 (di cui uno pari a 0) e in un solo caso tutti gli stimoli sono stati resi correttamente (con un *success rate* pari a 100). Nella maggior parte dei casi, gli errori si sono concentrati sui valori numerici (spesso oggetto di errori e di omissioni); in alcuni casi, la struttura dissimmetrica non è stata resa correttamente, seppur con scarse ripercussioni sulla correttezza semantica della resa. Di seguito, si riporta un esempio in cui è evidente lo *spillover effect* della densità informativa sulla resa dei numeri omessi (sia parzialmente che totalmente):

> *Task* NRS in B.1:
>
> Como saben, la Comisión Europea ha aprobado nuestro **plan de recuperación y resiliencia**, que ahora asciende a **163.200 millones** de euros, compuesto por **83.000 millones** en préstamos y **80.200 millones** en lo que se refiere a subvenciones para la transición energética.
>
> Resa P2:
>
> come saprete la Commissione Europea (1,1s) ha approvato il nostro piano di ripresa e resilienza (1,3s) che si acci/si aggira • circa •• a (2,3s) >cento sessantatré miliardi< di • euro con diversi prestiti •• e altrettanti euro riferi/ in riferimento a quelle che sono le sovvenzioni per la transizione energetica •
>
> Valutazione del *Success Rate*:
>
> plan de recuperación y resiliencia(k) // 163.200 millones(x50%) // 83.000 millones(o) // 80.200 millones(o) // en lo que se refiere a…(k)

I *task* NLP, UN, NRI e NNR presentano tutti valori medi superiori a 50. Tra questi, il *task* NLP presenta un valore minimo relativamente basso (17,7), ma in tutte le altre 9 rese del *task* sono stati ottenuti *success rati* uguali o superiori al 50% (in due casi tutti gli stimoli del *task* sono stati interpretati correttamente). Gli errori sono distribuiti su tutti gli stimoli presenti e, pertanto, non è generalizzabile uno *spillover effect* con maggiori ripercussioni su uno stimolo specifico, come possiamo osservare nelle rese seguenti:

> *Task* NLP in B.1:
>
> Señorías, les voy a dar un ejemplo: **justamente** a lo largo de la pasada legislatura se emitieron más de **58,5 millones** de toneladas de CO2 al año, y **no se** adoptaron medidas adecuadas **más que** en algunos sectores. Además, **debido a** la importación de combustibles fósiles, se gastaron más de **100**

millones de euros diarios. Ahora tenemos que **aferrarnos** a nuestros objetivos sostenibles.

Resa P3

onorevoli (1,5s) vi voglio dare •• un esempio •• proprio della scorsa >legislatura< • sono state emissi/e (1,1s) più di cinquantotto <miliardi di tonnellate ehm di C O 2> • all'anno e non sono state impiegate ehm•• misure adeguate • tranne in alcuni settori • inoltre a causa dell'importanza (1,3s)/>all'importazione di combustibili fossili<• sono stati sprecati più di •• cento milioni di euro • per/al giorno • ((tossisce)) ora • dobbiamo • cercare di raggiungere i nostri obiettivi sostenibili •••

Valutazione del *Success Rate*:

Cluster 1: justamente(k)//58,5 millones(e)//no se…más que…(k); cluster 2: debido a(k)//100 millones(k)//aferrarnos(e)

Resa P11

onorevoli • vi voglio dare un esempio (1,1s) eh • de:l mandato parlamentare • sono stati •• emessi (1.9s) più di •• eh cinque •• milioni di tonnellate di C O due e non sono state adottate le: misure • eh corrette ••• inoltre (2,9s) per l'importazione di combustibili • fossili • sono stati sprecati • più di cento milioni • di: euro >ogni giorno< (2,7s) dobbiamo: • >aggrapparci< ai nostri obiettivi sostenibili ••

Valutazione del *Success Rate*:

Cluster 1: justamente(o)//58,5 millones(e)//no se…más que…(e); cluster 2: debido a(k)//100 millones(k)//aferrarnos(k)

Un altro *task* con un valore minimo inferiore a 50 è il *task* UN (unità informativa con numerali). Nell'esecuzione di questo *task,* 4 partecipanti su 10 hanno ottenuto valori medi inferiori a 50 e in nessuna resa è stato ottenuto un *success rate* pari a 100. L'entità nominale e i valori numerici (in particolari quelli seguiti dalle unità di misura e dalle percentuali) hanno visto un numero frequente di errori, dovuti anche agli errori presenti nei suggerimenti del CAI *tool*. Tendenzialmente gli anni sono stati resi correttamente; di seguito, un esempio:

UN in A.1:

La Secretaría de Estado liderada por **Morán Fernández** cuenta con un presupuesto de **579 millones** de euros, cifra que supera con creces la de **2021**, la cual representa un **2,85%** más. En este contexto, España aspira a lograr los **22 megavatios** de capacidad antes de **2032**.

Resa P8:

il seg/• eh la segretaria dello Stato •• di Fernández •• investirà ••• cinquacento settantanove milioni di euro ••• ehm che supera molto • quella del duemila ven•tuno (1,2s) che presenta un • due virgola ottanta/•• >duecento ottantacinque per cento< in più (1,3s) e quindi • si vuole • aumentare la potenza prima del duemila • trentadue

Valutazione della resa:

Morán Fernández(x50%) // megavatios (e) // 2032(k) // 2021(k) // 22(e) //579 millones(x) // 2,85%(e)

Il *task* NNR (*cluster* con numerali non ridondanti) presenta un solo valore inferiore a 50 (46,15) e i valori ottenuti per questo *task* sono tendenzialmente compresi tra 63,84 e 90,4. Osservando la tendenza generale, le unità di misura complesse hanno causato molte difficoltà nella resa del testo originale, come nell'esempio seguente:

NNR in A.1:

Para el año 2030, en **España**, se prevé una **potencia total instalada** en el sector eléctrico de **157 gigavatios**. El **sector de movilidad y transporte** reducirá sus emisiones en **28 millones de toneladas de CO2 equivalente** [MtCO2-eq] **antes de 2027**, **en toda nuestra península**. **Antes de que acabe 2019**, solo en la región de **Murcia**, habrá un **ahorro de energía acumulado** en **una década** de más de **1.300 kilotoneladas equivalentes de petróleo** [ktep].

Resa P9:

in Spagna (2,2s) prevediamo che • entro il duemila trenta • la potenza totale installata nel settore elettrico • sarà di • cento cinquantasette • gigawatt • il

settore di mobilità e trasporto ridurrà • le: proprie emissioni (1,1s) di • ventotto ••• miliardi di tonnellate di C O due (2,2s) a/>entro il duemila ventisette< in tutta la Spagna (1,9s) prima • della >fine del duemila e diciannove< solo a Mursia • il risparmio energetico ••• sarà di più di (1,1s) mille • trecento • >chilotonnellate< (4,3s)

Valutazione del *Success Rate*:

Cluster 1: para el año 2030(k) // España(k) // potencia total instalada(k) // 157 gigavatios (k); *cluster* 2: sector de movilidad y transporte(k) // 28 MtCO2-eq(e) // antes de 2027(k) // en toda nuestra península(k); *cluster* 3: antes de que acabe 2019(k) // Murcia(k) // ahorro de energía acumulado(x75%) // una década(o) // 1.300 ktep(x50%)

Nel caso del *task* NRI (*cluster* con numerali ridondanti), osserviamo un valore minimo molto superiore al valore-soglia di 50 e in ben due casi è stato raggiunto un *success rate* pari a 100, nonostante la densità informativa. I valori elevati di questo *task* sono stati favoriti dall'omissione di stimoli ridondanti, che dimostrano una corretta selezione ed omissione degli stimoli ridondanti in questa condizione sperimentale. Come evidente anche nel *task* precedente, gli errori anche in questo caso si sono concentrati sui numerali e sulle unità di misura corrispondenti, come nell'esempio:

Task NRI in A.1

El **Gobierno socialista** propone instalar **102.500 megavatios** nuevos de energías renovables **en 7 años** solo en la **Comunidad Valenciana**. Solo en **esta Comunidad** el **Gobierno socialista** propone instalar **14.650 megavatios** nuevos al año **durante siete años**. Es decir, el **Gobierno socialista** propone aumentar un **23%** la producción de energías renovables **en un periodo de solo siete años** en la **Comunidad Valenciana**.

Resa P8

il governo socialista • propone • installare centododici mila • cinquecento • <megawatt nuovi> • di energie rinnovabili • in sette anni •• solo nella comunità valenziana •• solo in questa comunità•• il governo • ehm socialista • vuole installare nuovi megawatt •• ehm quindi il governo socialista propone

aumentare del ventitré per cento • la produzione di energie rinnovabili • in solo sette anni nella comunità valensiana •••

Valutazione del *Success Rate*:

Cluster 1: Gobierno socialista(k) // 102.500 megavatios(e) // en 7 años(k) // Comunidad Valenciana(k); *cluster* 2: esta Comunidad(k) // Gobierno socialista(k) // 14.650 megavatios(e) // durante siete años(k); *cluster* 3: Gobierno socialista(k) // 23%(k) // en un periodo de solo 7 años(k) // Comunidad Valenciana(k).

3.1.3 Analisi comparativa dei *success rate*

Osservando i dati ottenuti in ciascuna condizione sperimentale, è possibile procedere a una sintesi dei dati precedentemente descritti mettendo in relazione i valori medi dei *success rate* ottenuti nella condizione sperimentale con il CAI *tool* di SmarTerp (da adesso, ST) e quelli ottenuti nella condizione sperimentale contenente la trascrizione automatica di Google Meet (da adesso, GM). Per quanto riguarda i *task* a bassa complessità (NI, SI, EI, PI, AI, LI, TI), la relazione statistica tra i dati ottenuti in entrambe le condizioni sperimentali è visivamente esplicita nel grafico seguente:

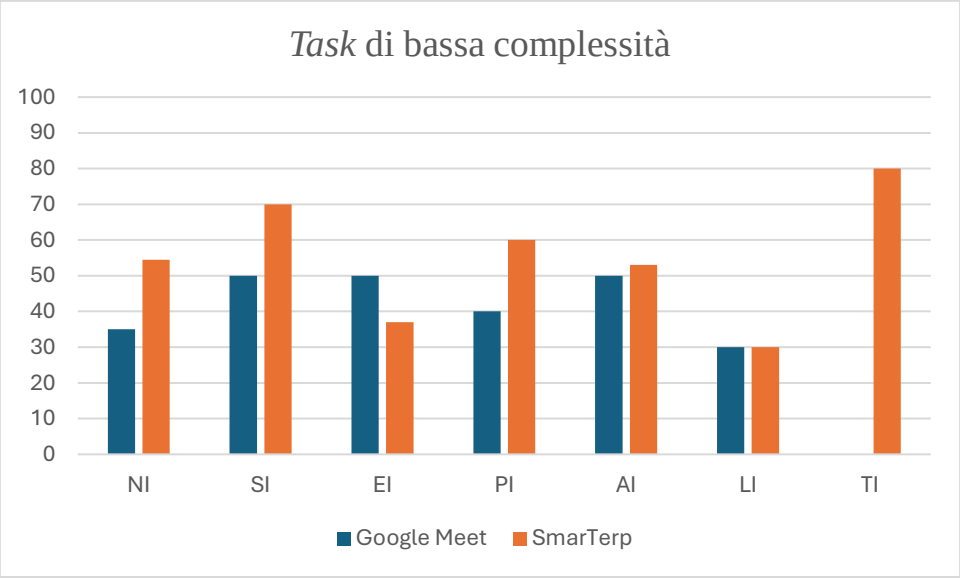


Figura 22 Success rate dei task di bassa complessità.

Osservando i sette *task* analizzati, in 5 di essi (NI, SI, PI, AI, TI) la condizione sperimentale ST ha portato a *success rate* più elevati rispetto all'altra condizione sperimentale, in 1 caso (LI) non vi è alcuna differenza e in 1 caso (EI) la condizione sperimentale GM ha portato risultati migliori. Tra tutti i *task* oggetto di quest'analisi, spicca il *task* TI (tecnicismo isolato), in quanto questo *task* non è mai stato reso correttamente nella condizione sperimentale GM, mentre nella condizione ST presenta un valore medio di 80 (abbondantemente superiore al valore-soglia di 50) e i risultati ottenuti in questo *task* ben illustrano la necessità di fornire all'interprete non solo la trascrizione dei termine specialistico (seppur funzionale a una corretta comprensione), ma anche il traducente degli stessi. Interessanti sono i risultati ottenuti nei *task* SI (dissimmetria superficiale) e PI (dissimmetria profonda); infatti, nonostante l'assenza della trascrizione di questi stimoli nel CAI *tool* di SmarTerp, sono stati resi con valori medi più elevati proprio in questa condizione sperimentale. Sulla base di questi risultati, è ipotizzabile un'elaborazione testuale più superficiale nella condizione GM, considerato il flusso informativo che viene trascritto senza alcuna interpretazione semantica dall'ASR; in entrambi i casi, la differenza dei valori medi è di 20 punti. Il *task* NI (numerale isolato) ha reiterato la stessa tendenza osservata per SI e PI, in quanto in molti casi nella condizione GM è stata determinante la presenza di calchi della struttura *x mil x millones* nella riduzione del *success rate* del *task*, mentre nella condizione ST è stata più frequente l'omissione di numerali (o la non corretta traduzione degli stessi) dovuta a una minor accuratezza dell'ASR. Il *task* EI (entità denominata isolata) mostra valori medi più elevati nella condizione GM; tuttavia, è necessario chiarire che lo stimolo per il *task* EI presente nel discorso A.1 (*programa DUS 5000*) non è mai stato reso correttamente (né in GM né in ST), a causa dell'inaccuratezza di entrambi i sistemi di ASR che in entrambi i casi hanno confuso *DUS* con il numerale *2* (*dos*, in spagnolo). Lo stimolo per il *task* EI presente nel discorso B.1 (ossia, il nome proprio *Francisco José Contreras*) ha ottenuto risultati decisamente migliori nella condizione GM in quanto è stato trascritto correttamente, mentre è stato omesso dal CAI *tool* di SmarTerp. Per il *task* LI (sfida lessicale isolata) non sono presenti differenze.

Per quanto riguarda i *task* di media complessità (RN, TP, SP, LS, LP), le differenze rilevate sono visivamente evidenziate nel grafico:

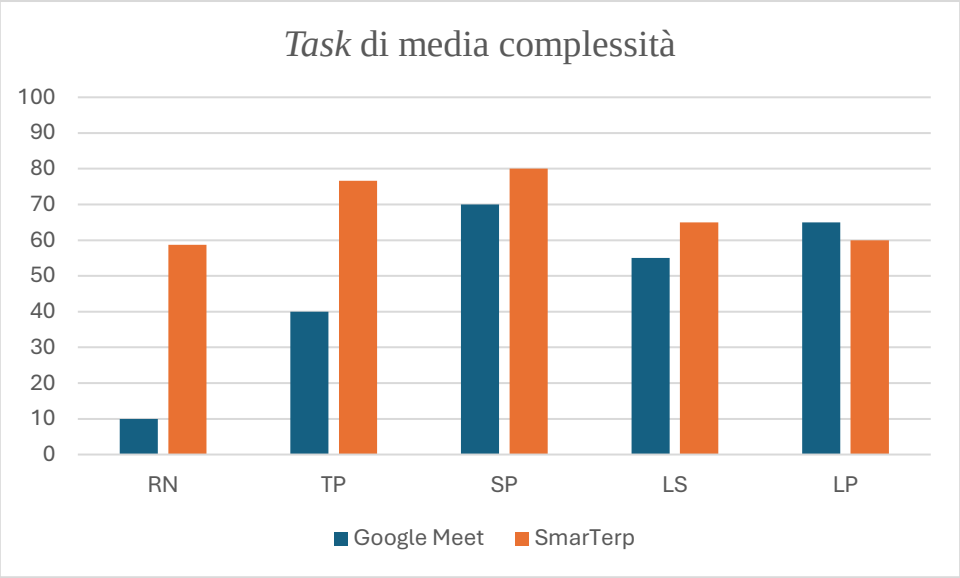


Figura 23 Success rate dei task di media complessità.

Osservando i dati raccolti, in 4 *task* su 5 (RN, TP, SP, LS) la condizione ST ha riportato dati più elevati rispetto a GM e in un solo *task* (LP) la condizione GM ha riportato valori leggermente più elevati. Osservando le differenze tra i valori medi, spicca la differenza tra GM e ST di circa 50 punti per il *task* RN (referente complesso + numerale): nella maggior parte dei casi, i *success rate* hanno avuto valori inferiori in GM per la mancata traduzione dei referenti complessi (FEDER e MRR in A.1 e B.1, rispettivamente) e per la presenza di molti calchi della struttura del numerale *x mil x millones*. Nonostante la difficoltà nella resa dei numerali in ST, il CAI *tool* di SmarTerp si è spesso rivelato un valido alleato per recuperare rapidamente i traducenti di FEDER e MRR. Un altro *task* che presenta una differenza dei valori medi interessante (di circa 40 punti) è il *task* TP (tecnicismi + dissimmetria profonda); infatti, nella condizione GM, molto spesso le difficoltà si sono concentrate sulla resa dei tecnicismi, spesso ben compresi ma con traducenti non immediatamente disponibili alle interpreti. La condizione ST ha invece permesso una rapida traduzione dei tecnicismi che ha permesso di concentrarsi maggiormente sull'input. Il *task* SP (dissimmetria superficiale + dissimmetria profonda) ha curiosamente riportato valori medi tra GM e ST con una differenza di 10 punti a favore della condizione ST; questi valori corroborano quanto osservato nei *task* a bassa complessità PI e SI; pertanto, si rafforza l'ipotesi di una maggior capacità di elaborazione testuale nella condizione ST da parte delle interpreti, dato il minor flusso testuale presente nell'input visivo. Anche per quanto riguarda il *task* LS (sfida lessicale + dissimmetria superficiale) si osserva una differenza tra GM e ST di 10 punti a favore della condizione ST. Per quanto riguarda il *task* LP (sfida lessicale + dissimmetria profonda), si osserva un valore medio di tutti i *success rate* superiore in GM (di 5 punti),

nonostante in entrambi i casi il valore sia superiore alla soglia di 50 (65 e 60, rispettivamente in GM e in ST).

Per quanto riguarda i *task* a elevata complessità (NRS, NLP, UN, NRI, NNR), i dati ottenuti in GM e ST sono visivamente espliciti nel grafico seguente:

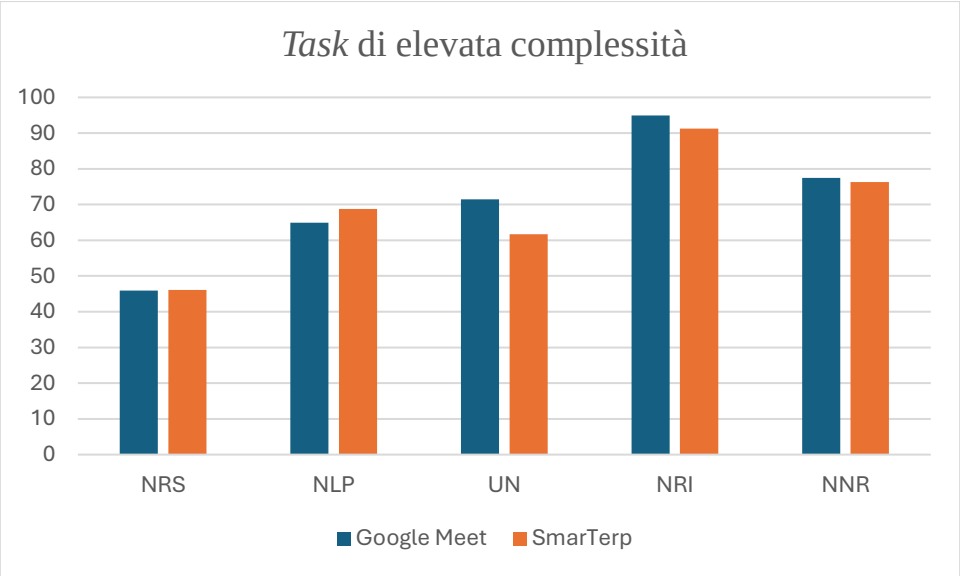


Figura 24 Success rate dei task di elevata complessità.

In termini generali, da un punto di vista meramente descrittivo, osserviamo differenze non particolarmente interessanti tra GM e ST anche se in 3 *task* su 5 (UN, NRI, NNR) la media di tutti i *success rate* è superiore nella condizione GM, mentre in soli due *task* (NRS, NLP) la media è leggermente superiore nella condizione ST. In termini generali, osserviamo che la maggiore densità informativa relativa alla presenza di numerali, acronimi e unità di misura complesse è favorita nella condizione GM; ma ancora una volta è proprio la condizione ST a presentare valori più elevati per quei *task* contenenti strutture dissimmetriche (NRS e NLP), nonostante queste siano presenti nella trascrizione automatica di Google Meet e non nel CAI *tool* di SmarTerp. Tra i *task* osservati, spicca il *task* UN (unità informativa con numerali) in cui si osserva una differenza tra i valori medi dei *success rate* di circa 10 punti a favore di GM. Il *task* ben illustra la maggiore utilità della trascrizione automatica di Google Meet per i numeri con ordini di grandezza inferiori alle migliaia e per le unità di misura complesse: entrambe le tipologie di stimoli infatti hanno spesso avuto problemi di riconoscimento da parte dell'ASR di SmarTerp e non da parte dell'ASR di Google Meet.

Terminata l'analisi comparativa tra le due condizioni sperimentali sulla base del valore medio ottenuto in ogni *task*, per ogni partecipante è stato calcolato il valore medio relativo all'accuratezza nella resa di tutti i *task* presenti all'interno del discorso, in modo da avere dati quantitativi utili per capire se ogni partecipante ha complessivamente reso meglio gli stimoli presenti in ST o in GM.

	Google Meet	SmarTerp
Media	50.0	62.9
Mediana	47.8	63.5
Deviazione standard	10.5	11.0
Minimo	34.5	44.8
Massimo	63.8	80.9

Tabella 22 Descrizione dei valori medi di tutti i success rate di tutte le partecipanti.

Come sintetizzato nella Tabella 22 i valori medi delle partecipanti risultano superiori di 12,9 punti nella condizione ST. Nella condizione GM, 5 partecipanti hanno ottenuto valori medi complessivi inferiori alla soglia di 50, mentre nella condizione ST in un solo caso il valore medio complessivo è stato inferiore a tale valore-soglia.

In particolare, questi sono i valori suddivisi per ogni partecipante:

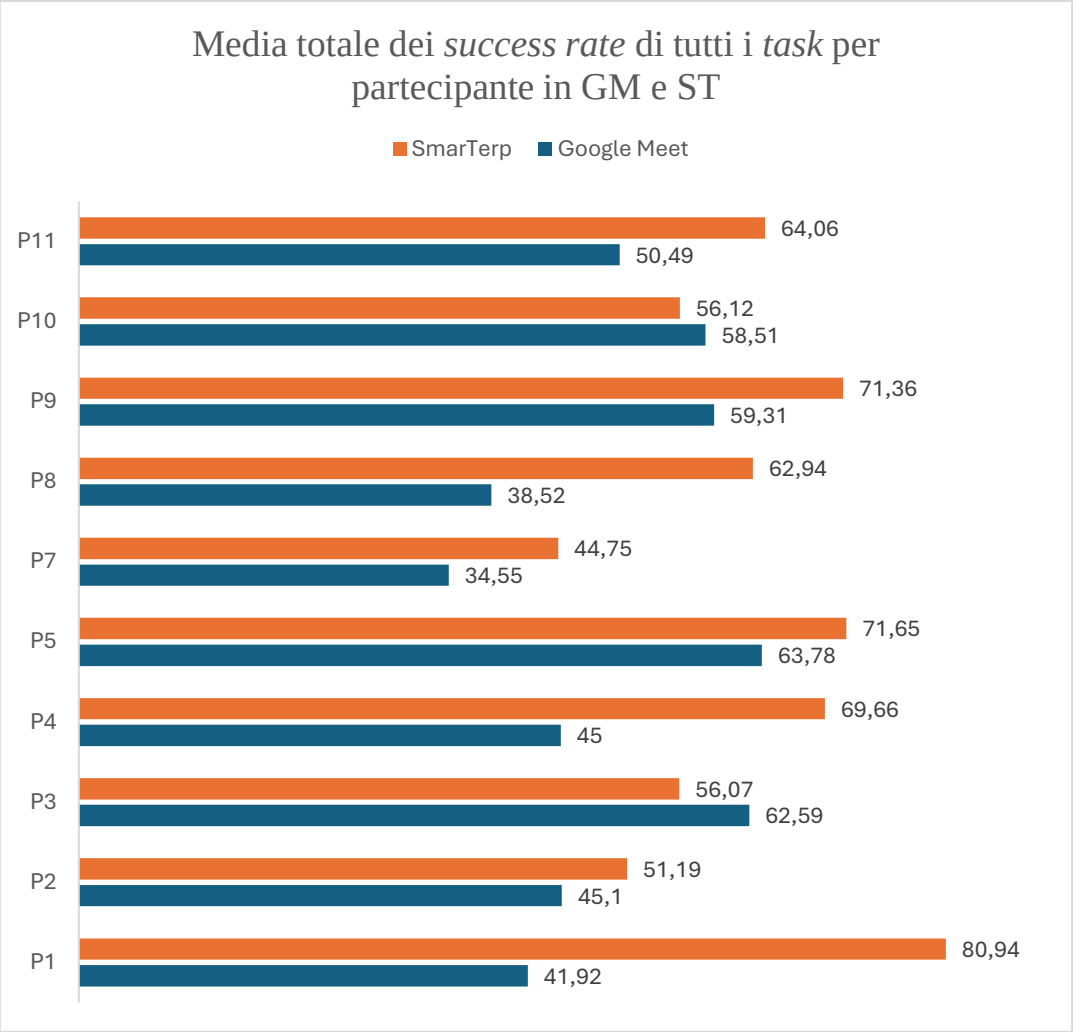


Figura 25 Descrizione comparativa relativa ai valori medi di tutti i task per partecipante.

Osservando i dati presenti in figura, è possibile affermare che solo in 2 casi (P3 e P10) sono stati raggiunti valori medi più elevati nella condizione GM, mentre negli altri 8 casi i valori medi sono risultati essere più elevati nella condizione ST. Quando i valori medi sono risultati maggiori nella condizione GM, la differenza tra i valori è risultata relativamente bassa (di circa 2 punti in P10 e di circa 6 punti in P3). Tra le otto partecipanti che hanno raggiunto valori medi superiori nella condizione ST, in 6 casi la differenza tra i valori in ST e GM è risultata essere maggiore di 10 punti e in 3 casi è risultata maggiore di 20 punti.

Come approfondito in §2.7.2, seguendo la metodologia proposta da Mellinger & Hanson (2017, p. 87 - 109), per valutare la significatività statistica delle differenze relative alle medie ottenute dallo stesso soggetto nelle due condizioni per quel che riguarda la media totale per ogni partecipante relativa ai *task success rate*, è stato condotto il *Wilcoxon signed-rank test* e abbiamo applicato la correzione di Bonferroni (Dunn, 1961) per ridurre il rischio di errori di tipo I. Abbiamo stabilito il nostro livello di significatività complessiva a 0,05; applicando la

correzione di Bonferroni, il livello di significatività di ogni test è stato dunque aggiustato a 0,0167. Il test di Wilcoxon ha rivelato una differenza statisticamente significativa tra le due condizioni (W = 4, p = 0,014). La correlazione biseriale di rango è risultata essere negativa e forte (r_b = -0,855), suggerendo una relazione inversa robusta tra le variabili esaminate.

Sulla base di questi risultati, possiamo concludere che: c'è una differenza statisticamente significativa tra GM e ST (p = 0,014); la grandezza della differenza è considerevole (d = -0,855) e GM (mediana 47,8) sembra essere inferiore rispetto a ST (mediana 62,9) per la variabile misurata. Questi risultati suggeriscono che le due condizioni non sono equivalenti e che la scelta tra la trascrizione automatica di Google Meet e il CAI *tool* di SmarTerp potrebbe avere un impatto significativo sui risultati misurati relativamente ai *task success rate*. I risultati relativi alla significatività statistica della differenza sono rilevanti, ma considerando l'esiguità del campione non possono essere generalizzabili.

3.2 Valutazione globale

Di seguito viene proposta l'analisi dei dati quantitativi relativi all'intelligibilità e all'informatività delle rese, necessari per avere dati utili a fornire informazioni sull'andamento generale della resa, oltre al *success rate* ottenuto in ogni *task*. Per procedere alla valutazione di entrambe, ogni resa è stata suddivisa in 35 *cluster* corrispondente alle *target sentence* e *control sentence* definite in fase di redazione del discorso. Prima di valutare l'informatività delle rese è stata valutata la loro intelligibilità.

3.2.1 Analisi dei dati relativi all'intelligibilità delle rese

A ogni *cluster* è stato attribuito un punteggio compreso tra 1 e 6, sulla base della sua intelligibilità. Nella valutazione dell'intelligibilità, alcuni aspetti che sono stati presi in considerazione sono: pause, disfluenze, chiarezza comunicativa, frasi aperte, legami logici, false partenze, pause piene e concordanze nel genere e nel numero. Di seguito viene proposta un'analisi dei dati raccolti suddivisa in tre sezioni: nella prima verranno approfonditi i dati raccolti per la condizione sperimentale GM, nella seconda verranno discussi i dati della

condizione ST e nella terza verrà proposta un'analisi comparativa relativa all'intelligibilità in GM e ST.

Per quanto riguarda la valutazione dell'intelligibilità nella condizione GM, di seguito viene proposta la descrizione dei valori relativi all'intelligibilità ottenuti nella resa di ogni partecipante.

	P1	P2	P3	P4	P5	P7	P8	P9	P10	P11
Media	4.60	4.54	4.37	4.00	5.00	3.66	4.20	5.14	4.80	4.40
Mediana	5	5	5	4	5	4	5	5	5	4
Deviazione standard	0.881	1.12	1.17	1.06	0.840	1.37	1.23	0.772	1.05	1.19
Minimo	2	2	2	2	3	1	2	3	2	2
Massimo	6	6	6	6	6	6	6	6	6	6

Tabella 23 Descrizione dei dati relativi all'intelligibilità nella condizione GM per partecipante, considerata la scala da 1 a 6.

Sulla base dei risultati ottenuti, 9 partecipanti su 10 hanno ottenuto un valore medio uguale o superiore a 4, di cui 5 hanno ottenuto valori medi superiori a 4,5. I valori medi sono compresi tra un minimo di 3,65 e un massimo di 5,14, con un valore medio pari a 4,47 (SD 0,452). Osservando le rese delle partecipanti, si osserva che nella condizione GM hanno particolarmente influito sulla valutazione dell'intelligibilità le incoerenze semantiche, le frasi aperte e le riformulazioni.

A titolo illustrativo, si propongono di seguito alcuni *cluster* in cui il sovraccarico informativo ha portato l'interprete a omettere parte della resa, senza l'adozione di strategie di sintesi, portando quindi a frasi aperte:

- Nel caso del *Cluster* 19 in P1 non è chiaro quali siano i due elementi che devono essere "separati", non a caso l'interprete ricorre a una pausa superiore a 1s omettendo parte dell'informazione fondamentale per poter saturare la valenza minima del verbo "separare":

  non possiamo • separare • le decisioni che hanno una ripercussione sulla nostra salute e benessere (1,6s) non possiamo far sì • che • le opzioni che vincano • siano quelle che ignorino i costi mh che hanno a che vedere con l'ambiente (3,3s)

- Nel caso del *Cluster* 22 in P7 osserviamo invece un esempio ancor più evidente di frase lasciata aperta, senza ricorrere a strategie di sintesi:

tuttavia •• mantenendo (1,5s) ehm/un/• >i piedi per terra< • dobbiamo essere coscienti del/•dei limiti che hanno i processi di depurazione nel nostro paese • dobbiamo pertanto • cercare di <modernizzare> (1,4s) alcuni • processi • che ci perme•ttano di

- Lo stesso *Cluster* precedente ha portato anche P11 a lasciare l'intera frase aperta (come dimostrato dalla pausa di 5,6s), portando a una difficile comprensione dell'intero *cluster*:

tuttavia •• essendo realisti siamo anche consapevoli delle limitazioni dei processi di depurazione del nostro paese •• per questa ragione • <stiamo lavorando> • co:n • scienziati di alto livello • per modernizzare (5,6s)

In altri casi le riformulazioni eccessive hanno causato problemi di comprensione (lato ascoltatore) della resa:

- Nel *Cluster* 30 in P1 osserviamo una quantità eccessiva di riformulazioni, nonostante venga proposto alla fine di esse un numero intelligibile in lingua italiana. Lo stesso numero proposto è seguito da una subordinata che presenta un verbo che non concorda in numero con la principale.

come sapete la commissione europea • ha approvato il nostro: • piano di ripresa e resilienza • che oggi arriva a **(1,2s) cento (3,1s) >che è composto da< ottantatré milioni in prestiti e ottanta mila (1,9s)/**ottanta miliardi e duecento milioni ••• che ha a che vedere con le sovvenzioni per la transizione energetica ••

- Nel *Cluster* 20 in P2 osserviamo un elevato numero di riformulazioni, dovute anche al calco lessicale "descontrollo". A causa della perdita di coerenza morfosintattica, si osservano errori relativi alle concordanze in genere e numero (il participio del verbo "produrre*"* dovrebbe infatti essere al femminile e la preposizione articolata prima di "servizi di estinzione" dovrebbe essere al plurale). Complessivamente, il *cluster* presenta una perdita significativa del legame logico tra la frase principale e quella subordinata.

gli esperti (1s) sottolineano la mancanza di un regolamento • **il de/descontrollo/il/la mancanza di controllo prodotto nel >servizi di estinzione< degli incendi boschivi** è • evidente ••

- Lo stesso *cluster* ha un'intelligibilità molto bassa anche nel caso di P3. Anche in questo caso l'abbondanza di riformulazioni e pause piene confonde la comprensione portando a una perdita dei legami logici tra principale e subordinata. A questo si aggiunge un errore di concordanza in genere e numero ("tutti una serie"), probabile sintomo di sovraccarico cognitivo.

gli esperti (2,1s) mh ehm dicono che spe/per mancanza di eh ehm regola/mh >di regolamenti< e mh di mancanza di controllo • nei servizi si producono tutti una serie di incendi forestali •••

- Nel *cluster* 7 in P8 osserviamo una riformulazione iniziale che confonde sull'avverbio utilizzato in apertura, ma l'intelligibilità viene ancora meno dopo la pausa piena (mh), in quanto non è chiaro se sia la scienza ad avvertirci circa le cause e le conseguenze dei cambiamenti climatici o coloro i quali non ascoltano la scienza.

n/ sfortunatamente/f/•• >nonostante< fortunatamente siano >una minoranza< • nel nostro Paese • c'è • chi no/• >chi non ascolta< la scienza **e •• mh • coloro che ci avvertono** delle cause e delle conseguenze del cambiamento climatico •••

Proseguendo l'analisi dei *cluster* con intelligibilità relativamente bassa, si osservano casi di incoerenze semantiche, ossia di errori dovuti alla non plausibilità dell'affermazione, dato il contesto, eccone alcuni esempi:

- Nella resa del *Cluster* 28, P7 presenta una misura di aumento delle bobine tra le politiche per il benessere sociale: un'affermazione di limitata plausibilità.

>modestamente< (1,5s) sostengo che sia • fondamentale • implementare politiche per il benessere sociale (1,1s) come •• quello dell'aumen/>come la misura dell'aumento delle< • bobine (2,3s)

- Nella resa del *Cluster* 15 in P10 si osserva un'incoerenza semantica simile, in cui è estremamente bassa la plausibilità dell'affermazione in grassetto, dato il contesto.

inoltre (1,1s) la preservazione della biodiversità • >marina< ha un budget di quasi mezzo milione di euro **••• il mare ci deve preoccupare (1,4s) soprattutto •• >perché è uno dei principali produttori di C O due**< (4,4s)

Per quanto riguarda la valutazione dell'intelligibilità nella condizione ST, di seguito viene proposta la descrizione dei valori relativi all'intelligibilità ottenuti nella resa di ogni partecipante.

	P1	P2	P3	P4	P5	P7	P8	P9	P10	P11
N	35	35	35	35	35	35	35	35	35	35
Mancanti	0	0	0	0	0	0	0	0	0	0
Media	4.54	4.80	4.51	4.23	5.03	4.71	4.54	5.40	4.91	5.00
Mediana	5	5	5	4	5	5	5	6	5	5
Deviazione standard	0.950	0.868	0.853	1.03	1.04	1.13	1.15	0.736	1.04	0.874
Minimo	3	3	2	2	2	1	1	3	2	2
Massimo	6	6	6	6	6	6	6	6	6	6

Tabella 24 Descrizione dei dati relativi all'intelligibilità nella condizione ST, considerata la scala da 1 a 6.

Sulla base dei risultati ottenuti, 10 partecipanti su 10 hanno ottenuto un valore medio uguale o superiore a 4 e 9 partecipanti hanno ottenuto valori medi superiori a 4,5, di cui 3 con valori medi uguali o superiori a 5. I valori medi sono compresi tra un minimo di 4,22 e un massimo di 5,40, con un valore medio pari a 4,76 (SD 0,339). Osservando le rese delle partecipanti, si osserva che nella condizione GM hanno particolarmente influito sulla valutazione dell'intelligibilità la mancanza di chiarezza comunicativa nelle parti più informative, le riformulazioni e la mancata comprensione di entità denominate, unità di misura complesse e numerali non riconosciuti dal CAI *tool*.

A scopo esemplificativo della tendenza generale, si propongono di seguito alcuni *cluster* in cui il sovraccarico informativo ha portato l'interprete a ricorrere a soluzioni traduttive non chiare, portando quindi a una mancanza di legami semantici chiari all'interno del *cluster*:

- Nel *Cluster* 9 in P2, la parola "invitati" confonde inutilmente una resa che in sé è tuttavia intelligibile:

> ed è sorprendente • come • >non ci sia un ascolto nei confronti dei cittadini< •• che siano giovani o che siano • le•/>la società civile organizzata< gli invitati o al/in generale la comunità internazionale (2,7s)

- Lo stesso *Cluster* ha portato anche la resa di P3 a una bassa intelligibilità, dovuta in questo caso non solo alle riformulazioni, ma anche all'attributo legato ai "giovani": "protestanti". In questo caso, la volontà di specificare l'orientamento religioso dei giovani risulta tutt'altro che plausibile, in quanto il discorso originale fa evidentemente riferimento ai giovani che protestano.

> ed è sorprendente >che non ascoltino nemmeno i cittadini< ••• nemmeno •• >sia la società/mh/ che non ascolti sia la so/società civile organizzata< • come i sindacati • >il consenso internazionale< o i giovani protestanti (2,1s)

- Il *Cluster* 12 ha generato numerose difficoltà (sia in ST che in GM) data la mancata identificazione del referente "Programa DUS 5000" da parte sia delle partecipanti che dei sistemi di ASR. Di seguito, viene proposta la resa di P4 in cui è evidente la mancanza di chiarezza nella resa di un *cluster* in cui il referente rimane sconosciuto da parte dell'interprete. La non chiarezza emerge, ad esempio, dalla carenza di una sintassi chiara ed esplicita.

> infatti voglio sottolineare • il successo • che: durante il periodo di apertura • e:hm cominciata: dieci giorni fa eh:m ehm (2,3) e che a:h • continua ad essere aperto il programma (1,2s)/>il programma< cinquemila (1,7s)

- Il *Cluster* 9 in P10 risulta poco plausibile il soggetto del verbo "ascoltare", dato il contesto in cui viene pronunciato il discorso, ossia il *Congreso de los Diputados*.

> ed è assurdo che i cittadini non ascoltino ••• sia • i giovani per le strade o anche la società civile • organizzata •• ma anche i sindacati e il consenso internazionale (4,5s)

In alcuni casi, il sovraccarico informativo ha portato le interpreti a riformulare, giungendo talvolta a soluzioni contenenti molte false partenze e riformulazioni che hanno inevitabilmente influito negativamente sull'intelligibilità della resa.

- Come evidente nel *Cluster* 5 in P2, le riformulazioni sono state talvolta causate da una vicinanza eccessiva rispetto al testo originale, che ha portato non solo alla presenza di

parole spagnole all'interno della resa ma anche alla presenza di strutture lessicali calcate dalla lingua spagnola (come in grassetto).

> vogliamo più fraternità e so/solidarietà •• dobbiamo biso• /avere una rivoluzione non violenta ma che possa gettare le basi per la Spagna del **siglo/>del secolo< ventuno •** ciò di cui abbiamo bisogno (1,6s)

- Come si può osservare nel *Cluster* 18 in P9, a volte le riformulazioni sono state dovute alla necessità di spiegare termini il cui traducente non era immediatamente disponibile; in grassetto è possibile osservare l'abbondante output testuale corrispondente a "bassa qualificazione energetica":

> per quanto riguarda •• >invece< l'efficienza energetica (1,2s) più della metà ••• delle case • in Spagna ••• ricevono **••pochissima/•>hanno una</•sono in una situazione •non esattamente ideale •• >per quanto riguarda la< • qualificazione energetica**
> •

- Nel *Cluster* 21 in P2 le riformulazioni della seconda parte della resa del *cluster* sono state causate dalla mancanza di comprensione della particella pronominale "nos", che è stata confusa con l'avverbio di negazione "no", portando a un crollo della coerenza semantica del *cluster*.

> per questo dobbiamo • sfi/dobbiamo cambiare il nostro modello produttivo • ma anche il modello energetico • ne vogliamo uno • che sia/>che cre</che creii posti di lavoro • che riduca le disuguaglianze • che non permetta di vivere in città (2,6s) più/che non siano più inabitabili/che siano/siano più abitabili (2,4s)

In altri casi, entità denominate, unità di misura complesse e numerali sono stati non compresi e resi in modo non corretto, ossia senza l'adozione di strategie che permettessero una chiarezza comunicativa nella resa.

- Nel *Cluster* 32 in P11 la chiarezza comunicativa è venuta meno a causa della presenza di un acronimo semanticamente neutro in lingua italiana e a causa della mancanza dell'elemento a cui fa riferimento il numerale "quaranta".

nel duemila e trenta (1,3s) >vogliamo che la Spagna< abbia ridotto le sue: • emissioni almeno del venti per cento (1,1s) in confronto a:l • mille novecento (1,1s) raggiungendo (1s) i **quaranta (3,6s)** così come raccomandato da **I A E** •• >abbiamo sette anni< ••

- Nel caso del *Cluster* 4 in P8 la difficoltà nella resa del numerale ha portato a una resa in cui il numerale vede la ripetizione dell'ordine di grandezza delle migliaia con tuttavia cifre diverse.

la prima sfida ••• della transizione energetica (1,1s) ha un ••• progetto di bilancio di • sette miliardi ••• e:h e >**tremila novecento settantasette< mila euro** •••

- Nel caso del *Cluster* 34 in P7, il sovraccarico informativo ha portato a una perdita del filo semantico del discorso, dovuta anche alla vuotezza semantica degli elementi in grassetto.

prima de:l (1,1s) duemila ventotto (1,2s) vogliamo che il venti per cento di **queste risorse** • possano essere in Spagna ••• <in quanto alla **legge dell' H due**> (2,5s) in Andalusia •• c'è stata una (1,5s) <riduzione di **ventuno miliardi di emissioni**> ((ride)) (1,5s) >va beh< nel mese di agosto sono state sviluppate nuove tecnologie •• per produrre energia >pulita< (3,8s) usando meno iridio • di • prima •

Per avere un termine di paragone utile per mettere in relazione i dati ottenuti nella condizione ST e GM per quanto riguarda l'intelligibilità ottenuta nelle due rispettive condizioni sperimentali, è stata calcolata la media di tutti i valori medi ottenuti dalle singole partecipanti, di seguito sintetizzati nel grafico:

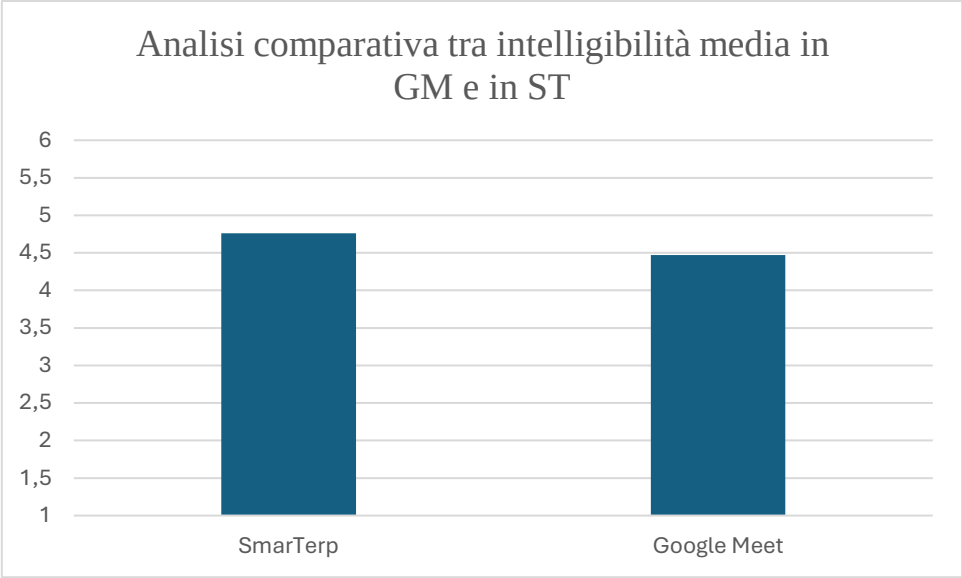


Figura 26 Intelligibilità media in ST e GM.

Come si può osservare in Figura 26, la media relativa a tutte le prestazioni delle interpreti è più elevata di 0,29 nella condizione ST, questo ci permette di poter dedurre che il minor input visivo-testuale della condizione ST favorisce un maggior distanziamento dall'input consentendo all'interprete di concentrarsi maggiormente sul proprio output.

Per ogni partecipante, i valori medi in ST e in GM sono i seguenti:

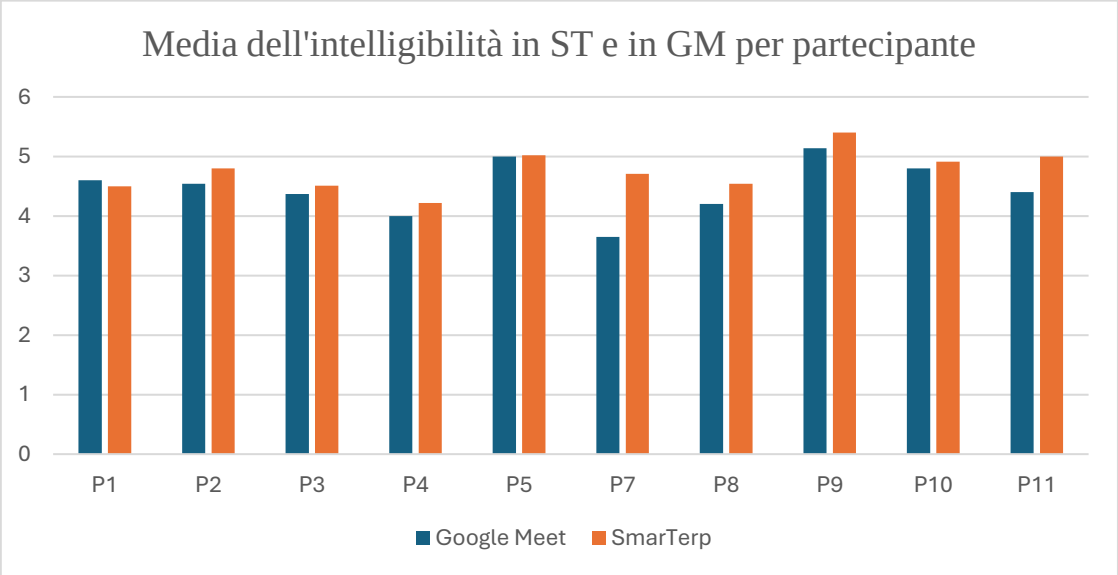


Figura 27 Analisi comparativa tra i valori medi dell'intelligibilità in GM e ST per partecipante.

Osservando il grafico, una sola partecipante su 9 (P1) ha ottenuto un valore medio relativo all'intelligibilità dei singoli *cluster* presenti in GM superiore rispetto al valore medio raggiunto in ST (con una differenza di 0,10). Le restanti 9 partecipanti hanno raggiunto sempre valori

medi relativi all'intelligibilità superiori in ST rispetto alla condizione GM (con differenze variabili da un minimo di 0,02 a un massimo di 1,06).

Per valutare la significatività statistica delle differenze relative alle medie ottenute nelle due condizioni sperimentali dallo stesso soggetto per quel che riguarda l'intelligibilità, è stato nuovamente condotto il *Wilcoxon signed-rank test* (come approfondito in §3.1.3). Il test di Wilcoxon ha rivelato una differenza statisticamente significativa tra le due condizioni ($W = 2$, $p = 0,006$), con alfa fissato a 0,0167 in seguito alla correzione di Bonferroni (§2.7.2). La correlazione biseriale di rango è risultata essere negativa e forte ($r_b = -0,927$), suggerendo una relazione inversa robusta tra le variabili esaminate.

Sulla base di questi risultati, possiamo concludere che: c'è una differenza statisticamente significativa tra GM e ST relativamente all'intelligibilità delle rese ($p = 0,006$); la grandezza della differenza è considerevole ($d = -0,927$) e l'intelligibilità sembra essere inferiore nella condizione GM (mediana 4,47) rispetto alla condizione ST (mediana 4,76), dato il valore negativo della dimensione dell'effetto. Inoltre, anche in questo caso, dato che abbiamo eseguito test multipli sullo stesso test di dati, abbiamo applicato la correzione di Bonferroni (Dunn, 1961) per ridurre il rischio di errori di tipo I. Abbiamo stabilito il nostro livello di significatività complessiva a 0,05; applicando la correzione di Bonferroni, il livello di alfa (significatività) di ogni test è stato dunque aggiustato a 0,0167. Considerando tale valore, si può affermare che la correzione di Bonferroni rafforza la significatività statistica emersa nel test di Wilcoxon. Questi risultati suggeriscono che le due condizioni non sono equivalenti e che la scelta tra GM e ST potrebbe avere un impatto significativo sui risultati misurati relativamente all'intelligibilità delle rese. Nonostante i risultati, è necessario chiarire che questi sono difficilmente generalizzabili, a causa del campione relativamente ridotto.

3.2.2 Analisi dei dati relativi all'informatività delle rese

Per ogni *cluster* è stato attribuito un punteggio compreso tra 0 e 6, sulla base della sua informatività. Nella valutazione dell'informatività, gli aspetti che sono stati presi in considerazioni sono: le omissioni, gli errori semantici, i calchi morfosintattici e le distorsioni di significato. Di seguito viene proposta un'analisi dei dati raccolti suddivisa in tre sezioni: nella prima verranno approfonditi i dati raccolti per la condizione sperimentale GM, nella seconda verranno discussi i dati della condizione ST e nella terza verrà proposta un'analisi

comparativa relativa all'informatività in GM e ST. Prima di procedere all'analisi dei dati, occorre ricordare che il valore relativo all'informatività è tanto maggiore quanta maggiore sarà la non corrispondenza dell'output dell'interprete rispetto al *cluster* del testo originale; pertanto, maggiore sarà il valore numerico e minore sarà la sua informatività.

Per quanto riguarda la valutazione dell'informatività nella condizione GM, di seguito viene proposta la descrizione dei valori relativi all'informatività ottenuta nella resa di ogni partecipante.

	P1	P2	P3	P4	P5	P7	P8	P9	P10	P11
Media	2.43	2.66	2.57	2.91	2.03	4.69	3.26	2.14	2.31	2.57
Mediana	2	3	2	3	2	5	3	2	2	2
Deviazione standard	1.24	1.19	1.12	1.36	0.785	1.13	0.950	1.19	1.55	1.27
Minimo	1	1	1	1	1	2	1	1	1	1
Massimo	6	5	6	5	4	6	5	5	6	5

Figura 28 Descrizione dei dati relativi all'informatività nella condizione GM per partecipante, considerata la scala da 0 a 6.

Sulla base dei risultati ottenuti, 8 partecipanti su 10 hanno ottenuto un valore numerico medio relativo all'informatività inferiore a 3, di cui 4 hanno ottenuto un valore numerico inferiore a 2,5. I valori numerici medi sono compresi tra un minimo di 2,03 e un massimo di 4,69 (in una scala da 0 a 6), con un valore medio pari a 2,76 (SD 0,768). Osservando le rese delle partecipanti, si osserva che nella condizione GM hanno particolarmente influito sulla valutazione dell'informatività: le omissioni, le strategie di sintesi che hanno portato a controsensi o frasi aperte, i calchi lessicali privi di significato in italiano, i calchi della struttura "x mil x millones" utilizzata in lingua spagnola per l'ordine di grandezza dei miliardi, le sostituzioni lessicali con controsensi, le omissioni dei legami morfosintattici e i calchi delle strutture dissimmetriche presenti nei discorsi.

A scopo illustrativo, di seguito osserviamo un esempio in cui la partecipante ha adottato una strategia di sintesi che porta a una vaghezza semantica in cui emerge con chiarezza una distorsione dell'intero significato del *cluster*:

> *Cluster* 9 in B.1:
>
> Es sorprendente que no escuchen tampoco a los ciudadanos, ya sea a los jóvenes en las calles o a la sociedad civil organizada, a los sindicatos o al consenso internacional. [pause]

Resa in P1:

ed è sorprendente •• che (1,3s) i cittadini (3,6s)/che/che i cittadini continuino a:•/a:hm•••/ad andare contro questa tendenza •

La tendenza alla semplificazione linguistica riscontrata a fronte di *cluster* con difficoltà interne ha talvolta portato a soluzioni interpretative in cui si osserva l'incapacità di ricostruire un senso e un messaggio non reso chiaramente, ma con omissioni e lievi distorsioni:

Cluster 20 in A.1

Los expertos achacan a la falta de regulación, el descontrol que se ha producido en los servicios de extinción de los incendios forestales.

Resa in P3

gli esperti (2,1s) mh ehm dicono che spe/per mancanza di eh ehm regola/mh >di regolamenti< e mh di mancanza di controllo • nei servizi si producono tutti una serie di incendi forestali •••

I calchi dei referenti complessi (in particolare, acronimi e nomi propri di associazioni) hanno portato spesso a errori dovuti a una non reperibilità immediata del traducente dell'acronimo nella memoria a lungo termine, portando tuttavia a un messaggio non informativo:

Cluster 12 in B.1:

No, señorías, las previsiones de la CMNUCC [Convención Marco de las Naciones Unidas sobre el Cambio Climático] ni son alarmistas ni son mentirosas. Al contrario, lamentablemente se ven confirmadas por los hechos.

Resa in P1:

no onorevoli non si tratta (1,7s)/ le/le previsioni della • C M N N non sono • allarmiste e neanche • bugiarde purtroppo sono confermate dai fatti (1,1s)

In alcuni casi, le sostituzioni lessicali messe in campo dall'interprete si sono rivelate semanticamente contrarie rispetto al significato del termine utilizzato nel discorso originale, portando a una distorsione dell'intero messaggio del *cluster*:

Cluster 15 in A.1:

> Además, el capítulo para la preservación de biodiversidad marina tiene un presupuesto de casi medio millón de euros. El mar debe preocuparnos, principalmente porque es uno de los principales **sumideros** de CO2. [pause]
>
> Resa in P10:
>
> inoltre (1,1s) la preservazione della biodiversità • >marina< ha un budget di quasi mezzo milione di euro ••• il mare ci deve preoccupare (1,4s) soprattutto •• >perché è uno dei principali **produttori** di C O due< (4,4s)

La stessa tendenza è a volte emersa nella resa di sfide lessicali:

> *Cluster* 20 in A.1:
>
> Los expertos **achacan** a la falta de regulación, el descontrol que se ha producido en los servicios de extinción de los incendios forestales.
>
> Resa in P10
>
> gli esperti • **sottolineano** la mancanza di regolamentazione (1,5s) che eh >è stata prodotta nel</nei servizi di estinzione degli incendi forestali •••

In altri casi, gli errori semantici dovuti a specifiche sfide lessicali sono stati preceduti da una resa sintetica e semanticamente vaga; ed è pertanto ipotizzabile uno *spillover effect* dovuto a un input testuale percepito come difficile (e quantitativamente abbondante da dover essere elaborato) sulla sfida lessicale alla fine di esso:

> *Cluster 28* in A.1:
>
> Con toda modestia, pienso que es fundamental impulsar políticas para la cohesión territorial y social: desde la rebaja de las facturas eléctricas hasta el aumento de las **nóminas**.
>
> Resa in P7:
>
> >modestamente< (1,5s) sostengo che sia • fondamentale • implementare politiche per il benessere sociale (1,1s) come •• quello dell'aumen/>come la misura dell'aumento delle< • **bobine** (2,3s)

In alcuni casi, sono state commesse delle omissioni di alcune frasi, senza l'adozione di strategie di sintesi delle stesse:

> *Cluster* 25 in B.1:

> Ejemplo de que las propuestas europeas trabajan en nuestra misma dirección. Dejar atrás los combustibles fósiles y sustituirlos por energías renovables genera empleo cualificado y fortalece nuestro tejido empresarial e industrial. *[pausa]*
>
> Resa in P4:
>
> e questo è un esempio del fatto che le proposte europee• lavorano • in linea con le nostre proposte (12,4s)

Tra questi casi, si osservano anche alcune rese in cui sono state omesse alcune parti all'interno della frase stessa senza l'adozione di strategie di sintesi, portando quindi alla presenza di frasi aperte:

> *Cluster* 22 in A.1:
>
> Sin embargo, desde el realismo, somos conscientes de las limitaciones de los procesos de depuración de nuestro país. Por ello estamos trabajando con científicos de gran nivel para ir modernizando las fases de percolación y floculación.
>
> Resa in P11:
>
> tuttavia •• essendo realisti siamo anche consapevoli delle limitazioni dei processi di depurazione del nostro paese •• per questa ragione • <stiamo lavorando> • co:n • scienziati di alto livello • per modernizzare (5,6s)

In alcuni casi, le strategie di sintesi hanno portato l'interprete a produrre un output testuale totalmente semanticamente nuovo rispetto al discorso originale; pertanto, informativamente distante rispetto al discorso originale:

> *Cluster* 16 in A.1:
>
> El mar ejerce de termostato fundamental y las cuestiones relacionadas con el mar son más de voluntad política que de carencia de recursos.
>
> Resa in P7:
>
> realizz/ <i dati ricevuti> (2,8s) sono •• non sono affatto • incoraggianti

Per quanto riguarda la valutazione dell'informatività nella condizione ST, di seguito viene proposta la descrizione dei valori relativi all'informatività ottenuti nella resa di ogni partecipante.

	P1	P2	P3	P4	P5	P7	P8	P9	P10	P11
Media	2.91	2.80	2.49	2.40	2.17	3.94	2.97	2.14	2.26	2.34
Mediana	3	3	2	2	2	4	3	2	2	2
Deviazione standard	1.12	1.32	1.25	1.19	1.18	1.43	1.20	1.12	1.17	1.24
Minimo	1	1	1	1	1	1	1	1	1	1
Massimo	5	6	5	6	6	6	6	6	5	5

Figura 29 Descrizione dei dati relativi all'informatività nella condizione ST per partecipante, considerata la scala da 0 a 6.

Sulla base dei risultati ottenuti, 9 partecipanti su 10 hanno ottenuto un valore numerico medio relativo all'informatività inferiore a 3, di cui 6 hanno ottenuto un valore numerico inferiore a 2,5 (in una scala da 0 a 6). I valori numerici medi sono compresi tra un minimo di 2,14 e un massimo di 3,94, con un valore medio pari a 2,64 (SD 0,546). Osservando le rese delle partecipanti, si osserva che nella condizione ST hanno particolarmente influito sulla valutazione dell'informatività: gli errori semantici a livello lessicale (dovuti spesso a errori di comprensione, anche relativamente ai numerali), le omissioni senza l'adozione di strategie di sintesi plausibili (oppure con l'omissione delle parti più informative), le omissioni di unità di misura (lasciando quindi il valore numerale spoglio del proprio referente) e i controsensi.

In alcuni casi, la mancata comprensione di referenti relativamente semplici ha portato a errori semantici che hanno avuto ripercussioni sull'intera semantica del *cluster*, portando a errori gravi. Si noti in questo esempio come la mancata resa semantica di "democracia" abbia di fatto portato l'interprete a stravolgere l'intero significato della seconda parte del *cluster*:

> *Cluster* 27 in A.1
>
> Otro reto al que hacía referencia al inicio de mi intervención, no menos importante, es el denominado reto demográfico. Creemos que se trata de una cuestión de calidad de la **democracia**. [pause]
>
> Resa in P1:
>
> inoltre (1,8s) come facevo/come dicevo prima/ all'inizio ••• si parla della sfida demografica •• è una questione di qualità della demografi/del/della **demografia** (4,4s)

In alcune rese, quando l'informatività del *cluster* aumenta, si osserva l'adozione di strategie di sintesi non plausibili e con un'alta frequenza di errori semantici e di omissioni. In questo caso, oltre agli errori relativi alla resa dei numerali, si osserva una pausa di 8,5s con cui l'interprete lascia la frase aperta, senza ricorrere a strategie di sintesi:

> *Cluster 32* in A.1
>
> El Gobierno socialista propone instalar 102.500 megavatios nuevos de energías renovables en 7 años solo en la Comunidad Valenciana. Solo en esta Comunidad el Gobierno socialista propone instalar 14.650 megavatios nuevos al año durante siete años. Es decir, el Gobierno socialista propone aumentar un 23% la producción de energías renovables en solo siete años en la Comunidad Valenciana.
>
> Resa in P1:
>
> il governo socialista • propone • di • installare (3,1s) cento(2s)/**cento ventidue mila (8,5s)** il/il governo socialista vuole proporre di aumentare del ventitré per cento la produzione di energie rinnovabile in solo sette anni nella comunità valensiana (1,4s)

Come menzionato precedentemente, in alcuni casi l'omissione delle unità di misura ha influito particolarmente sull'informatività dell'intero *cluster*, si noti come in questo caso la resa semanticamente corretta del valore numerale sia di fatto totalmente annullata e semanticamente vuota data la mancanza della propria unità di misura; inoltre, è notevole l'impatto del sovraccarico informativo sulla resa dell'acronimo:

> *Cluster 32* in B.1:
>
> En 2030 proponemos que España haya reducido sus emisiones al menos un 20 % con respecto a las de 1990, alcanzando los 40 **gramos de dióxido de carbono equivalentes por megajulio** [gCO2eq/MJ], tal y como nos recomienda la AIE [Agencia Internacional de la Energía]. Tenemos 7 años.
>
> Resa in P11:
>
> nel duemila e trenta (1,3s) >vogliamo che la Spagna< abbia ridotto le sue: • emissioni almeno del venti per cento (1,1s) in confronto a:l • mille novecento

(1,1s) raggiungendo (1s) i quaranta (3,6s) così come raccomandato da I A E •• >abbiamo sette anni< ••

In alcuni casi, l'adozione di strategie di sintesi ha comportato l'omissione di parti significative del cluster originale e ha portato l'interprete a unire due frasi diverse, come nell'esempio illustrativo:

Cluster 4 in B.1:

En los últimos años las desigualdades no han hecho sino aumentar; ahora, queremos más igualdad de oportunidades y en el acceso a beneficios ambientales, a través de una política todo lo arriesgada que se quiera, pero necesaria.

Resa in P10:

negli ultimi anni • le disuguaglianze • >non hanno fatto altro che< aumentare (1,4s) ora vogliamo più uguaglianza di: opportunità (1,2s) nell'accesso ••• a:l/le/••• alle >varie politiche< ••

Osservando la tipologia delle omissioni nella condizione ST, si osserva che talvolta alcune omissioni sono state commesse in parti centrali della frase, portando l'interprete a cadere in controsensi. In questo caso, l'interprete omette una parte centrale della frase (in grassetto) e giunge a un'asserzione dal significato opposto rispetto al messaggio originale:

Cluster 14 in B.1:

La fase de adaptación y mitigación en la que estamos representa un momento fundamental. Y no está de más **hablar de las claves necesarias para contestar a las grandes preguntas que nos plantea** la transición ecológica.

Resa in P7:

la fase di adattamento e di mitigazione • in cui ci troviamo • rappresenta un momento ••• fondamentale •e:d è superfluo >forse< •• il fatto che • è fondamentale trovare le soluzioni per ••• questa sfida ecologica ••

La perdita di entità denominate nella fase di comprensione ha portato l'interprete a omettere l'intero referente senza il ricorso a strategie linguistiche. In questo caso, la perdita del referente in grassetto porta a una vuotezza semantica dell'output:

Cluster 12 in A.1:

De hecho, cabe destacar el gran éxito que ha tenido durante el periodo de apertura iniciado apenas hace diez días y que sigue abierto del **Programa DUS 5000**.

Resa in P5:

infatti bisogna: ricordare il grande successo ne:l >periodo di apertura< • cominciato: poco tempo fa (4,2s)

Come già illustrato, i numerali sono stati spesso oggetto di errori semantici. Di seguito vengono proposte due rese: una in cui il numerale è stato reso correttamente senza tuttavia un impatto significativo sulla resa dell'input testuale successivo (R1) e una in cui la difficoltà relativa alla resa del numerale ha avuto un impatto sulla comprensione dell'intera seconda parte del *cluster* (R2).

Cluster 15 in A.1

Además, el capítulo para la preservación de biodiversidad marina tiene un presupuesto de casi **medio millón** de euros. El mar debe preocuparnos, principalmente porque es uno de los principales **sumideros** de CO2. [pause]

Resa in P8 (R1):

e >in più< per • la/l'ambito della/la tutela della biodiversità marina • ha • un• >progetto di bilancio< **di un milione • di euro ••** e ci deve preoccupare perché è uno dei più grandi **assorbitori** di C O due **••**

Resa in P4 (R2):

per >preservare la biodiversità marina< (2,3s) e il bilancio in quest'ambito è di più **di un milione di euro** • il mare ci deve preoccupare perché è uno dei/<dei principali **fattori che>/di >emissioni di C O due< (2,6s)**

Per avere un termine di paragone utile dei dati ottenuti nella condizione ST e GM per quanto riguarda l'informatività ottenuta nelle due rispettive condizioni sperimentali, è stata calcolata la media di tutti i valori numerici medi ottenuti dalle singole partecipanti, di seguito sintetizzati nel grafico:

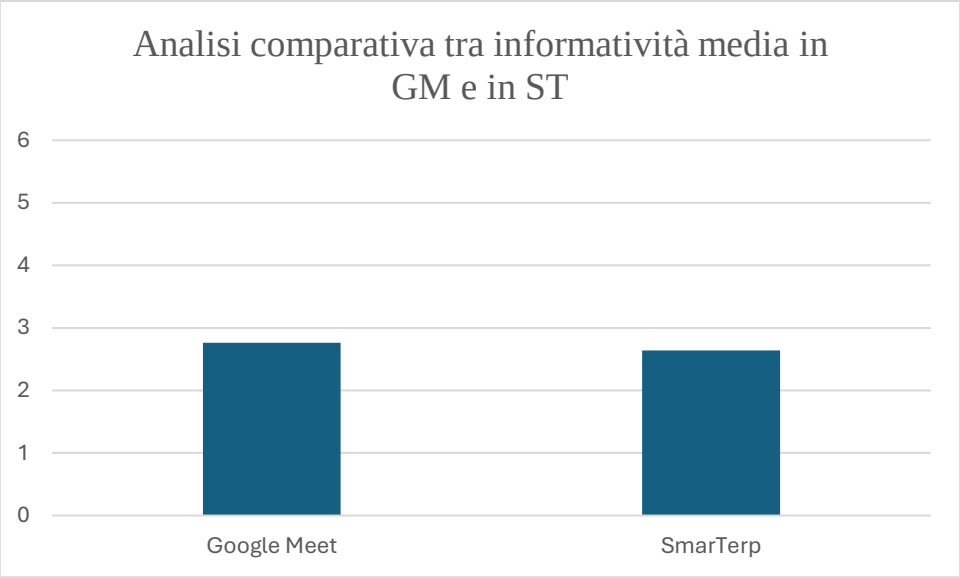


Figura 30 Informatività media in GM e in ST

Come si può osservare in Figura 30, la media relativa a tutte le prestazioni delle interpreti è più elevata di 0,12 in Google Meet; pertanto, le rese delle interpreti nella condizione ST sono risultate leggermente più informative rispetto alle rese in GM. Dato lo scarso margine tra GM e ST, si può dedurre che le condizioni GM e ST non abbiano avuto un impatto molto diverso sull'informatività della resa. Ciò che tuttavia è estremamente rilevante ai fini del presente studio è che i risultati ottenuti per la condizione ST siano molto simili a quelli della condizione GM, nonostante la minor presenza nell'input visivo di una maggior quantità di input verbale trascritto.

Per ogni partecipante, i valori medi in ST e in GM sono i seguenti:

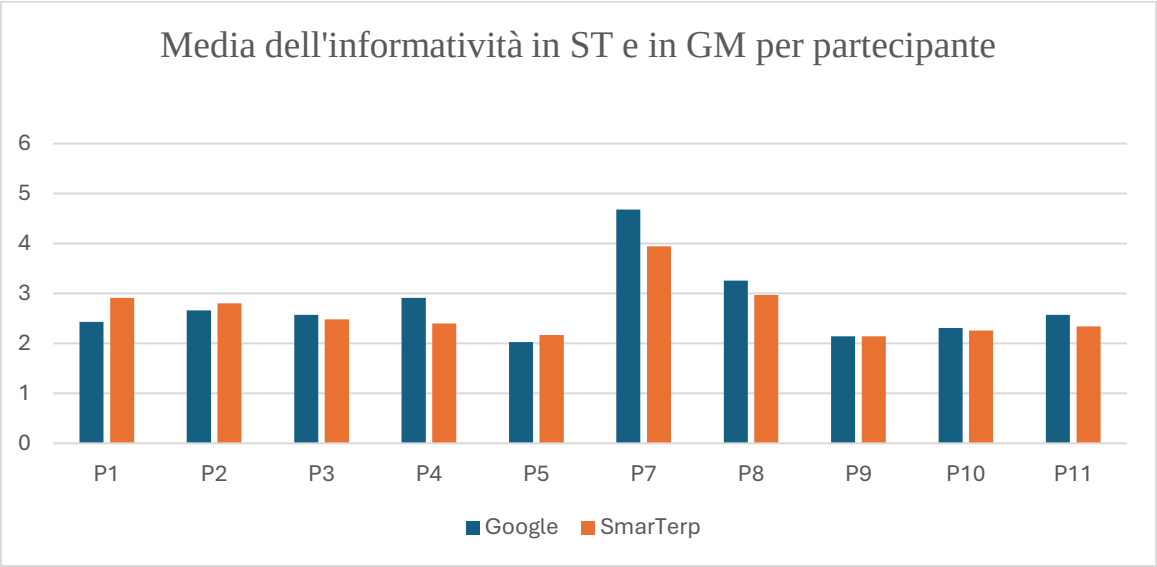


Figura 31 Analisi comparativa tra i valori medi relativi all'informatività in GM e ST per partecipante.

Osservando il grafico, una sola partecipante su 9 (P9) ha ottenuto un valore numerico medio relativo all'informatività dei singoli *cluster* uguale sia per GM che per ST. Inoltre, 6 partecipanti su 10 hanno ottenuto valori numerici relativi all'informatività superiori in GM rispetto a ST (con differenze variabili da un minimo di 0,05 a un massimo di 0,74). Le 3 partecipanti restanti hanno invece ottenuto valori numerici medi leggermente superiori in ST rispetto a GM (con differenze variabili da un minimo di 0,14 a un massimo di 0,48). Pertanto, considerata la proporzionalità inversa tra valore numerico medio e informatività (più alto è il valore numerico e più bassa è l'informatività), è possibile concludere che il 60% delle partecipanti ha prodotto rese leggermente più informative nella condizione ST, seppur con differenze tra la media in GM e in ST sempre inferiori a 1. Nonostante la differenza relativamente bassa, il dato interessante è che la trascrizione automatica di Google Meet non ha portato a rese più informative, nonostante il maggior input visivo-verbale contenente una maggior quantità di informazioni.

Per valutare la significatività statistica delle differenze relative alle medie ottenute nelle due condizioni sperimentali per quel che riguarda l'informatività, è stato condotto anche in questo caso il test di Wilcoxon (§2.7.2); in seguito alla correzione di Bonferroni, il livello di alfa è stato fissato a 0,0167. Il test di Wilcoxon ha rivelato una differenza statisticamente non significativa tra le due condizioni ($W = 31$, $p = 0,343$). La correlazione biseriale di rango è risultata positiva ($r_b = 0,378$), suggerendo una relazione non statisticamente differente tra le variabili esaminate.

Sulla base di questi risultati, possiamo concludere che non c'è una differenza statisticamente significativa tra GM e ST ($p = 0,343$); la grandezza della differenza è piccola-moderata ($d = 0,378$) e ST (mediana 2,44) sembra avere un piccolo vantaggio rispetto a GM (mediana 2,57) per quanto riguarda l'informatività delle rese, ma questa differenza non è significativa. Questi risultati suggeriscono che le due condizioni sono sostanzialmente equivalenti in termini di impatto sulla variabile dell'informatività ed eventuali differenze osservate potrebbero essere attribuibili al caso piuttosto che a un effetto reale sulle due condizioni, come ancor più evidente se viene applicata la correzione di Bonferroni (§3.2.1). Anche in questo caso, è necessario ribadire che i risultati sono di difficile generalizzazione, data la ridotta dimensione del campione.

3.3 *User Experience Questionnaire*

In questo paragrafo, verranno presentati i dati raccolti tramite l'*User Experience Questionnaire* (UEQ), mediante il quale sono stati ottenuti dati quantitativi circa la percezione dell'usabilità dello strumento proposto, così come quantificata da parte delle partecipanti. Nel grafico seguente (Figura 32) viene visivamente presentata la media dei dati ottenuti in GM e in ST per le seguenti categorie: attrattività, apprendibilità, efficienza, controllabilità, stimolazione e originalità.

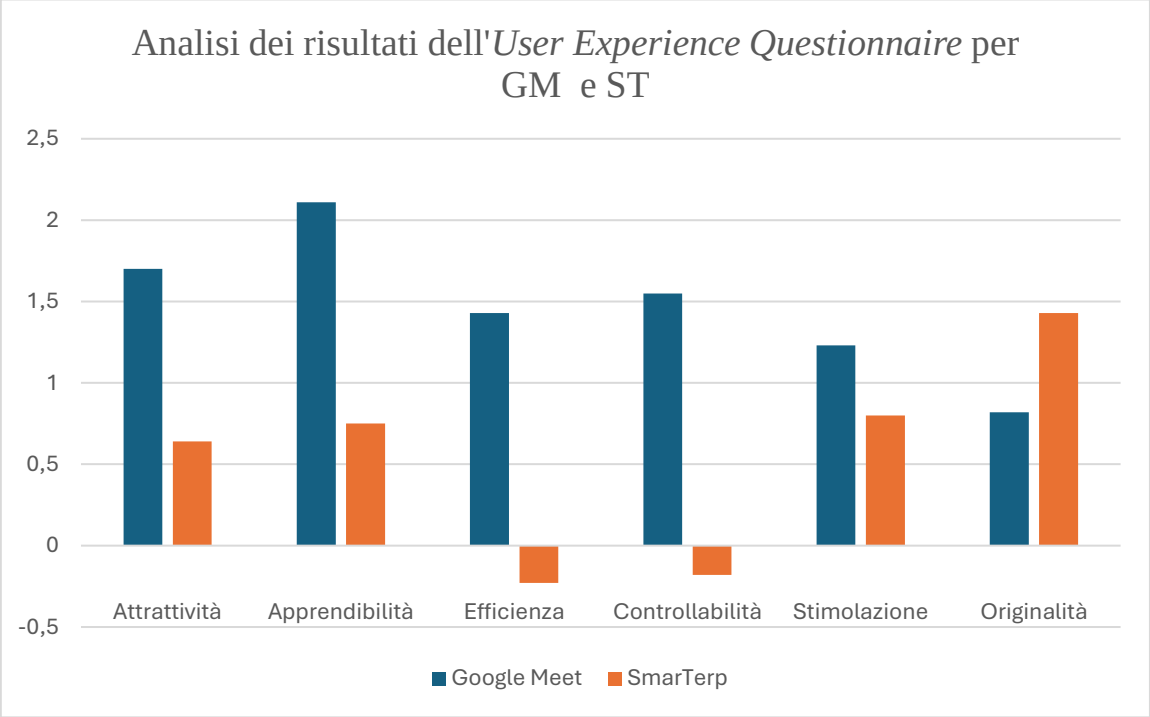


Figura 32 Risultati dell'User Experience Questionnaire per GM e ST a confronto.

In particolare, la descrizione dei dati ottenuti in GM e in ST è la seguente:

Categorie	Google Meet					
	Media	STD	N	Confidenza	Intervallo di confidenza	
Attrattività	1,70	0,72	11	0,43	1,27	2,12
Apprendibilità	2,11	0,52	11	0,31	1,81	2,42
Efficienza	1,43	0,56	11	0,33	1,10	1,76
Controllabilità	1,55	0,79	11	0,47	1,08	2,01
Stimolazione	1,23	1,05	11	0,62	0,61	1,85
Originalità	0,82	1,28	11	0,76	0,06	1,57

Tabella 25 Descrizione dei dati di UEQ per GM.

	SmarTerp					
Categorie	**Media**	**STD**	**N**	**Confidenza**	**Intervallo di confidenza**	
Attrattività	0,64	1,37	11	0,81	-0,17	1,45
Apprendibilità	0,75	1,66	11	0,98	-0,23	1,73
Efficienza	-0,23	0,84	11	0,50	-0,72	0,27
Controllabilità	-0,18	1,28	11	0,76	-0,94	0,58
Stimolazione	0,80	1,17	11	0,69	0,10	1,49
Originalità	1,43	0,80	11	0,47	0,96	1,90

Tabella 26 Descrizione dei dati di UEQ per ST.

Osservando la Figura 32, si osserva che in una sola categoria (originalità) la media ottenuta per SmarTerp supera la media ottenuta per Google Meet. Per le categorie relative alla stimolazione e all'originalità non si osserva una notevole differenza; mentre per le categorie relative all'attrattività, l'apprendibilità, l'efficienza e la controllabilità le differenze risultano estremamente pronunciate. Curiosamente, i risultati relativi alla percezione delle partecipanti sembrano collidere con il miglior andamento nella condizione ST relativamente all'accuratezza. Tuttavia, è proprio nei dati relativi all'UEQ che emerge una delle particolarità del campione coinvolto, ossia il coinvolgimento di interpreti in formazione.

Frittella (2023) ha condotto uno studio sperimentale mirato alla valutazione dell'usabilità di SmarTerp coinvolgendo un campione di sole interpreti professioniste. In particolare, nel suo studio, è stato utilizzato un modello di UEQ contenente tre scale in comune con il presente studio sperimentale (apprendibilità, efficienza, controllabilità). Di seguito, vengono paragonati i dati medi ottenuti nello studio di Frittella (2023) con quelli emersi in questo studio:

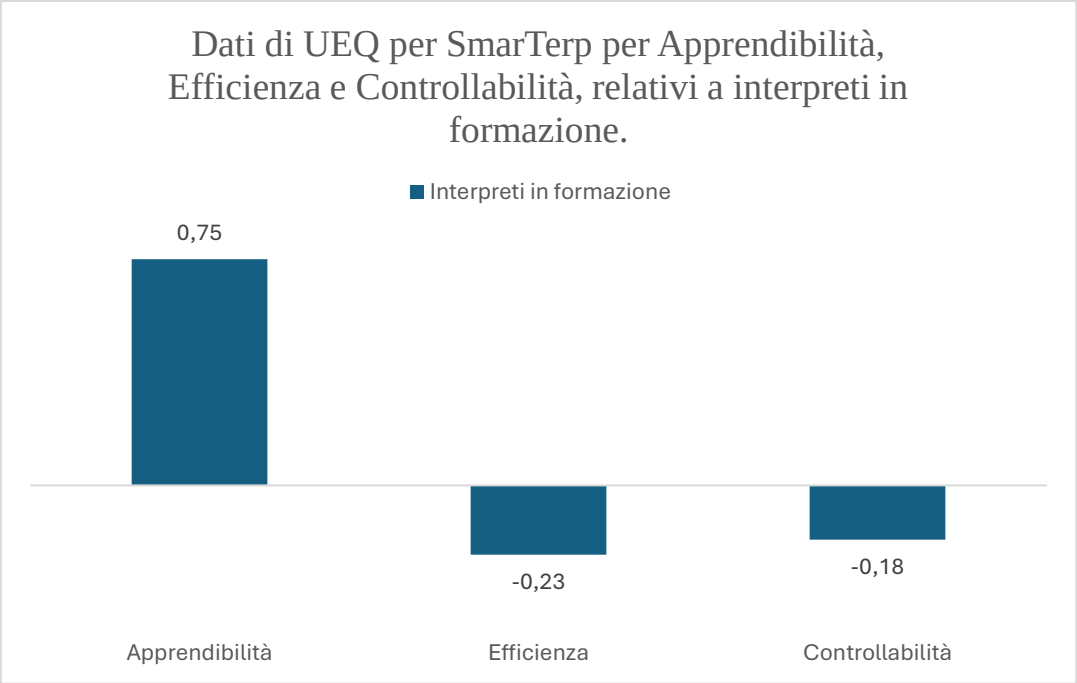


Figura 33 Dati di UEQ relativi a SmarTerp di interpreti in formazione.

Il valore medio relativo all'apprendibilità è pari a 0,75 ed è molto vicino al valore medio di 0,8 ottenuto da interpreti professioniste; pertanto, è possibile dedurre che entrambe le categorie di interpreti hanno percepito pressoché la medesima difficoltà nell'imparare a utilizzare il CAI *tool* di SmarTerp. Il valore negativo -0,23 ottenuto per la categoria dell'efficienza è nettamente distante dal valore medio pari a 1,9 che è invece stato ottenuto dalle interpreti professioniste; pertanto, è ipotizzabile che le interpreti professioniste abbiano saputo gestire in modo più efficiente tutti gli input presenti. Il valore negativo -0,18 raggiunto per la scala della controllabilità è molto distante dal valore medio di 1,4 ottenuto dalle interpreti professioniste: è quindi ipotizzabile che le interpreti professioniste abbiano avuto una maggior sicurezza nella gestione di tutti gli input, con una maggior capacità di controllo di tutti gli input.

Nonostante le sole tre scale messe a confronto, si osserva che i dati relativi alle interpreti professioniste corroborano quanto osservato anche per le interpreti in formazione relativamente all'apprendibilità di SmarTerp: l'apprendibilità dello strumento presenta infatti in entrambi i casi valori positivi ma non superiori alla soglia di 1. Tuttavia, il campione di interpreti con esperienza mostra medie nettamente distanti per quanto riguarda l'efficienza e la controllabilità dello strumento. Osservando i dati sopra menzionati, sembra evidente che la discrepanza sia dovuta non tanto allo strumento quanto alla diversa esperienza e sembra plausibile che le interpreti professioniste siano in grado di adattarsi più facilmente al nuovo strumento, grazie

alla maggiore esperienza. Questo dato già anticipa una necessità didattica, più che un problema di usabilità dovuto all'organizzazione dell'interfaccia del CAI.

CAPITOLO 4. Analisi e discussione dei dati di *eye-tracking* e delle attività di retrospezione

Per avere una prospettiva più ampia circa la modalità con cui gli strumenti sono stati utilizzati dalle partecipanti durante l'attività sperimentale, in questo capitolo verranno analizzati i dati quantitativi di *eye-tracking:* verranno da un lato descritti i risultati relativi al comportamento oculare prima e dopo la comparsa di 51 stimoli selezionati e dall'altro lato verranno illustrati i dati quantitativi relativi al *dwell time* trascorso nelle aree di interesse preconfigurate. Al termine dell'analisi dei dati quantitativi a disposizione, verrà proposta un'analisi qualitativo-descrittiva che metterà in luce proposte didattiche, difficoltà ricorrenti e approfondimenti specifici sui processi cognitivi così come emerso durante le attività retrospettive.

4.1 Dati di *eye-tracking*

In questa sezione verranno presentati i dati relativi al tracciamento del movimento oculare durante le due prove di IS sottoposte alle partecipanti in entrambe le condizioni sperimentali. Verranno presentati in un primo momento i dati riguardanti l'analisi quali-quantitativa, il cui obiettivo è stabilire se le partecipanti abbiano utilizzato gli strumenti prevalentemente in una fase di comprensione dell'input e prima di iniziare l'output vocale o se gli strumenti siano stati utilizzati solo dopo aver iniziato l'output vocale (e, quindi, nella fase di monitoraggio). Successivamente, verranno presentati i dati relativi alla distribuzione delle fissazioni sulle AOI nel periodo di interesse selezionato.

4.1.1 Dati quali-quantitativi

Dopo aver selezionato 51 stimoli, presenti sia nei suggerimenti del CAI *tool* di SmarTerp che nella trascrizione automatica di Google Meet, per ogni stimolo è stato osservato il movimento oculare della partecipante durante l'attività di simultanea ed è stato determinato se: (1) la partecipante guarda lo stimolo prima di iniziare l'output; (2) la partecipante guarda lo stimolo dopo aver iniziato la traduzione e non si riformula; (3) la partecipante guarda lo stimolo dopo aver iniziato la traduzione e si riformula migliorando l'output; (4) la partecipante guarda lo

stimolo dopo aver iniziato la traduzione e si riformula peggiorando l'output; (5) la partecipante non guarda lo stimolo; (6) il dato relativo al movimento oculare non è disponibile. Di seguito viene proposta un'analisi dei dati raccolti suddivisa in tre sezioni: nella prima verranno approfonditi i dati raccolti per la condizione sperimentale GM, nella seconda verranno discussi i dati della condizione ST e nella terza verrà proposta un'analisi comparativa relativa alle differenze in GM e ST. Per l'analisi dei dati, sono stati calcolati i valori medi dei punteggi raggiunti da ogni partecipante per ciascuna categoria. Per i dati di *eye-tracking*, è indispensabile ribadire che il campione preso in considerazione è stato di 11 partecipanti, a differenza del campione di 10 partecipanti considerato per la valutazione della *performance*.

Per quanto riguarda i dati raccolti relativamente al movimento oculare nella condizione GM in corrispondenza dei 51 stimoli selezionati, di seguito viene proposta la loro descrizione:

	Media	Mediana	SD	Minimo	Massimo
Guarda stimolo prima dell'output	35.091	38	9.492	17	46
Guarda stimolo dopo output e non si riformula	2.364	2	2.111	0	7
Guarda stimolo dopo traduzione e migliora output	1.818	2	1.328	0	4
Guarda stimolo dopo traduzione e peggiora output	1.455	1	1.572	0	4
Non guarda stimolo	9.727	6	10.209	0	28
Dato n/d	0.545	0	0.934	0	3

Tabella 27 Descrizione dei dati relativi al movimento oculare in corrispondenza degli stimoli selezionati nella condizione GM.

Come emerge chiaramente dalla tabella riassuntiva, in media oltre il 50 per cento degli stimoli sono stati guardati prima di iniziare l'output; infatti, solo 2 partecipanti su 11 hanno guardato un numero di stimoli inferiore al 50 per cento del totale prima di iniziare l'output, in quanto hanno preferito non guardare gli stimoli nella maggior parte dei casi. Nel caso delle 9 partecipanti restanti, in ben 5 casi più di 40 stimoli sono stati guardati prima di iniziare l'output raggiungendo un valore massimo pari a 46 stimoli guardati prima di iniziare l'output. Quando lo stimolo è stato guardato dopo aver iniziato la traduzione, tendenzialmente non si osservano riformulazioni dell'output. Infatti, in media 2,36 stimoli sono stati guardati dopo aver iniziato la traduzione e non sono state osservate riformulazioni: 2 partecipanti non hanno mai guardato gli stimoli dopo aver iniziato la traduzione senza riformularsi, 8 partecipanti hanno guardato lo stimolo senza riformularsi in un numero di casi compreso tra 1 e 4 e 1 partecipante ha guardato ben 7 stimoli dopo aver iniziato l'input senza riformularsi.

Nei casi in cui le partecipanti si sono riformulate dopo aver iniziato l'output, si osserva che in media sono stati leggermente maggiori i casi in cui le partecipanti si sono riformulate

migliorando l'output rispetto ai casi in cui le partecipanti si sono riformulate peggiorando l'output. Per quanto riguarda i casi in cui le partecipanti hanno guardato gli stimoli dopo aver iniziato la traduzione riformulandosi migliorando l'output, si osserva che solo 2 partecipanti non si sono mai riformulate dopo aver guardato la trascrizione di GM migliorando l'output, mentre le 9 partecipanti restanti si sono riformulate dopo aver guardato l'output scritto della trascrizione automatica in un numero di casi compreso tra 1 e 4, migliorando l'output verbale. Per quanto riguarda i casi in cui le partecipanti hanno guardato gli stimoli dopo aver iniziato la traduzione riformulandosi peggiorando l'output, si osserva che ben 4 partecipanti non si sono mai riformulate dopo aver guardato la trascrizione automatica peggiorando l'output, mentre le 7 partecipanti restanti si sono riformulate dopo aver guardato gli stimoli in un numero di casi compreso tra 1 e 4, peggiorando la loro resa.

Nei dati raccolti emerge una tendenza relativamente nitida: nella condizione GM la maggior parte delle partecipanti guarda lo stimolo o prima o dopo l'inizio dell'output vocale. Infatti, solo 2 partecipanti hanno ottenuto un valore numerico medio relativo agli stimoli non guardati superiore al 50 per cento degli stimoli (con valori pari a 27 e 28), 2 partecipanti non hanno guardato un numero di stimoli compreso tra 10 e 20, 5 partecipanti non hanno guardato un numero di stimoli compreso tra 1 e 10 e 2 partecipanti hanno guardato tutti gli stimoli oppure il dato non era disponibile. Per quanto riguarda i dati non disponibili, è opportuno ricordare che questi possono essere sinonimo di perdita del dato oppure di fissazioni fuori dalle AOI presenti sullo schermo. Sette partecipanti non hanno presentato dati n/d sugli stimoli selezionati, mentre 4 partecipanti hanno presentato un numero di stimoli compreso tra 1 e 4 in concomitanza dei quali non è presente nel *software* di analisi dei dati *EyeLink Data Viewer* alcun indizio circa il movimento oculare.

Per quanto riguarda i dati raccolti relativamente al movimento oculare nella condizione ST in corrispondenza dei 51 stimoli selezionati, di seguito viene proposta la loro descrizione:

	Media	Mediana	SD	Minimo	Massimo
Guarda stimolo prima dell'output	16.727	17	6.451	8	28
Guarda stimolo dopo output e non si riformula	11.727	11	6.901	2	23
Guarda stimolo dopo traduzione e migliora output	4.909	5	2.468	0	9
Guarda stimolo dopo traduzione e peggiora output	0.364	0	0.674	0	2
Non guarda stimolo	12.273	12	7.115	3	23
Dato n/d	5.000	0	9.306	0	31

Tabella 28 Descrizione dei dati relativi al movimento oculare in corrispondenza degli stimoli selezionati nella condizione ST.

Come si può osservare nella descrizione relativa ai dati per la condizione ST, la media degli stimoli guardati prima di iniziare l'output è nettamente inferiore a 25 (corrispondente al 50 per cento degli stimoli). In un solo caso, nella condizione ST, sono stati guardati più del 50 per cento degli stimoli prima di iniziare l'output. Entrando nei dettagli delle prestazioni delle singole partecipanti, solo 3 partecipanti hanno guardato prima di iniziare l'output vocale un numero di stimoli compreso tra 20 e 25, 5 partecipanti hanno invece guardato prima di iniziare la traduzione un numero di stimoli compreso tra 10 e 20 e 2 partecipanti hanno guardato solo 8 stimoli prima di iniziare l'output. È ipotizzabile, sulla base del ridotto valore numerico relativo agli stimoli guardati prima di iniziare l'output, un impatto significativo della latenza del CAI *tool* di SmarTerp sull'adozione delle strategie durante l'IS. Questo è infatti confermato dal fatto che un numero significativo di stimoli è stato guardato dopo aver iniziato la traduzione (e quindi nella fase di monitoraggio dell'output), pur senza riformulazioni da parte delle partecipanti. Infatti, si osserva una media di 11 stimoli guardati dopo aver iniziato la traduzione che tuttavia non hanno portato ad alcuna riformulazione: 4 partecipanti hanno guardato un numero di stimoli compreso tra 1 e 10 senza riformularsi, 6 partecipanti hanno guardato tra 10 e 20 stimoli dopo aver iniziato la traduzione degli stessi e senza riformularsi e 1 partecipante ha guardato un numero superiore a 20 stimoli dopo aver iniziato la resa degli stessi e senza riformularsi. Nei casi in cui le partecipanti si sono riformulate dopo aver guardato gli stimoli, si osserva che le riformulazioni hanno portato a un miglioramento dell'output nella maggior parte dei casi, con una media di 4,9 stimoli guardati dopo l'inizio della traduzione e riformulati migliorandone l'accuratezza. In particolare, è opportuno osservare che solo 1 partecipante non ha presentato almeno una riformulazione con miglioramento della resa dopo aver guardato l'output del CAI *tool* ed è opportuno osservare che per questa partecipante non sono stati disponibili i dati per ben 31 stimoli (quindi per un numero di stimoli superiore al 50 per cento del totale). Per quanto riguarda le 10 partecipanti restanti, hanno tutte presentato un numero di riformulazioni dopo aver guardato il suggerimento del CAI (pur avendo iniziato la traduzione dello stimolo) compreso tra un minimo di 2 stimoli e un massimo di 9 stimoli, con miglioramenti significativi relativamente all'accuratezza della resa. Nonostante gli errori e le inaccuratezze del CAI, spicca un dato particolarmente curioso, ossia la media di 0,3 per quanto riguarda i casi in cui le interpreti hanno guardato lo stimolo dopo aver iniziato la traduzione riformulandosi peggiorando l'output. Infatti, ben 8 partecipanti hanno riportato un numero di stimoli riformulati con peggioramento della resa dopo aver guardato lo stimolo pari a 0 e in sole 3 partecipanti gli stimoli riformulati con peggioramento della resa sono stati o 1 o 2.

Osservando i dati relativi agli stimoli non guardati dalle partecipanti, emerge che in tutti i casi il numero degli stimoli non guardati è inferiore al 50 per cento: 4 partecipanti non hanno guardato (prima o dopo l'inizio dell'output) un numero di stimoli inferiore a 10, 5 partecipanti non hanno guardato un numero di stimoli compreso tra 10 e 20 e solo 2 partecipanti non hanno guardato un numero di stimoli superiore a 20 (in entrambi i casi uguale a 23). In merito ai dati non disponibili per la condizione ST, si osserva che per 6 partecipanti su 11 non sono presenti dati non disponibili, mentre in 4 casi il numero di stimoli i cui dati non sono disponibili è di 3 o di 9 stimoli; tuttavia, in 1 caso per ben 31 stimoli (corrispondenti a oltre il 50 per cento del totale) i dati relativi al movimento oculare in corrispondenza di essi non sono disponibili a causa di una perdita di calibrazione durante la prova di IS nella condizione ST.

Di seguito viene proposto un grafico in cui vengono sintetizzati e paragonati i valori medi dei dati raccolti per la condizione ST e GM per ogni categoria:

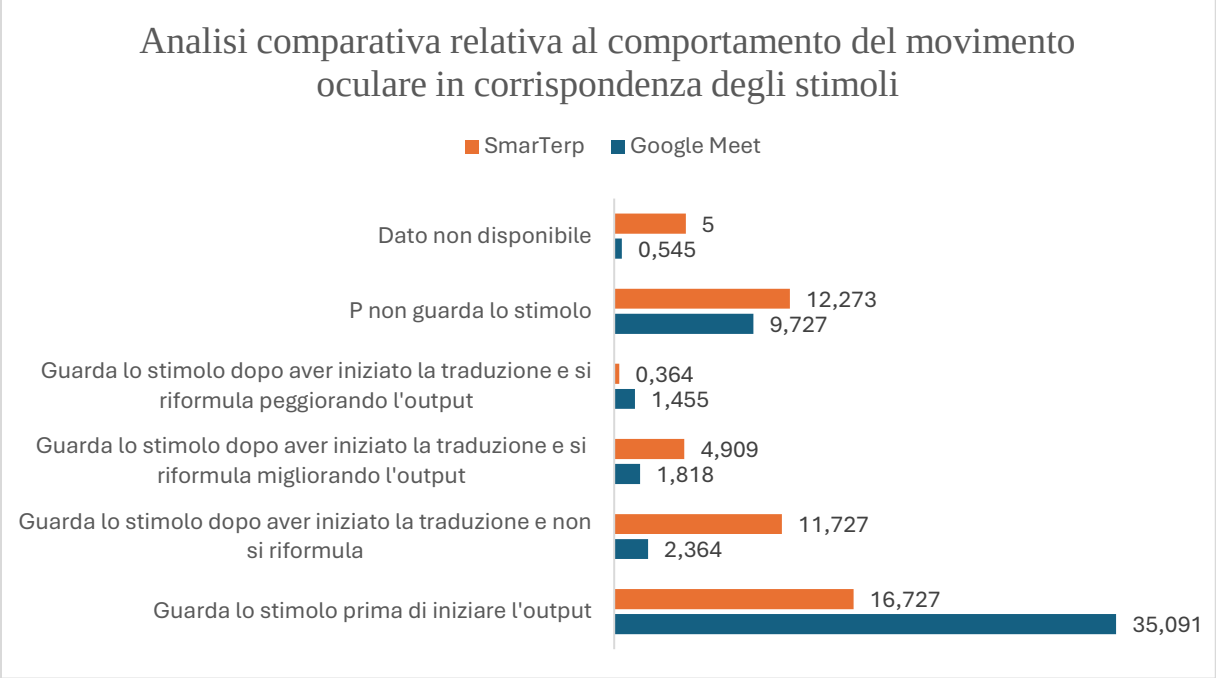


Figura 34 Grafico comparativo relativo al comportamento del movimento oculare in corrispondenza dei 51 stimoli selezionati.

Osservando il grafico, risulta visivamente chiara la tendenza a guardare lo stimolo prima di iniziare l'output nella condizione GM, con una differenza di circa il doppio rispetto al numero di stimoli che sono stati guardati prima di iniziare l'output nella condizione ST. Nonostante questo valore elevato sia probabilmente dovuto alla ridotta latenza dell'ASR di Google Meet e alla tendenza a eseguire una traduzione a vista da parte di molte partecipanti, è ipotizzabile un vantaggio per la condizione ST: il valore numerico più basso nella condizione ST può infatti essere sinonimo di una minore fiducia da parte delle interpreti nei confronti del CAI *tool* e al

contempo può essere indice di una eccessiva fiducia delle partecipanti nei confronti del testo scritto della trascrizione automatica di Google Meet. Tra i possibili vantaggi di un numero inferiore di stimoli guardati prima di iniziare l'output può esserci una maggior concentrazione sull'input auditivo e sulla comprensione di esso, tralasciando in alcune parti l'input verbale-visivo. Se si osserva il confronto tra i dati relativi agli stimoli che sono stati guardati solo dopo l'inizio della traduzione, si osserva che sia in GM che in ST la maggior parte delle volte non si sono osservate riformulazioni, ma lo strumento è stato utilizzato in fase di monitoraggio. In particolare, nella condizione ST è nettamente superiore il numero di stimoli guardati dopo l'inizio della traduzione e non riformulati: tra le cause possibili si può rinvenire ancora una volta l'elevata latenza del CAI e l'utilizzo del CAI durante il monitoraggio della resa. Nei casi in cui si sono osservate delle riformulazioni della resa a seguito della comparsa dei suggerimenti o del testo scritto, si osserva che nella condizione ST il numero di riformulazioni che hanno portato a un miglioramento della resa è superiore di circa 3 stimoli rispetto alla condizione GM. Nonostante siano stati sporadici i casi di peggioramento della resa dopo aver guardato gli stimoli nella trascrizione automatica o nei suggerimenti del CAI, si osserva una differenza media di circa 1 stimolo tra GM e ST che ha visto un valore medio più basso in ST. Per quanto riguarda gli stimoli non guardati (né prima né dopo l'inizio dell'output) si osserva che in ST una media di 12 stimoli non sono stati guardati, mentre in GM in media 9 stimoli non sono stati guardati. Nonostante la differenza sia relativamente bassa (di soli 3 stimoli), questa sembra confermare la tendenza delle interpreti a eseguire una traduzione a vista nella condizione GM, nel corso della quale non vengono selezionati gli stimoli presenti nell'input visivo e in cui si assiste a una maggior probabilità di perdita di concentrazione sull'input auditivo. Relativamente ai dati non disponibili, è necessario chiarire che il valore nettamente più elevato di dati non disponibili per la condizione ST è stato dovuto a una perdita della calibrazione nel corso della prova di IS di una partecipante, per la quale non sono stati disponibili i dati per 31 stimoli. Inoltre, come vedremo successivamente nel corso della discussione delle indagini retrospettive, alcune partecipanti hanno adottato alcune strategie per cercare di gestire la complessità dell'input verbale-visivo sovrapposto all'input verbale-auditivo che hanno portato a una perdita dei dati relativi al comportamento del movimento oculare in corrispondenza di alcuni stimoli: alcune hanno chiuso gli occhi per qualche secondo e altre hanno guardato fuori dallo schermo per qualche istante. È necessario, tuttavia, sottolineare che la frammentarietà dei dati (data la non disponibilità di alcuni dati) e la dimensione ridotta del campione rendono i dati difficili da generalizzare.

4.1.2 Dati quantitativi

In questa sezione verranno presentati i dati di *eye-tracking* relativi al tempo trascorso dall'occhio di ogni partecipante (o *dwell time*) in ogni AOI configurata, in termini percentuali. Verranno presentati e descritti in primo luogo i dati relativi alla condizione GM; successivamente verranno illustrati i dati relativi alla condizione ST e verranno infine discussi i dati per la condizione GM e per la condizione ST messi a confronto.

Prima di procedere alla descrizione dei dati raccolti relativamente alle AOI per la condizione GM, è necessario ricordare che per questa condizione sperimentale sono state configurate tre AOI: una contenente l'intero schermo del PC utilizzato dalle partecipanti per la prova di IS (*ScreenG*); una contenente il riquadro con l'oratrice e uno sfondo bianco (*SpeakerG*) e una contenente la trascrizione automatica in costante movimento (*SubtitlesG*). Considerate queste tre AOI, di seguito si propone la descrizione per i dati raccolti per le 11 partecipanti, in termini percentuali:

	ScreenG	SpeakerG	SubtitlesG
N	11	11	11
Mancanti	0	0	0
Media	96.5	28.6	62.6
Mediana	97	21	70
Deviazione standard	2.66	23.3	24.9
Minimo	90	4	17
Massimo	99	71	88

Tabella 29 Descrizione dei dati relativi al tempo trascorso dall'occhio delle partecipanti nelle tre AOI configurate per la condizione GM, in termini percentuali.

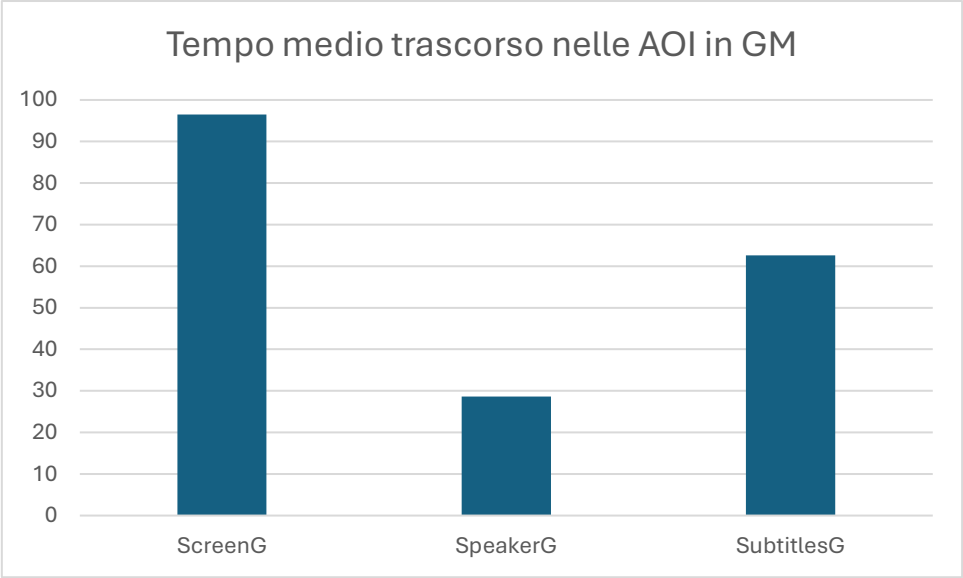


Figura 35 Grafico dei valori medi del dwell time nelle AOI nella condizione GM.

Osservando i dati raccolti, è possibile affermare che il campione coinvolto nello studio ha guardato complessivamente lo schermo per un tempo uguale o superiore al 90 per cento del tempo totale: ben 5 partecipanti su 11 hanno guardato lo schermo per un tempo uguale o superiore al 98 per cento del tempo totale. Osservando i *dwell time* delle AOI relative all'oratrice e alla trascrizione automatica, si può notare che la media del tempo trascorso sull'AOI *SubtitlesG* supera l'AOI *SpeakerG* di 34 punti percentuali; infatti, in sole 3 partecipanti il tempo trascorso sull'AOI relativa all'oratrice è superiore rispetto al tempo trascorso sull'AOI relativa alla trascrizione automatica (con differenze che oscillano dai 2 ai 47 punti percentuali). In 3 casi il tempo trascorso sull'AOI relativa all'oratrice è risultato uguale o inferiore al 10 per cento del tempo totale e in 10 casi su 11 il tempo trascorso sull'AOI relativa all'oratrice è risultato inferiore al 50 per cento del tempo totale. Per quanto riguarda l'AOI relativa alla trascrizione automatica, si osserva che 7 partecipanti su 11 hanno trascorso oltre il 50 per cento del tempo totale proprio in quest'area di interesse (da un minimo pari al 69 per cento a un massimo pari all'88 per cento). Nei restanti 4 casi, i valori vedono un'oscillazione tra un minimo pari al 17 per cento e un massimo uguale al 48 per cento di tempo trascorso in quest'AOI.

Sulla base di questi risultati è ipotizzabile una minor concentrazione delle partecipanti sull'input audio e para-verbale dell'oratrice (ad esempio, relativo al movimento delle labbra), dato il maggior *focus* sul testo scritto, in costante movimento, presente nell'AOI *SubtitlesG*.

Prima di procedere alla descrizione dei dati raccolti relativamente alle AOI per la condizione ST, è necessario ricordare che per questa condizione sperimentale sono state configurate sei

AOI: una contenente l'intero schermo del PC utilizzato dalle partecipanti per la prova di IS (*Screen_S*); una contenente il riquadro con l'oratrice e uno sfondo bianco (*Speaker_S*); una contenente le tre colonne con i suggerimenti generati automaticamente dal CAI *tool* (*Suggestions_S*); una contenente la colonna del CAI relativa ai numeri (*Numbers_S*); una contenente la colonna del CAI relativa alla terminologia specialistica (*Terms_S*) e una contenente la colonna del CAI relativa alle entità denominate (*Named_Entities_S*). Considerate queste tre AOI, di seguito si propone la descrizione dei dati raccolti per le 11 partecipanti, in termini percentuali:

	Screen_S	Speaker_S	Suggestions_S	Numbers_S	Terms_S	Named_Entities_S
N	11	11	11	11	11	11
Mancanti	0	0	0	0	0	0
Media	94.2	68.0	25.2	5.64	15.5	4.00
Mediana	97	72	22	5	13	4
Deviazione standard	9.15	15.8	15.0	3.07	10.9	1.73
Minimo	67	41	10	1	4	2
Massimo	99	89	57	12	37	7

Tabella 30 Descrizione dei valori relativi al tempo trascorso dall'occhio delle partecipanti nelle sei AOI configurate per la condizione ST, in termini percentuali.

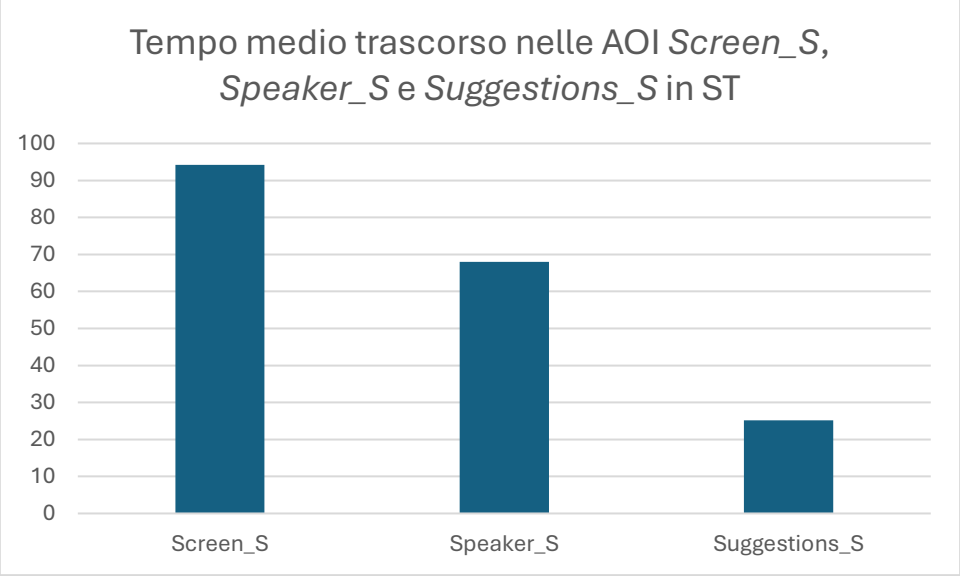


Figura 36 Grafico dei valori medi del dwell time nelle AOI Screen_S, Speaker_S e Suggestions_S nella condizione ST.

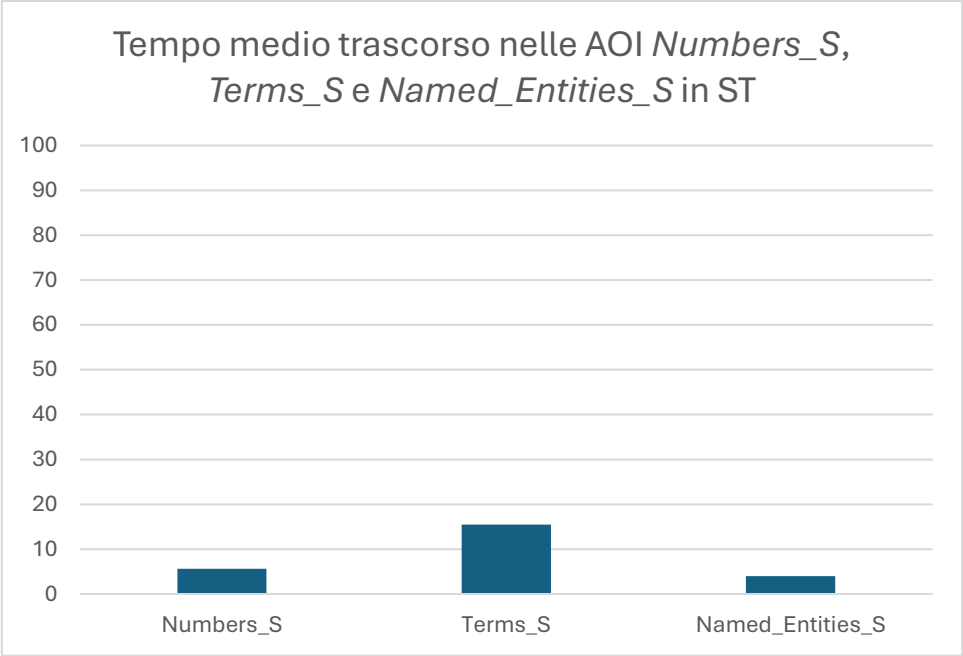


Figura 37 Grafico dei valori medi del dwell time nelle AOI Numbers_S, Terms_S e Named_Entities_S in ST.

Come osserviamo dal valore minimo del tempo trascorso sull'AOI *Screen_S*, tutte le partecipanti hanno guardato lo schermo del computer per oltre il 50 per cento del tempo; in particolare, 9 partecipanti su 11 hanno guardato lo schermo per un tempo uguale o superiore al 95 per cento del tempo totale e, di queste, 4 partecipanti hanno guardato lo schermo per un tempo uguale o superiore al 98 per cento del tempo totale. In un solo caso il *dwell time* sullo schermo è stato inferiore al 70 per cento (67%).

Osservando i valori relativi al *dwell time* sull'AOI relativa all'oratrice, è possibile affermare che 10 partecipanti su 11 hanno guardato l'AOI *Speaker_S* per un tempo uguale o superiore al 50 per cento del tempo totale (fino a un massimo pari all'89%); in particolare, in 3 casi il tempo trascorso sull'AOI contenente l'oratrice è stato uguale o superiore all'80 per cento del tempo totale.

I *dwell time* ottenuti dalle singole partecipanti per l'AOI relativa all'oratrice superano in 10 casi su 11 i valori relativi al tempo trascorso sull'AOI contenente i suggerimenti del CAI *tool* (*Suggestions_S*). In particolare, una sola partecipante ha trascorso sull'AOI *Suggestions_S* un tempo superiore al 50 per cento del tempo totale; le altre partecipanti hanno ottenuti valori compresi tra un minimo pari al 10 per cento e un massimo pari al 43 per cento.

Per quanto riguarda le 3 AOI contenenti le tre diverse colonne presenti nell'interfaccia utente del CAI *tool* di SmarTerp, si osserva che la colonna che ha riportato un valore maggiore relativamente al *dwell time* è quella contenente la terminologia specialistica (*Terms_S*); infatti,

10 partecipanti su 11 hanno raggiunto valori maggiori in questa AOI rispetto a quelle contenenti le altre due colonne del CAI. Nel caso in cui le partecipanti hanno trascorso un maggior tempo su un'altra AOI tra le tre relative ai suggerimenti del CAI, la colonna relativa ai numeri ha riportato un valore più alto rispetto alle altre due. Relativamente al tempo trascorso sull'AOI *Terms_S* si osserva, inoltre, che 7 partecipanti hanno guardato questa area di interesse per oltre il 10 per cento del tempo, raggiungendo un valore massimo pari al 37 per cento.

Tra le tre AOI corrispondenti alle tre colonne dell'interfaccia del CAI *tool* di SmarTerp, emerge chiaramente un tempo trascorso nettamente inferiore per quanto riguarda l'AOI corrispondente alle entità denominate (*Named_Entities_S*): i valori medi non hanno in nessun caso superato il 10 per cento e in 9 casi su 11 il tempo trascorso su questa AOI è stato uguale o inferiore al 5 per cento.

Per quanto riguarda l'AOI relativa ai numeri (*Numbers_S*), in soli due casi il tempo trascorso su di essa è risultato superiore al 10 per cento e in 7 casi il tempo trascorso su questa è risultato uguale o inferiore al 5 per cento del totale.

Sulla base di questi risultati, sembra emergere una tendenza a una maggiore osservazione dell'AOI relativa all'oratrice, probabile sintomo di una maggior concentrazione sull'input auditivo-verbale nella fase di comprensione e di un minor attaccamento alle risorse verbali scritte. Inoltre, durante il tempo trascorso sulle AOI relative ai suggerimenti del CAI le partecipanti hanno trascorso la maggior parte del tempo sulla colonna relativa alla terminologia specialistica, pur utilizzando anche le altre due.

Sulla base delle AOI configurate, è possibile individuare 3 AOI paragonabili tra la condizione ST e la condizione GM: quella contenente tutto lo schermo del computer; quella relativa all'oratrice e quella contenente tutti i suggerimenti dello strumento a disposizione (la trascrizione automatica in GM e le tre colonne contenenti numeri, termini specialistici e entità denominate in ST). Come discusso precedentemente, in entrambe le condizioni risulta particolarmente elevato il tempo trascorso sull'AOI relativa a tutto lo schermo; pertanto, si procede adesso a una breve descrizione comparativa delle due AOI più rilevanti: quella corrispondente all'oratrice e quella contenente i suggerimenti dello strumento tecnologico. Osservando i valori medi raggiunti in ciascuna di queste due AOI nelle rispettive condizioni sperimentali, si ottiene il grafico seguente:

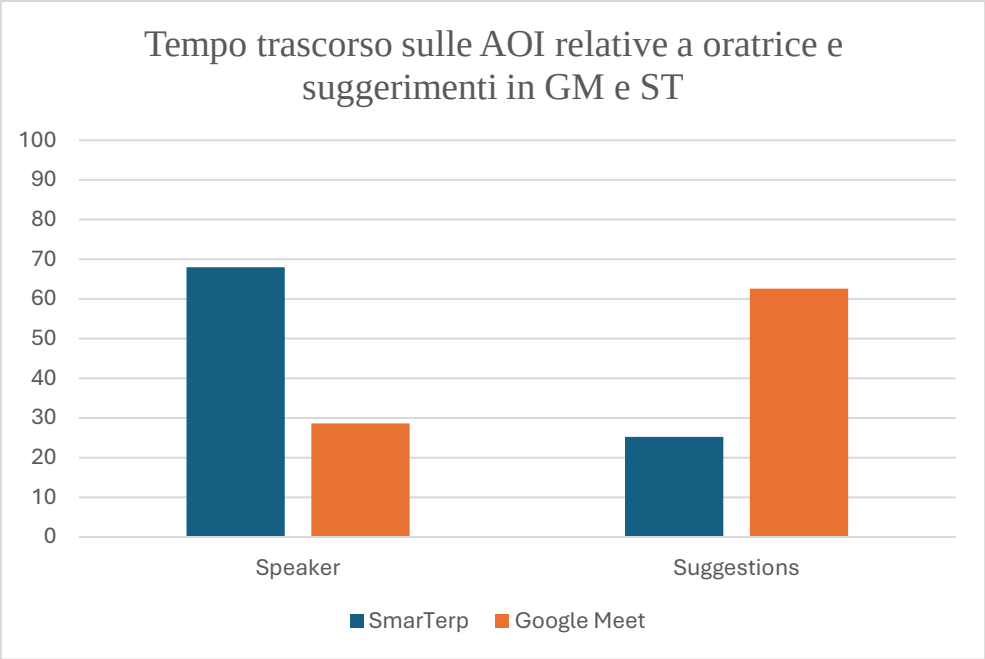


Figura 38 Analisi comparativa dei valori medi relativi al dwell time sulle AOI contenenti oratrice e suggerimenti in GM e ST.

Osservando il grafico emerge chiaramente la tendenza a un maggior tempo trascorso sull'AOI relativa all'oratrice nella condizione sperimentale contenente il CAI *tool* di SmarTerp, con una differenza di 39,4 punti percentuali. Si nota parallelamente un maggior tempo trascorso sull'AOI contenente i suggerimenti nella condizione GM, con una differenza di 37,4 punti percentuali. Integrando questi dati con il miglioramento osservato nell'intelligibilità delle rese nella condizione ST, si osserva che un tempo maggiore trascorso sull'oratrice sembra essere direttamente proporzionale a una maggior concentrazione sulle risorse auditivo-verbali e sull'output dell'interprete stessa. Pertanto, guardare con tempi più prolungati l'oratrice sembra aver avuto una ricaduta positiva sulla prestazione. Il maggior allontanamento dall'input visivo-scritto – favorito dal CAI *tool* di SmarTerp – sembra favorire un maggior distaccamento dalla struttura del testo originale, portando inevitabilmente l'interprete a una maggior concentrazione nella fase di comprensione del testo originale.

4.2 Attività di retrospezione

Dopo aver sintetizzato e descritto i dati quantitativi raccolti nel corso dello studio sperimentale, in questa sezione verranno presentati i dati qualitativi ottenuti mediante le retrospezioni. Prima di descrivere i dati raccolti, è necessario ricordare che ogni retrospezione ha avuto luogo subito dopo ogni prova di IS e il *retrieval cue* utilizzato per ogni retrospezione è costituito dall'input

auditivo originale e dall'input video del discorso originale con la sovrapposizione del movimento oculare nel corso dell'IS. Ogni partecipante ha fermato il video nei momenti in cui lo ha ritenuto più opportuno evidenziandone gli aspetti a suo avviso più importanti, senza alcuna interferenza da parte del ricercatore.

Per entrambe le condizioni sperimentali le partecipanti hanno più volte proposto commenti e riflessioni circa le difficoltà riscontrate nella resa di numeri, entità denominate e termini specialistici. Pertanto, la disamina dei dati qualitativi inizierà proprio con la presentazione delle riflessioni riguardanti queste tre categorie lessicali. In secondo luogo, verranno presentati i commenti relativi alla percezione della latenza dell'ASR, per poi passare alla descrizione dell'impatto dello strumento utilizzato (sia in GM che in ST) sull'adozione delle strategie dell'interprete; in particolare, verranno discusse situazioni problematiche che hanno portato a omissioni, calchi e allontanamento dello sguardo dagli stimoli. Successivamente, verranno discussi gli impatti degli strumenti sull'andamento prosodico dell'interprete, dovuti molto spesso a una percezione della prosodia del discorso diversa tra GM e ST. Verrà poi fornita un'ampia panoramica sulla gestione degli input visivo e auditivo, che più volte è stata oggetto di riflessioni e commenti da parte delle partecipanti. Posteriormente, verranno presentate alcune riflessioni in merito alle strategie adottate per quanto riguarda il monitoraggio della resa e per recuperare eventuali informazioni perse nella fase di comprensione con l'ausilio degli strumenti proposti.

Dopo aver illustrato alcuni elementi di riflessione emersi sia per GM che per ST, verranno discussi fattori emersi solo per GM e fattori emersi esclusivamente per ST. In particolare, per quanto riguarda la condizione sperimentale contenente la trascrizione automatica di Google Meet verrà discussa la punteggiatura della trascrizione automatica (e il potenziale impatto sulla resa) e verrà illustrata la tendenza di molte partecipanti all'esecuzione di una traduzione a vista di frammenti molto ampi del discorso. Per quanto riguarda la condizione sperimentale contenente il CAI *tool* di SmarTerp verranno invece affrontati alcuni elementi relativi alla percezione dell'organizzazione dell'input visivo di SmarTerp e alcune difficoltà riscontrate a causa della ridondanza informativa del sistema.

Dopo aver osservato alcuni elementi specifici per ogni condizione sperimentale, verranno discussi alcuni commenti relativi alla comparazione diretta di Google Meet e SmarTerp come sistemi di ausilio nel corso dell'IS. Successivamente, verranno sintetizzate alcune riflessioni relative alla percezione dell'usabilità degli strumenti, con particolare enfasi sulle aspettative delle partecipanti.

4.2.1 Numeri

Per quanto riguarda la condizione GM, P1, P6, P8 e P11 hanno fornito interessanti riflessioni relativamente alla difficoltà nella gestione dei numeri. P1 ha affermato che la difficoltà nella gestione dei numeri è una difficoltà indipendente dalla trascrizione automatica, ma ritiene che la presenza nella trascrizione automatica della struttura "mil millones" abbia acuito la tendenza al calco e la difficoltà nell'elaborazione del numero: afferma infatti che la struttura dissimmetrica del numerale ha generato confusione nella fase di comprensione e chiarisce la necessità di un numero già con la struttura italiana, se lo strumento vuole essere di ausilio all'interprete. P6 afferma di aver chiuso gli occhi nel 90 per cento dei casi, in quanto vedere il numerale nella trascrizione avrebbe inficiato negativamente nella fase di percezione e ascolto del numerale. P6 ritiene di aver a volte chiuso gli occhi subito dopo aver osservato rapidamente il numerale scritto: questo atteggiamento può essere interpretato come probabile sintomo di sovraccarico cognitivo dovuto alla presenza di molti stimoli che esercitano una maggior pressione sulla comprensione del discorso. P8 afferma di aver fatto ricorso alla trascrizione ogniqualvolta ci fossero numeri, percentuali, dati nell'input audio; tuttavia, evidenzia che la costante aspettativa di una trascrizione con un ottimo output ha portato più volte a un'interruzione nella fase di ascolto/comprensione che nei casi con numeri di ordine di grandezza dei miliardi ha invece avuto esito negativo (un maggior sovraccarico cognitivo e una percezione del peggioramento della resa) data la difficoltà nell'elaborazione del numerale scritto. L'eccessiva fiducia di P8 nell'accuratezza dell'ASR ha portato la partecipante a ridurre il proprio *décalage*, rimanendo attaccata alla trascrizione automatica e questo l'ha portata a seguire le false partenze e le riformulazioni compiute dall'oratrice (e, quindi, dalla trascrizione automatica). P8 afferma di aver adottato delle strategie efficienti, relativamente alla gestione dei numerali, solo nella seconda parte della resa, quando ha deciso di fidarsi di meno della trascrizione nella fase di comprensione e di utilizzarla solo nella fase di monitoraggio della resa dei numeri. P8 sostiene che la trascrizione automatica non abbia un criterio uniforme nel trascrivere i numeri (ad esempio, alcune volte i miliardi vengono proposti con la struttura "tre zeri + millones" e altre volte vengono trascritti con la struttura "mil millones") e questa mancanza di uniformità ritiene che abbia generato tanta confusione da aver confuso l'ordine di grandezza dei miliardi con quello dei milioni, più di una volta. P11 afferma che la trascrizione del numero ha portato a una deconcentrazione dall'ascolto di esso e afferma che l'unica

strategia possibile, data la perdita di concentrazione nella fase di ascolto, fosse leggere direttamente il numero.

Sempre per quanto riguarda la condizione GM, P3, P9 e P11 hanno evidenziato invece l'utilità della trascrizione automatica nella resa dei numeri. P3, infatti, sottolinea che la struttura dissimmetrica "mil millones" sia più facile da tradurre se viene trascritta letteralmente (e quindi senza convertirla in numerale con tanti zeri), in quanto si tratterebbe di una struttura più volte incontrata e automatizzata. P9 afferma di aver cercato nella trascrizione quasi esclusivamente i numeri, in quanto erano affidabili e accurati. Anche P10 ribadisce l'utilità della trascrizione automatica per una resa accurata dei numeri.

Per quanto riguarda la condizione ST, P1, P2, P3, P4, P6, P7, P8, P10, P11 hanno riscontrato notevoli difficoltà nella gestione dei numeri, in parte a causa della latenza del CAI *tool* e in parte a causa dell'output del numerale nel CAI *tool* (in cui si osserva che l'ordine di grandezza non viene scritto verbalmente ma viene determinato dal numero di zeri presenti). P1 ha evidenziato la difficoltà nella comprensione di alcuni numeri che il CAI *tool* ha riformulato più volte (nonostante non fossero riformulati dall'oratrice), generando una confusione nella comprensione che si sovrappone a una difficoltà a priori, ammessa dalla stessa partecipante, nella gestione dei numeri. Inoltre, P1 ritiene che la latenza del sistema del CAI porti spesso a una perdita dell'elemento cui fa riferimento il numerale e afferma che sarebbe di grande supporto un CAI in cui viene mostrato il referente cui fa riferimento il numerale accanto ad esso. P2 ritiene che gli zeri presenti nei numeri con ordine di grandezza dei milioni o dei miliardi incrementino il carico cognitivo nella fase di elaborazione e proprio a causa della lunghezza del numerale molto spesso ha optato per omissioni volontarie dello stesso. P3 ritiene che la latenza del CAI abbia avuto un impatto negativo ancor più notevole, in termini di carico cognitivo, nei momenti in cui l'oratrice si è riformulata, in quanto in quei momenti l'output nel CAI della prima formulazione dell'oratrice giungeva solo dopo la seconda formulazione dell'oratrice, generando confusione e smarrimento. Afferma inoltre che, dopo una prima fase iniziale di confusione, ha optato per un ascolto autonomo e indipendente dal CAI (cercando di comprendere i numeri attraverso l'immaginazione di essi con cifre grandi) in cui ha optato più volte per l'approssimazione dei numeri. Come accennato anche dalle partecipanti precedenti, P4 ha affermato che molto spesso la confusione generata dal CAI era dovuta principalmente alla latenza. P6 sostiene che sono di difficile elaborazione i numeri che nel CAI *tool* abbiano un numero di zeri superiore a tre, in quanto inutilizzabili e ribadisce l'impatto della latenza sulla perdita del referente del numerale. P7 afferma di aver spostato il proprio sguardo

immediatamente sulla colonna del CAI relativa ai numerali ogniqualvolta sentisse un numero nell'input audio e ritiene di essersi *fidata ciecamente* del numero proposto dal CAI, nonostante la sua difficoltà nel leggere i numeri con un numero elevato di zeri: dice infatti di aver spesso letto i numeri andando per intuito in base alla loro lunghezza. P8 afferma di aver interrotto molto spesso l'ascolto dei numerali data la certezza della comparsa di questi nella colonna preposta del CAI: la perdita di attenzione ha portato, tuttavia, a una confusione e a una minor sicurezza nella resa dei numeri. P8 ritiene estremamente significativi gli errori del CAI nella resa di alcuni numerali ("2,85%" e "medio millón", resi rispettivamente con "285%" e "un milione" dal CAI, ad esempio), in quanto l'incongruenza audio/suggerimenti del CAI ha portato un inutile sovraccarico cognitivo. P10 ribadisce l'impatto della lunghezza dei numeri (con molti zeri) e della latenza dei CAI sulla resa dei numeri e afferma che nei pochi casi in cui il suggerimento del CAI sarebbe servito è arrivato troppo tardi per poter essere utilizzato. P10 critica particolarmente gli errori del CAI in cui i numeri sono stati divisi in più numeri, portando a un inutile aumento degli elementi presenti nell'input visivo, generando confusione. P11 aggiunge che sarebbe stato molto utile avere le unità di misura di una maggior quantità di numeri, in quanto alcuni numerali, seppur chiari nell'input visivo, sono stati inutilizzati a causa della perdita (nella memoria a breve termine) dell'unità di misura menzionata. Complessivamente, nelle retrospezioni, nessuna partecipante ha messo in luce vantaggi significativi circa l'uso del CAI per la gestione dei numerali.

In sintesi, per entrambe le condizioni le partecipanti hanno riportato una difficoltà nell'elaborazione dei numerali indipendente dalle tecnologie proposte, sintomo di inesperienza e di necessità di formazione relativamente alla gestione dei numeri. Inoltre, le sporadiche inaccuratezze del CAI e della trascrizione hanno portato molte interpreti a perdere la fiducia nello strumento proposto, segnale di un mancato consolidamento previo relativamente alla gestione consapevole di un input che è inevitabilmente imperfetto. A causa di questa sfiducia e della percezione di un ritardo eccessivo nell'output dei numerali, molte partecipanti hanno affermato di aver utilizzato gli strumenti in fase di monitoraggio. Da queste opinioni emerge dunque la necessità di formare le interpreti affinché acquisiscano una maggior capacità nell'elasticizzare il proprio *décalage*, in modo da poter sfruttare l'input visivo-verbale non solo in fase di monitoraggio, ma anche in fase di ascolto e percezione dell'input.

4.2.2 Entità denominate

Per quanto riguarda la condizione GM, P2, P5, P6 e P10 hanno fornito alcuni commenti in merito alla difficoltà nella gestione delle entità denominate con l'ausilio della trascrizione automatica di Google Meet. P2 ha infatti affermato che quando ha guardato la trascrizione dell'entità denominata "Programa DUS 5000" (trascritto con "Programa 2 5000") si è resa conto che il nome non avesse alcun senso ed ha percepito un aumento del carico cognitivo in quanto ha dovuto perdere tempo nel trovare una strategia di emergenza che fosse in grado di salvare un'entità denominata che non aveva compreso, anche a causa dell'incongruenza tra l'input visivo-verbale e l'input auditivo-verbale**.** P5 afferma di aver percepito molto spesso, per quanto riguarda gli acronimi, un'incongruenza tra input audio e input visivo-verbale, che ha portato la partecipante a fidarsi poco dello strumento quando vi fossero acronimi nella trascrizione. P6 ha lamentato invece l'assenza di maiuscole per alcuni acronimi quale elemento potenzialmente fuorviante nella comprensione dell'input. Anche P10 ha percepito diversi errori nella trascrizione di alcuni acronimi ed ha preferito omettere direttamente le sigle con le minuscole, in quanto non affidabili. P10 ha inoltre affermato che l'assenza di traduzione delle entità denominate nell'input audio ha spesso portato a soluzioni traduttive improvvisate che hanno richiesto talvolta una maggior concentrazione sull'input: P10, ad esempio, afferma che per interpretare l'entità denominata "bono social térmico" ha dovuto rileggere più volte la trascrizione per poter trovare una soluzione affidabile e accurata anche in base al contesto sintattico in cui veniva utilizzata.

Nonostante le difficoltà riscontrate, per quanto riguarda le retrospezioni relative alla condizione GM, P9 e P10 hanno sottolineato alcuni vantaggi della trascrizione automatica. Sia P10 che P9 hanno ribadito l'utilità della trascrizione nella resa del nome proprio "Francisco José Contreras": l'elevata accuratezza dell'ASR di Google Meet, infatti, permette un'ottima comprensione dei nomi propri, in quanto questi non necessitano una traduzione verso la lingua di arrivo. P10 ribadisce l'utilità della trascrizione anche per il nome proprio del piano menzionato dall'oratrice ("Plan España Puede"). P10 ribadisce come in alcuni casi l'output verbale-scritto sia stato utile per il riconoscimento di specifiche entità denominate ma non per la traduzione delle stesse; P10 menziona infatti l'esempio di "fondos FEDER": nonostante l'accurata trascrizione della sigla, la mancata disponibilità della traduzione ha avuto un impatto significativo sull'accuratezza della resa.

Per quanto riguarda la condizione ST, P2, P3, P5, P6 e P10 hanno commentato alcuni problemi riscontrati nella gestione delle entità denominate con il CAI *tool* di SmarTerp e la maggior parte delle difficoltà si sono concentrate nei punti dei discorsi in cui l'aspettativa di un suggerimento affidabile relativamente alle entità denominate è stata delusa. P2 afferma infatti di aver atteso a lungo la sigla AEH2 (che fa riferimento al nome di un'associazione) osservando la colonna di destra, ma questa non è comparsa: P2 ha salvato la frase optando per una generalizzazione, nonostante la destabilizzazione abbia portato a una perdita di fiducia nello strumento. P3 si aspettava che anche il nome proprio "Francisco José Contreras" comparisse nell'interfaccia del CAI e ha pertanto interrotto l'ascolto nella fase di comprensione (sia per stanchezza che per l'aspettativa che le apparisse il nome), giungendo a una totale perdita del nome. P5 ritiene che molte unità lessicali comparse nella colonna delle entità denominate fossero estremamente ridondanti e inutili (ad esempio, fa riferimento alla parola "Governo", più volte comparsa nella colonna di destra, ma che viene ritenuta inutile in quanto ad elevata disponibilità per un'interprete, anche in formazione). P6 ha reiterato la problematicità dell'omissione del nome proprio "Francisco José Contreras" e "Asociación AEH2" tra i suggerimenti del CAI. P10 afferma che la mancanza della sigla "CMNUCC" tra i suggerimenti del CAI ha comportato una perdita totale di fiducia nei confronti del CAI, che ha portato l'interprete a fare affidamento sul CAI solo ed esclusivamente nella fase di monitoraggio delle entità denominate successive, optando per un ascolto puro nella fase di comprensione che fosse autonomo e indipendente dal CAI. Per quanto riguarda i casi in cui il CAI *tool* ha avuto un impatto positivo sulla gestione delle entità denominate, P1 menziona l'utilità del CAI nei casi in cui esso è stato in grado di riconoscere e tradurre correttamente le entità denominate presenti nel testo, nonostante la latenza. Infatti, P1 menziona l'ottima accuratezza del CAI nella resa di "bono social térmico" e di "fondos FEDER": l'interprete si è infatti ricordata della presenza di queste due entità denominate nei glossari proposti durante la formazione, pur non ricordandosi la traduzione e ha deciso di aspettare l'output del CAI, certa della sua utilità.

Riassumendo quanto esposto, molte partecipanti hanno affermato che la condizione GM non ha aiutato in alcun modo nella resa delle entità denominate con traducente equivalente in lingua italiana; ma risulta essere stata più utile per la resa di nomi propri che non presentano un traducente nella lingua di arrivo. Il CAI di SmarTerp è risultato più utile in quanto ai traducenti, ma l'omissione di alcuni nomi propri ha portato le interpreti a non fare più affidamento sui suoi suggerimenti. Sembra dunque rafforzarsi la necessità di una formazione mirata su una gestione critica dell'input visivo-verbale, in modo da non far totalmente perdere all'interprete la fiducia nei suggerimenti del CAI, solo a causa di sporadiche omissioni.

4.2.3 Terminologia specialistica

Nelle retrospezioni riguardanti la condizione sperimentale GM, P2, P3, P4, P6, P8 e P10 hanno sottolineato notevoli difficoltà nella comprensione e nella resa della terminologia specialistica presente nei discorsi, a causa della scarsa disponibilità immediata dei traducenti dei termini specialistici. P2 afferma di non aver potuto ricorrere a una traduzione rapida per i termini "percolación" e "floculación" e ha messo in atto una strategia di sintesi parlando di "due fasi molto importanti". Anche P3 ha riscontrato particolari difficoltà nella resa di questi due termini specialistici, ma ha optato per una strategia diversa rispetta a quella osservata per P2: P3 ha infatti sostituito in entrambi i casi il suffisso "-ción" con l'equivalente italiano "-zione", pur perdendo sicurezza circa l'accuratezza del proprio output. P6 afferma di aver inventato i traducenti, pur di evitarne l'omissione. P4 ha invece fornito un commento interessante relativamente alle unità di misura poco conosciute che ha incontrato durante la prova di IS: P4 ritiene che avrebbe preferito non leggere un'unità di misura poco nota (nella fattispecie, "megajulios"), in quanto la ridondanza informativa di un elemento non noto sia nell'input visivo-verbale che nell'input auditivo-verbale ha portato solo a una maggior confusione nella comprensione, ostacolando il reperimento del traducente dalla memoria a lungo termine. P8 afferma che soffermarsi sul termine "megajulios" ha avuto un impatto significativo su tutta la parte successiva del *cluster*, giungendo a generalizzazioni, omissioni ed errori semantici: quanto descritto risulta particolarmente interessante in quanto sintetizza una dinamica da interprete non-esperta. Oltre a ribadire la necessità di adottare una strategia di sintesi per i termini "percolación" e "floculación", P10 ha aggiunto che per il lessema "plantas potabilizadoras" è stata necessaria l'adozione di una strategia di ampliamento, con cui ha cercato di spiegare un concetto chiaro nella comprensione ma con un traducente non immediatamente disponibile. Tuttavia, la stessa partecipante sottolinea che l'adozione di una strategia così temporalmente significativa ha portato a una perdita di concentrazione sulla parte immediatamente successiva al lessema, che ha in parte recuperato grazie alla presenza della trascrizione di quella parte sullo schermo, ma con una selezione delle informazioni abbastanza superficiale. P10 ha infine aggiunto che alcune unità di misura non chiare hanno portato a una resa incerta e con alcune omissioni, oltre che a un'intonazione percepita come più robotica nelle parti più altamente informative. Per quanto riguarda una gestione relativamente positiva della terminologia specialistica, P8 afferma di aver utilizzato la trascrizione automatica solo nella

fase di monitoraggio per i termini specialistici, in modo da evitare un'interferenza diretta durante il processo di ascolto/traduzione.

Per quanto riguarda la condizione sperimentale ST, le partecipanti hanno tendenzialmente apprezzato il supporto del CAI *tool* di SmarTerp per la resa dei termini specialistici. P2, P3 e P10 sono d'accordo nell'affermare che senza il suggerimento di SmarTerp avrebbero sbagliato la traduzione del termine "iridio", nonostante la somiglianza con la lingua italiana. P1 afferma di aver chiuso gli occhi in alcune parti del discorso e di averli aperti esclusivamente quando sentiva un termine specialistico nell'input auditivo, in quanto sapeva che sarebbe stato di grande aiuto l'ausilio del CAI per la resa di quegli stimoli. P4 afferma che la colonna della terminologia specialistica è stata particolarmente utile anche per disambiguare possibili dubbi sulla traduzione di termini (anche noti): P4 ritiene infatti che molti suggerimenti di questa colonna abbiano in molti casi aiutato l'interprete a non cadere in calchi. P6, P10 e P11 evidenziano l'utilità del CAI per la resa del termine "ecologización", nonostante la latenza del sistema. P8 ritiene che la terminologia specialistica sia stata abbastanza di aiuto, ma al contempo sostiene che la traduzione di termini già noti (ad esempio, "sistema energetico" o "coesione territoriale") abbia istintivamente portato l'interprete a guardare i suggerimenti nella fase di monitoraggio della resa, distraendola nella fase di comprensione dei frammenti testuali immediatamente successivi. P8 ha riportato la difficoltà di trattenere le informazioni immediatamente successive ai termini specialistici, data la latenza del sistema nel proporre i traducenti per i termini specialistici. P10 ritiene che i traducenti dei termini specialistici proposti dal CAI *tool* di SmarTerp permettano una maggior accuratezza nella resa, permettendo l'utilizzo di termini leggermente più accurati e precisi rispetto a quelli che l'interprete avrebbe utilizzato (seppur con una semantica equivalente). P6 afferma invece che in molti casi il CAI *tool* ha presentato omissioni di termini specialistici, non apparentemente motivate, e la cui traduzione sarebbe stata di gran beneficio per le interpreti (in particolare, per i termini "vehículos compartidos" e "circulación rodada").

In sintesi, nonostante le strategie applicate da molte interpreti per la ricerca di equivalenti italiani sfruttando la presenza di strutture morfologicamente simili tra italiano e spagnolo nella condizione GM, la condizione ST risulta essere stata molto più utile per raggiungere una maggior accuratezza nella resa dei termini specialistici. In molti casi, le partecipanti hanno evidenziato come la presenza dei termini specialistici nell'input visivo-verbale di SmarTerp abbia di fatto alleggerito l'elaborazione del discorso, giungendo a strategie di successo e a una maggior sicurezza da parte delle interpreti circa il proprio output.

4.2.4 Percezione della latenza dell'ASR

Per quanto riguarda la condizione GM, solo P3 ha riportato un impatto significativo della latenza dell'ASR di Google Meet sull'elaborazione degli input visivo e auditivo. In particolare, P3 afferma di aver utilizzato la trascrizione solo nella fase di monitoraggio della resa e non nella fase di comprensione, proprio a causa della latenza dell'ASR: il ritardo della trascrizione non avrebbe consentito alla partecipante una gestione ottimale dei diversi input. P3 afferma di aver solo sporadicamente aspettato la trascrizione in caso di elenchi e in quei casi ha sfruttato la trascrizione per recuperare rapidamente le informazioni, pur interrompendo l'ascolto dell'input auditivo. P3 afferma che avrebbe preferito appuntarsi i numeri presenti nel discorso su un foglio cartaceo autonomamente rispetto a leggerli nella trascrizione generata automaticamente, in quanto ritiene di essere stata in grado aspettare e leggere la trascrizione solo nei punti in cui l'oratrice ha fatto delle pause nel discorso.

A differenza della condizione descritta precedentemente, per la condizione ST 8 partecipanti su 11 (P1, P3, P4, P5, P6, P9, P10 e P11) hanno sottolineato l'impatto negativo della latenza dell'ASR sull'elaborazione del discorso in input. P1 ha affermato di aver fatto dei silenzi molto lunghi a causa della latenza del sistema nella resa dei numeri, in quanto molto spesso, pur avendo a disposizione il numero, la perdita del referente cui faceva riferimento ha portato l'interprete a non essere in grado di ricollegare il numerale con il proprio referente. P3 ritiene che la latenza nell'output dei numerali nel CAI sia stata ulteriormente aggravata dalla struttura dei numeri con un elevato numero di zeri; pertanto, P3 ha preferito molto spesso evitare il CAI per la resa dei numeri, affidandosi quasi esclusivamente a una comprensione dei numerali indipendente e autonoma. P3 ribadisce a più riprese che molto spesso la latenza del sistema ha portato a un'incapacità di ricollegare i diversi stimoli in una resa accurata rispetto all'input originale. P4 osserva invece che il CAI ha un'elevata latenza soprattutto per quanto riguarda l'output dei numeri e non tanto per quanto riguarda la terminologia specialistica, che viene percepita come sufficientemente rapida da poter essere utilizzata in modo sicuro. Anche P5 osserva una latenza eccessivamente elevata per quanto riguarda i numeri: ha infatti preferito affidarsi esclusivamente all'ascolto puro dei numerali, senza aspettare gli stimoli. P6 e P9, come già osservato dalle partecipanti precedenti, hanno ribadito che la latenza nell'output dei numeri ha talvolta portato l'interprete a perdere il referente cui facessero riferimento. P9 ha inoltre osservato che l'attesa stessa della comparsa dell'output dei numerali è stata un fattore di distrazione, in quanto difficile da gestire e portatrice di ulteriore ansia. P10 ritiene di non aver

sfruttato in modo ottimale lo strumento proprio a causa della sua latenza, che ha portato la partecipante a utilizzarlo prevalentemente in fase di monitoraggio, pur riconoscendo che molto spesso lo strumento proponeva solo traducenti con lievi differenze semantiche e stilistiche rispetto a quelli proposti dall'interprete. P11 sottolinea che la latenza dell'ASR l'ha spesso portata a non utilizzare lo strumento nella resa prevalentemente per due motivi: l'attesa dell'output del CAI avrebbe portato a uno smarrimento dell'interprete e l'attesa dei suggerimenti l'avrebbe portata a interrompere l'ascolto dell'audio. Quasi tutte le partecipanti che hanno commentato la latenza dell'ASR hanno anche sottolineato che molto spesso la latenza dell'ASR ha comportato problemi nella resa, anche data la scarsa capacità delle stesse nell'avere un *décalage* elastico e flessibile; quest'ultimo fattore mette in luce la necessità di maggior esperienza e formazione per poter utilizzare a proprio vantaggio un CAI come quello di SmarTerp.

In sintesi, nella condizione GM il sistema è stato percepito tendenzialmente con una latenza minima, mentre nella condizione ST l'inesperienza delle partecipanti le ha portate a non poter sfruttare in molti casi l'input proprio a causa di una latenza non pienamente gestibile. La mancata flessibilizzazione del proprio *décalage* ha infatti portato molte interpreti a utilizzare il CAI di SmarTerp solo in fase di monitoraggio passivo.

4.2.5 Impatto degli strumenti sulle strategie dell'interprete

Sia la trascrizione automatica di Google Meet che il CAI *tool* di SmarTerp hanno portato le interpreti a una gestione strategica di numerose difficoltà, alcune volte con esiti positivi e altre volte cadendo in errori. Per quanto riguarda la condizione GM, P2, P4, P6, P7 e P8 hanno proposto alcune riflessioni circa le strategie che hanno messo in atto per poter sfruttare a proprio vantaggio la trascrizione automatica. P2 ritiene che la latenza estremamente bassa di Google Meet abbia permesso di leggere la trascrizione anticipando la parte di input ancora non elaborata dall'interprete e in questo modo ha potuto correggere, nella fase di comprensione, alcuni errori di trascrizione. Pur non facendo riferimento a esempi specifici nel discorso, P3 sostiene di essere sempre stata molto attaccata alla trascrizione, ma ritiene al contempo che nelle parti più informative la lettura della trascrizione abbia portato a un'elevata insicurezza circa l'accuratezza della propria resa, in quanto non pienamente consapevole dell'accuratezza dell'output rispetto all'input. P4 afferma di aver potuto riformulare più agevolmente l'input

auditivo proprio grazie alla presenza della trascrizione dello stesso. P6 ritiene di aver migliorato notevolmente il proprio output solo dal momento in cui ha deciso di non dire tutto quello che vedeva trascritto e, quindi, da quando ha cominciato a selezionare le informazioni nel corso di un ascolto più attivo. P7 afferma che nei casi in cui ha optato per la nominalizzazione del verbo la trascrizione ha rappresentato un ostacolo nel mantenimento della coerenza sintattico-semantica, in quanto la trascrizione portava l'interprete a rimanere più attaccata alle strutture sintattico-grammaticali del testo originale. P8 ha inizialmente optato per una resa più attaccata al testo originale e ha cercato di ridurre il proprio *décalage*, pur di rimanere attaccata all'input auditivo, cercando di sfruttare le pause per terminare alcune frasi; tuttavia, nella seconda parte del discorso, P8 ha deciso di guardare meno la trascrizione, in quanto ha cominciato a essere un fattore di deconcentrazione sull'input auditivo-verbale. Nella seconda parte, P8 ha infatti deciso di utilizzare la trascrizione solo nei casi in cui rimaneva indietro, per recuperare le informazioni perse nella fase di ascolto.

Per quanto riguarda la condizione ST, P1, P2, P3, P6, P8 e P10 hanno proposto alcune riflessioni circa le strategie che hanno messo in atto per poter sfruttare a proprio vantaggio la trascrizione automatica. P1 ha manifestato un'elevata sfiducia nei confronti dell'input visivo del CAI di SmarTerp: la partecipante ha infatti affermato di essere stata confusa dalla molteplicità dei riquadri presenti nell'interfaccia dello strumento e ritiene di essere stata in grado di concentrarsi maggiormente sull'ascolto dell'input auditivo solo nei momenti in cui ha deciso di chiudere gli occhi. P2 afferma che nei momenti in cui sigle o entità denominate non sono comparse nell'output del CAI è stato utile mettere in atto delle strategie di generalizzazione del concetto, sintomo dell'avvenuta comprensione della struttura semantica sottostante. P2, P3, P6 e P8 affermano di avere in più punti generalizzato per arrotondamento i numeri spesso non immediatamente chiari, data la loro lunghezza nell'output del CAI. Ad esempio, P6 afferma di aver interpretato "10 veces menos iridio" con "meno iridio", sintetizzando un concetto che seppur meno accurato dimostra l'avvenuta comprensione. P10 ha ritenuto indispensabile allungare il proprio *décalage* per poter sfruttare a proprio vantaggio il CAI nella seconda parte del discorso, pena il suo mancato utilizzo nella fase di comprensione del discorso. Al contempo, P10 ha definito il CAI come relativamente imprevedibile, in quanto solo alcune sigle sono comparse durante il discorso: P10 ha infatti optato per una gestione strategica in cui ha preferito non affidarsi troppo ai suggerimenti, assicurandosi di aver autonomamente capito tutti i concetti prima di controllarli nel CAI in fase di monitoraggio. Questa strategia ritiene che l'abbia portata, ad esempio, a generalizzare il nome proprio "Asociación AEH2" con "associazione", a causa della mancata comparsa dello stimolo tra i suggerimenti del CAI.

Per quanto riguarda le omissioni commesse nella resa dei discorsi nella condizione GM, P1 afferma di aver commesso prevalentemente delle omissioni nei momenti in cui non avesse immediatamente disponibili i traducenti di elementi lessicali presenti nell'input, pur avendo chiaro il concetto in lingua spagnola. Sempre P1 ritiene di aver fatto molta confusione nella gestione degli elenchi e di non essere stata in grado di mettere in campo delle strategie utili per una gestione ragionata degli elenchi. È dunque ipotizzabile che la trasposizione della densità informativa nell'input visivo-verbale abbia portato a un sovraccarico nella partecipante che non ha permesso una selezione ragionata degli elementi presenti in elenchi semanticamente densi. Per quanto riguarda le omissioni commesse nella resa dei discorsi nella condizione ST, P4 afferma che la densità informativa per la presenza di numeri ha portato in un momento l'interprete a perdere consapevolezza circa la struttura del discorso, obbligandola a commettere una lunga omissione.

Per quanto riguarda la percezione dei calchi commessi nelle rispettive condizioni sperimentali, P3, P7 e P8 hanno descritto alcuni casi in cui la trascrizione automatica di Google Meet ha portato le interpreti a commettere alcuni calchi. P3 ha descritto la percezione di aver commesso una maggior quantità di calchi nei momenti in cui si è concentrata sul testo scritto anziché sull'ascolto puro dell'input auditivo-verbale. P7 ha riscontrato la stessa percezione ed ha affermato di aver deciso di non guardare la trascrizione in alcune parti del discorso proprio per evitare di cadere in inutili calchi. P8 ritiene di aver calcato la struttura iniziale del discorso e di essere stata in grado di commettere meno calchi da quando ha deciso di guardare di meno la trascrizione. P2 e P6 hanno descritto alcuni casi in cui l'omissione di elementi del discorso nell'output del CAI ha portato le interpreti a cadere in una maggior quantità di calchi. In particolare, P2 afferma che l'omissione della sigla "AIE" nell'output del CAI *tool* ha portato l'interprete a calcare la sigla ascoltata senza aver potuto riformularla. P6 ritiene che sarebbe stata utile la presenza tra i suggerimenti del CAI di parole ad alta frequenza nel vocabolario della lingua spagnola ma note per essere morfologicamente dissimmetriche rispetto agli equivalenti in lingua italiana (ad esempio "conocimiento").

In sintesi, nella condizione GM è emersa la tendenza a eseguire una traduzione a vista della trascrizione automatica, riconducibile all'inesperienza delle partecipanti. Questa tendenza le ha infatti portate a rimanere attaccate alle strutture sintattico-grammaticali della lingua spagnola, con un'inibizione delle strategie interpretative. Non a caso, nonostante le scarse riformulazioni, le omissioni che sono state commentate non erano frutto della volontà di applicare una strategia di selezione delle informazioni, ma erano dovute esclusivamente al sovraccarico dato

dall'abbondante flusso informativo presente nella trascrizione. La condizione ST sembra invece aver stimolato maggiormente l'applicazione di strategie quali generalizzazioni e omissioni critiche.

4.2.6 Esempio di gestione strategica dell'input: allontanamento dello sguardo dagli stimoli

Sia per la condizione GM che per la condizione ST, in determinati punti del discorso le partecipanti hanno deciso volontariamente di allontanare lo sguardo dagli stimoli presenti nella trascrizione automatica di Google Meet o nel CAI *tool* di SmarTerp. Nel corso delle retrospezioni le partecipanti hanno fornito commenti e riflessioni in merito a questa scelta strategica.

In questa prima parte vengono presentati i commenti relativi ai motivi sottostanti all'allontanamento dello sguardo dagli stimoli nella condizione sperimentale contenente la trascrizione automatica di Google Meet. P1 afferma che ha dovuto allontanare lo sguardo dalla trascrizione a causa di un'elevata lacrimazione dell'occhio, dovuta principalmente all'eccessiva elaborazione del discorso affidata alla lettura della trascrizione. P3 ritiene di aver allontanato lo sguardo dagli stimoli nei momenti in cui fosse necessaria una riformulazione più profonda del testo e particolarmente in concomitanza di elenchi e parti altamente informative. P3 ritiene infatti che l'allontanamento dalla trascrizione favorisca una maggior concentrazione e concretizzazione del testo, oltre a una maggior libertà creativa nell'elaborazione dell'output, soprattutto nei punti in cui la partecipante ha dovuto mettere in atto strategie di sintesi del discorso dato il ritardo percepito. P3 chiarisce infatti che l'aumento della velocità d'eloquio dell'oratrice è stato spesso concomitante a un maggior allontanamento dello sguardo dalla trascrizione. Come accennato anche da P3, P6 e P7 ribadiscono che per le parole i cui traducenti non erano immediatamente disponibili, la partecipante ha preferito focalizzare il proprio sguardo sul labiale dell'oratrice e non sulla trascrizione. P7 ha infatti sottolineato che ha allontanato lo sguardo dalla trascrizione in particolare quando venivano presentate parole simili tra l'italiano e lo spagnolo e per le strutture dissimmetriche (ad esempio, per la struttura "desde el realismo") per cercare di mantenere uno stile naturale e idiomatico. P7 afferma inoltre di aver allontanato lo sguardo dalla trascrizione nei momenti in cui fosse necessario attivare un po' la creatività oppure quando fosse necessaria un'operazione di ripescaggio nella memoria. P3 ha affermato di aver allontanato lo sguardo dal testo scritto nelle parti più argomentative e meno

tecniche, in quanto l'input scritto in quelle parti risultava più confusionario. P5 afferma invece di essersi allontanata dalla trascrizione automatica nei punti in cui rimaneva più indietro, in quanto in quei punti la trascrizione non avrebbe permesso un'astrazione sufficiente del discorso da permettere una sintesi rapida ed efficace. P6, nel corso dell'attività di retrospezione, ha inoltre sottolineato che nei punti più densamente informativi è stato necessario chiudere gli occhi, in quanto se non avesse chiuso gli occhi non avrebbe potuto ascoltare l'input auditivo ed elaborarlo correttamente.

In questa seconda parte vengono presentati i commenti relativi ai motivi sottostanti all'allontanamento dello sguardo dagli stimoli nella condizione sperimentale contenente il CAI *tool* di SmarTerp. P1 afferma di aver chiuso gli occhi per visualizzare meglio i numeri nella parte più densamente informativa del testo, in quanto il numero elevato di input visivi rappresentava un ostacolo in quella parte. P1 ritiene infatti che in quella parte gli occhi chiusi aiutassero a visualizzare meglio il significato delle cose. Come osservato anche per la condizione GM, P3 ribadisce la necessità di allontanare lo sguardo dai suggerimenti del CAI nelle parti più discorsive del discorso. P7 ritiene di aver guardato fuori dallo schermo più frequentemente in questa condizione sperimentale rispetto alla condizione GM.

Riassumendo, per entrambe le condizioni molte partecipanti hanno allontanato lo sguardo dagli stimoli nelle parti più discorsive, potenziale sintomo dell'utilità dei suggerimenti nelle parti più informative. Inoltre, l'abbondanza dell'input visivo-verbale in ST ha portato le interpreti a guardare più frequentemente la trascrizione, pur non avendone necessità: questa tendenza sembra ribadire la necessità di una formazione che miri all'apprendimento di strategie di selezione informativa dinanzi a un input come quello della trascrizione automatica.

4.2.7 Impatto degli strumenti sulla prosodia dell'interprete

Seppur inaspettatamente, alcune partecipanti hanno individuato alcune correlazioni tra strumento utilizzato e andamento prosodico del proprio output. Ad esempio, per la condizione GM, P1, P3 e P11 hanno evidenziato alcuni fattori interessanti. P1 sostiene che la presenza della trascrizione automatica ha reso più robotica la propria resa rispetto a una resa basata sull'ascolto puro. P3 ritiene di aver perso la carica emotiva delle parti meno informativamente dense proprio a causa della trascrizione, che ha portato l'interprete a concentrarsi di più sulle parole che sull'obiettivo comunicativo. Anche P11 sottolinea la percezione di un andamento prosodico

robotico della propria resa a causa della densità informativa, che a suo avviso veniva amplificata dalla trascrizione generata automaticamente. Per la condizione ST, P2 ha sottolineato che la latenza del sistema (sovrapposta a un *décalage* molto corto) ha portato diverse volte la partecipante a interrompere la propria resa, portando a frasi prosodicamente frammentate.

Complessivamente, emerge chiaramente la necessità di aiutare le interpreti in formazione nel selezionare l'input visivo, in modo da poter acquisire una maggior sicurezza circa la fluidità della propria resa, focalizzandosi sull'obiettivo comunicativo e non esclusivamente sulla resa di quante più informazioni possibili.

4.2.8 Gestione dell'input visivo e dell'input auditivo

Per entrambe le condizioni sperimentali, le partecipanti hanno fornito riflessioni generali sulla gestione dell'input verbale-auditivo e quello verbale-visivo, cercando di sottolinearne strategie e difficoltà riscontrate.

In questa prima parte verranno sintetizzati i commenti relativi alla gestione dei due input nella condizione sperimentale contenente la trascrizione automatica di Google Meet. P10 afferma di aver cercato di non guardare la trascrizione, in quanto fonte inutile di distrazioni; P10 ritiene infatti di averla solo utilizzata per recuperare alcuni numeri persi nella fase di ascolto e per recuperare l'input testuale delle parti in cui rimaneva indietro. Anche P11 ribadisce di aver utilizzato l'input del canale visivo prevalentemente per i numeri, nonostante i numeri siano stati spesso difficili da elaborare e quindi fonte di confusione. P11 ritiene inoltre che l'input verbale-visivo abbia portato la partecipante a confondere la gerarchia semantica degli elementi presenti nel testo, perdendo talvolta il verbo della frase principale, proprio a causa della non capacità di ricollegare gli elementi presenti nell'input visivo con quelli provenienti dall'input auditivo. P11 ha tuttavia affermato di aver fatto ricorso all'input verbale-visivo nei momenti in cui non fossero chiari elementi presenti nell'input auditivo, ma solo successivamente alla constatazione di un'incomprensione. P3 e P6 ritengono di aver monitorato la resa nelle frasi più discorsive e idiomatiche con un movimento costante dell'occhio dall'oratrice alla trascrizione e viceversa (e quindi, passando da ascolto puro ad ascolto con trascrizione, e viceversa). Anche P7 ritiene di aver prevalentemente sfruttato la trascrizione solo nelle parti più densamente informative (ossia con un'elevata concentrazione di liste con numeri e unità di misura complesse). Non a caso, P3 sottolinea di aver guardato solo l'oratrice nelle parti più emotivamente cariche.

Curiosamente, P5 afferma di aver riposto il proprio sguardo sulla trascrizione (piuttosto che sull'oratrice) a causa dell'intonazione neutra dell'oratrice e a causa dell'assenza di elementi para-verbali significativi; aggiunge infatti di aver guardato l'oratrice solo durante le pause. P2 sottolinea di aver utilizzato la trascrizione automatica di Google Meet solo nei momenti in cui rimaneva indietro con il *décalage* e, quindi, quando aveva bisogno di recuperare rapidamente il materiale testuale, a discapito di una minor elaborazione per astrazione del discorso. P3 ritiene di aver preferito l'input auditivo puro (quindi, senza guardare la trascrizione) nei momenti in cui sono comparsi nel discorso dei lessemi la cui traduzione era assodata nella memoria a lungo termine della partecipante (ad esempio, "papel de liderazgo"): P3 ritiene infatti che l'input scritto avrebbe potuto trascinarla in calchi e sostiene che sia stato fondamentale in quei momenti non fare più affidamento sul testo. P1 afferma di aver seguito prevalentemente l'input della trascrizione, nonostante si sia resa conto che in alcuni punti seguire l'input scritto l'abbia portata a interrompere l'ascolto; infatti, P1 afferma di aver riattivato l'ascolto attivo durante l'IS solo nei momenti in cui l'assenza della punteggiatura non permetteva la ricostruzione della coesione morfosintattica dei sintagmi e la comprensione del senso della frase.

In questa seconda parte verranno sintetizzati i commenti relativi alla gestione dei due input nella condizione sperimentale contenente il CAI *tool* di SmarTerp. P1 ritiene che la gestione di un numero elevato di input abbia portato a una perdita totale di controllo su quello che stesse dicendo e aggiunge che ritiene di aver eseguito una miglior prestazione nei momenti in cui ha concentrato il proprio sguardo sull'oratrice senza soffermarsi sugli stimoli presenti nell'input visivo. P1 ritiene di essere stata particolarmente distratta da quegli stimoli che sono comparsi nel CAI di SmarTerp ma la cui traduzione era già nota alla partecipante: la ridondanza informativa nell'input scritto di vocaboli già presenti nella memoria a lungo termine della partecipante sembra aver avuto un impatto molto negativo nel corso di tutta la resa. P1 ribadisce a più riprese di aver abbandonato l'input visivo-verbale nei momenti di difficoltà, cercando di focalizzarsi solo sull'oratrice, e afferma di aver chiuso gli occhi nelle parti testuali più discorsive, per evitare qualsiasi tipo di interferenza. P1 dice che la gran quantità di input le ha tolto qualsiasi facoltà cognitiva, in quanto ha portato la partecipante a un sovraccarico di informazioni e di input, a suo avviso, troppo difficile da gestire. Anche P4 sostiene di aver riposto il proprio sguardo sull'oratrice piuttosto che sugli stimoli del CAI e dice di aver cominciato a guardare con maggior frequenza gli stimoli solo nella seconda parte, ma più per stanchezza che per necessità, andando ad alleggerire la comprensione di molti numeri. P4 dice di aver mosso più rapidamente lo sguardo in tutte le parti dello schermo nei momenti in cui sono comparsi i numeri nella parte finale del discorso; in particolare, la partecipante ritiene che

la difficoltà nell'associazione dei numeri al concetto cui facevano riferimento abbia influito negativamente sulla gestione ragionata e critica dei diversi input. Anche P5 afferma di aver mosso molto rapidamente lo sguardo in più parti del discorso a causa dell'elevata quantità di input presenti nell'interfaccia del CAI *tool* di SmarTerp, che hanno portato la partecipante a stancarsi molto e a un utilizzo poco ragionato e motivato dell'input visivo. P9 afferma di essere ricorsa all'input verbale-visivo del CAI anche per parole note all'interprete ma perse nella fase di ascolto del discorso originale (ad esempio, per la parola "hito"). P9 afferma di aver fatto ricorso all'input verbale-visivo prevalentemente per recuperare gli equivalenti di termini specialistici, pur chiarendo che i termini specialistici la cui traduzione era già nota all'interprete hanno provocato alcune distrazioni.

In sintesi, sembra emergere una chiara necessità didattica mirata alla gestione degli input presenti nel canale visivo. L'inesperienza nel gestire la trascrizione automatica, ad esempio, ha portato alcune interpreti a interrompere l'ascolto attivo e a confondere la gerarchia semantico-sintattica degli elementi, proprio a causa di una comprensione affidata alla trascrizione automatica. Nella condizione ST, la maggior parte delle interpreti hanno riportato una difficoltà significativa nel gestire la complessità degli input, che in molti casi ha generato confusione e inibizione delle proprie strategie.

4.2.9 Monitoraggio della resa

A causa della latenza di entrambi i sistemi di ASR e della difficoltà nello sfruttare a proprio vantaggio l'input visivo-verbale elasticizzando il proprio *décalage*, diverse partecipanti hanno chiarito nelle retrospezioni di aver utilizzato l'input visivo-verbale solo in fase di monitoraggio in alcune parti dei discorsi.

Per quanto riguarda la condizione GM, P3, P6, P8, P9 e P11 hanno sottolineato l'utilizzo della trascrizione generata automaticamente da Google Meet in fase di monitoraggio. P3 afferma di aver utilizzato la trascrizione solo dopo aver compreso il discorso originale e iniziato il proprio output; quindi, solo per avere un costante *feedback* circa l'andamento del proprio output: P3 ritiene che quest'utilizzo dello strumento non abbia scompensato troppo la propria resa. P6 afferma di aver avuto qualche difficoltà nel monitoraggio di alcuni numerali che sono comparsi e scomparsi più volte nella trascrizione. P6 ritiene che la trascrizione automatica permetta un monitoraggio neutrale della propria resa, in quanto il fatto che la trascrizione non fornisca gli

equivalenti in italiano porta la partecipante ad essere sicura del fatto che se commette errori non ne viene a conoscenza immediatamente nella fase di monitoraggio. P8 ritiene che sia stato molto utile monitorare le entità denominate e i nomi propri, in quanto invariati tra lo spagnolo e l'italiano. P9 sottolinea di aver monitorato la propria resa guardando gli elementi già interpretati per concentrarsi maggiormente sull'input del discorso non ancora interpretato. P11 afferma di aver sfruttato la trascrizione durante le pause dell'oratrice, per monitorare quanto detto prima di ciascuna pausa: P11 ritiene infatti che osservare l'input della trascrizione durante le pause sia stato fondamentale per assicurarsi che la trascrizione fosse accurata e completa; in questo modo la partecipante ha controllato di aver detto cose sensate prima della pausa rafforzando la sua fiducia sullo strumento utilizzato.

Per quanto riguarda la condizione ST, P1, P3, P6, P10 e P11 hanno invece commentato l'utilizzo del CAI *tool* di SmarTerp in fase di monitoraggio. P1 afferma che il fatto che gli stimoli suggeriti rimangano sullo schermo ha confuso la partecipante, portandola talvolta a chiudere gli occhi per la presenza di troppi stimoli da gestire anche in fase di monitoraggio. P3 ha detto di avere optato per soluzioni traduttive diverse rispetto a quelle suggerite dal CAI: P3 afferma infatti di aver osservato le soluzioni proposte dal CAI in fase di monitoraggio ma di non essere stata confusa dalla traduzione fornita dallo strumento. P6 osserva che la comparsa delle traduzioni sotto ad ogni stimolo individuato dal CAI ha portato la partecipante a una maggior ansia nella fase di monitoraggio, in quanto il CAI ha portato la partecipante a essere immediatamente consapevole degli errori traduttivi commessi. Al contempo, la stessa partecipante chiarisce che dovrebbe sapere gestire meglio i suggerimenti del CAI capendo quando riformularsi e quando invece andare avanti senza riformularsi, pur mantenendo scelte lessicali con lievi sfumature stilistiche e semantiche rispetto alle soluzioni proposte dal CAI. Anche P10 ribadisce la maggior ansia generata nella fase di monitoraggio dall'osservazione di stimoli tradotti dal CAI con soluzioni diverse (ma semanticamente simili) rispetto a quelle dell'interprete. P11 ritiene di aver guardato molti suggerimenti del CAI solo in fase di monitoraggio per vedere se avesse commesso errori e per avere una rassicurazione circa i traducenti utilizzati nel corso della resa; inoltre, P11 afferma di aver guardato alcune parole solo in fase di monitoraggio per poi magari utilizzare i traducenti proposti dal CAI nel caso in cui si fossero ripresentati nel corso del discorso. P11 chiarisce infatti che il monitoraggio attivo della resa è stato particolarmente stimolato dall'evidenziazione dei suggerimenti del CAI, che, data la latenza dello strumento, nella maggior parte dei casi sono comparsi solo quando l'output dell'interprete per quello stimolo era già iniziato.

Sia nella condizione GM che nella condizione ST, alcune partecipanti hanno affermato che l'input verbale-visivo dello strumento è stato fondamentale nella fase di monitoraggio per recuperare alcune informazioni che altrimenti sarebbero andate perdute. In particolare, per la condizione GM, P7 e P9 hanno osservato situazioni simili. P7 ha affermato di aver avuto un *décalage* troppo lungo che non le ha consentito di sfruttare sufficientemente la trascrizione automatica durante l'elaborazione dell'input e, pertanto, ha fatto ricorso alla trascrizione solo per il recupero sporadico di informazioni andando a tradurre alcune parti della trascrizione facendo brevi traduzioni a vista. P9 ha invece affermato di aver sfruttato la trascrizione di Google Meet per recuperare l'inizio di alcune frasi che, a causa della sovrapposizione con la fine della frase precedente, non era stato elaborato dall'interprete. P9 ha inoltre aggiunto che la trascrizione è stata utile per avere un rapido punto di riferimento per comprendere il contesto di utilizzo di alcune unità di misura poco note (in particolare, "megajoule"): la partecipante afferma infatti di essere stata molto aiutata dalla frase intera e di aver migliorato la comprensione della frase, nonostante i potenziali errori. Per la condizione ST, P3 e P6 hanno evidenziato alcune tendenze interessanti relativamente al recupero delle informazioni durante il monitoraggio del proprio output. In particolare, se per la condizione GM vengono prevalentemente recuperati elementi relativi all'organizzazione sintattica, sia P3 che P6 evidenziano che il CAI *tool* è stato utilizzato prevalentemente per il recupero di terminologia specialistica non immediatamente disponibile alle partecipanti, ma di cui già è stata compresa la sua funzione nell'organizzazione sintattica.

In sintesi, entrambi gli strumenti hanno avuto una funzione positiva nella fase di monitoraggio della resa, permettendo alle partecipanti di acquisire una maggior sicurezza circa il proprio andamento. Tuttavia, l'eccessiva difficoltà nell'elasticizzare il proprio *décalage* nella condizione ST ha portato spesso le partecipanti a utilizzare il CAI prevalentemente in fase di monitoraggio senza una selezione critica dei suggerimenti. È pertanto necessaria una didattica che permetta un monitoraggio critico e selettivo, in cui l'interprete vada a monitorare solo quelle informazioni di cui ha bisogno.

4.2.10 CAI *tool* di SmarTerp: tra vantaggi e svantaggi

In questa sezione verranno presentati alcuni commenti relativi specificatamente alla condizione sperimentale contenente il CAI *tool* di SmarTerp. Le partecipanti hanno infatti sottolineato

principalmente due aspetti nel corso delle indagini retrospettive che riguardano questo strumento: da una parte hanno evidenziato l'impatto dell'organizzazione dell'input visivo di SmarTerp e dall'altra parte hanno posto particolare enfasi sulla ridondanza informativa percepita come talvolta fonte di distrazione. Per quanto riguarda l'impatto dell'organizzazione dell'input visivo dello strumento sulla prestazione, P1, P2, P8, P9 e P11 hanno condiviso alcune riflessioni particolarmente interessanti. P1 ha affermato che la colonna contenente le entità denominate si è rivelata piuttosto inutile nel corso della sua simultanea e ha invece affermato di aver nettamente preferito le altre due colonne; in quanto, se ha avuto dubbi, questi erano prevalentemente relativi alla terminologia specialistica o i numeri. P1 ha infatti sottolineato che la colonna relativa alle entità denominate confonde molto in quanto molto spesso propone parole già note all'interprete (come, ad esempio, "Governo", "Spagna" o "istituto"). Anche P2 osserva di aver piuttosto ignorato la colonna relativa alle entità denominate e ritiene che la colonna più importante sia stata quella relativa alla terminologia specialistica, più utile anche rispetto alla colonna relativa ai numerali. Anche P9 afferma di aver guardato solo sporadicamente la colonna relativa alle entità denominate e dice che nei casi in cui l'ha guardata non ha fatto affidamento su ciò che ci fosse scritto; tuttavia, ritiene di essersi fidata prevalentemente dei numerali (nonostante la latenza) e della terminologia specialistica. P8 ha invece sottolineato che molto spesso l'evidenziazione automatica dei suggerimenti ha portato l'interprete a guardare automaticamente la maggior parte dei suggerimenti, ma afferma che talvolta l'evidenziazione automatica l'ha distratta dall'ascolto del discorso originale. Anche P11 ribadisce di essere stata portata a guardare la colonna relativa alle entità denominate dall'evidenziazione automatica, nonostante in molti casi non abbia letto il suggerimento, in quanto non necessario.

Per quanto riguarda la percezione della ridondanza informativa, P1, P5, P6 e P11 hanno riscontrato un'elevata ridondanza informativa nei suggerimenti del CAI. P1 ritiene che il fatto che lo strumento proponeva più volte lo stesso termine nel corso del discorso (ad esempio, "bilancio") sia stato un elemento di distrazione nel corso di tutta la *performance*. P1 dice infatti che in una situazione ideale sarebbe perfetto se comparissero solo termini di cui non sa la traduzione, in quanto la ridondanza informativa nell'input visivo-verbale di elementi noti nell'input auditivo-verbale l'ha confusa nella fase di comprensione del discorso. Anche P5 ribadisce che la comparsa nell'input visivo-verbale di parole già consolidate nella memoria a lungo termine abbia distratto la partecipante. P6 critica invece la proposta di più traduzioni da parte del CAI per uno stesso termine specialistico, portando l'esempio di "empleo cualificado" che è stato tradotto sia come "occupazione" che come "posti di lavoro qualificati": in questo

caso la ridondanza informativa data dalla proposta di più suggerimenti per uno stesso sintagma nominale ha portato la partecipante a riformularsi più volte, pur non avendone necessità. P11 ribadisce il problema riscontrato da P6 portando l'esempio di "plan de recuperación y resiliencia", che è stato tradotto dal CAI, anche in questo caso, con due traducenti.

Riassumendo, è necessario il rafforzamento delle capacità di gestione della complessità dell'input visivo presente nel CAI di SmarTerp proprio durante la formazione delle interpreti ed è necessario al contempo preparare le interpreti alla possibilità di ridondanza informativa nelle tecnologie per l'interpretazione.

4.2.11 Trascrizione automatica di Google Meet: tra vantaggi e svantaggi

In questa sezione verranno presentati alcuni commenti relativi specificatamente alla condizione sperimentale contenente la trascrizione automatica di Google Meet. Nell'analisi delle retrospezioni emergono due elementi particolarmente rilevanti relativamente a questa condizione sperimentale: da un lato alcune partecipanti sottolineano alcuni fattori relativi alla punteggiatura di Google Meet e dall'altro lato alcune partecipanti affermano di aver utilizzato la trascrizione generata automaticamente facendo una traduzione a vista di gran parte del discorso (come del resto è emerso anche nei dati di *eye-tracking*).

Per quanto riguarda la punteggiatura della trascrizione, P1 ritiene di essere stata tendenzialmente aiutata dalla presenza di una punteggiatura piuttosto affidabile per gran parte del discorso; tuttavia, la stessa partecipante afferma che la punteggiatura è venuta meno nei punti in cui la velocità d'eloquio è aumentata. Curiosamente, P1 afferma che il fatto che in alcuni punti non ci fosse più la punteggiatura le ha permesso di porre più attenzione sul canale auditivo, rafforzando la concentrazione e alleggerendo l'elaborazione della trascrizione. P2 ritiene invece che la mancanza di punteggiatura sia stata particolarmente problematica, in quanto in alcuni casi l'ha portata a non capire quale fosse il soggetto della frase (probabile sintomo di una comprensione basata quasi esclusivamente sul canale visivo-verbale e non prevalentemente sul canale auditivo). Secondo P2, l'assenza di punteggiatura ha avuto un impatto negativo anche nella resa di elenchi e nella segmentazione dei periodi (e, quindi, nel capire dove finisse una frase e ne iniziasse un'altra).

Per quanto riguarda la tendenza a eseguire una traduzione a vista della trascrizione automatica, questa è stata ribadita più volta da molte partecipanti e confermata dai dati di *eye-tracking*. Di seguito verranno riportate alcune riflessioni interessanti. P1 dice di esercitarsi molto con la traduzione a vista e ritiene che sia più facile per lei fare la simultanea con il testo sotto, pur mantenendo sempre un *décalage* molto corto. P1 ritiene inoltre che fare la traduzione a vista sia particolarmente utile per le strutture dissimmetriche tra lo spagnolo e l'italiano, in quanto, a suo avviso, la traduzione a vista permette di andare avanti con lo sguardo, potendo cogliere rapidamente il contesto sintattico-semantico della struttura. Anche P3 ritiene di aver fatto una traduzione a vista per gran parte del discorso, in quanto secondo lei è di grande aiuto per mantenere un ritmo sostenuto dell'output e per mantenere un'elevata informatività nella resa. P4 ritiene, curiosamente, che la traduzione a vista l'abbia aiutata nel riformulare il discorso e ritiene che se avesse soltanto ascoltato avrebbe tralasciato una maggior quantità di informazioni. P8 afferma invece che l'aver fatto una traduzione a vista per la maggior parte del discorso ha avuto un impatto negativo relativamente ai calchi, giungendo in alcune occasioni a pronunciare parole italiane con una pronuncia spagnola (ad esempio, "esensiale", anziché "essenziale").

Riassumendo, l'attenzione riposta da molte interpreti sulla punteggiatura della trascrizione automatica mette in luce la tendenza di molte interpreti a non fare affidamento sul canale auditivo per l'elaborazione più profonda del discorso e, in questo senso, è necessaria una didattica che sottolinei l'importanza dell'interprete quale soggetto pensante che elabora la semantica di un discorso, solo in parte reso dalla macchina. Inoltre, è necessario sottolineare nella formazione i pericoli potenzialmente comportati dalla tendenza a eseguire la traduzione a vista di una trascrizione automatica.

4.2.12 Trascrizione automatica di Google Meet e CAI *tool* di SmarTerp a confronto

P1, P2, P3, P4, P6, P7, P8 e P11 hanno proposto alcune riflessioni in cui hanno messo a confronto gli strumenti proposti per le due condizioni sperimentali. P1 ha affermato di preferire nettamente la trascrizione automatica di Google Meet, in quanto, a suo avviso, nonostante spesso induca in calchi, le consente di svolgere una prestazione meno cognitivamente impegnativa. P1 ritiene tuttavia che non ci siano state differenze significative nella gestione dei numeri e afferma che, per quanto riguarda i numeri, SmarTerp permette una gestione

leggermente più efficiente, in quanto propone una struttura dei numerali in cui non viene presentata la struttura dissimmetrica "mil millones" per i miliardi. P2 ritiene invece di essersi trovata meglio con Google Meet per i numeri, proprio perché aiutata dalla struttura "mil millones" presente nella trascrizione dei miliardi, in quanto automatizzata e semplice. P2 dice di aver nettamente preferito Google Meet in generale, nonostante la mancanza della punteggiatura in alcune parti. Anche P3 afferma di preferire Google Meet, in quanto la struttura dei numerali è stata ritenuta poco efficiente nel caso in cui si sono presentati numerali con un elevato numero di zeri. P6 dice di essersi fidata più di SmarTerp che di Google Meet e aggiunge di aver chiuso gli occhi nel 90 per cento dei casi in cui questi sono comparsi nella trascrizione automatica di Google Meet, in quanto la trascrizione distraeva nella fase di ascolto e comprensione per l'abbondanza del flusso informativo nell'input visivo-verbale. P7 percepisce invece una maggior latenza del CAI di SmarTerp rispetto all'ASR di Google Meet e, ciononostante, ritiene che SmarTerp sia stato più accurato nella resa dei numeri. P8 ritiene di aver preferito la trascrizione automatica di Google Meet, in quanto questa le ha permesso di svolgere una traduzione a vista di porzioni significative del discorso, potendo seguire maggiormente il filo tematico del discorso, data la familiarità con la tecnica della traduzione a vista. P11 afferma di preferire SmarTerp, poiché a suo avviso si tratta di uno strumento che permette di focalizzarsi sull'ascolto dell'oratrice e di gestire più autonomamente la comprensione del discorso. P11 afferma infatti di aver letto meno suggerimenti presenti nel CAI *tool* di SmarTerp rispetto alla trascrizione di Google Meet, in quanto si è sentita meno obbligata a leggere tutti gli stimoli presenti nell'input visivo-verbale.

In sintesi, nonostante la maggior difficoltà nell'input visivo, diverse partecipanti hanno evidenziato l'utilità di SmarTerp nel concentrarsi sull'input auditivo e nello stimolare le strategie interpretative.

4.2.13 Percezione dell'usabilità degli strumenti

Oltre ai dati quantitativi raccolti relativamente alla percezione dell'usabilità della trascrizione automatica di Google Meet e del CAI di SmarTerp, alcune partecipanti hanno condiviso alcune ulteriori riflessioni in merito. Per quanto riguarda la percezione della trascrizione di Google Meet, P3, P6, P8 e P9 hanno condiviso alcuni commenti interessanti. P3 ritiene che la trascrizione sia molto utile, ma al contempo ribadisce che l'assenza di punteggiatura e le sigle

trascritte con lettere minuscole abbiano avuto un impatto negativo sulla resa. P3 ritiene complessivamente lo strumento di supporto e di aiuto all'interprete, ma è consapevole di non averlo saputo utilizzare in un modo efficiente. Anche P6 ritiene che la trascrizione di Google Meet sia di supporto e aggiunge che secondo lei la trascrizione permette un monitoraggio attivo della resa in cui l'interprete non acquisisce una consapevolezza immediata dei propri errori, in quanto ha solo un riscontro circa il testo originale. P8 ritiene invece che la trascrizione abbia complessivamente ostacolato la resa e chiarisce che pensa di aver migliorato il proprio output solo dal momento in cui ha deciso di non fidarsi di Google Meet, affidandosi solo all'ascolto puro dell'oratrice. P9 dice di aver percepito lo strumento come di supporto nei punti del testo più informativamente densi oppure quando è rimasta indietro, ma afferma di aver percepito la trascrizione come di ostacolo nelle parti del discorso in cui c'erano frasi chiare.

Per quanto riguarda la percezione del CAI *tool* di SmarTerp, P1 sottolinea come lo strumento sia stato particolarmente d'ostacolo e poco efficiente per la sua intrinseca ridondanza informativa: il fatto che il programma ripetesse parole già presentate più volte, ad esempio, è stato percepito come un elemento di distrazione. P1 aggiunge che la molteplicità degli input nella condizione ST abbia portato l'interprete a perdere il controllo sul proprio output. Da queste opinioni emerge con chiarezza la necessità di una prospettiva didattica che permetta all'interprete un alleggerimento del carico cognitivo durante l'IS traendo vantaggio da questo strumento.

4.2.14 Prospettiva didattica

Giungiamo adesso alla ragion d'essere sottostante al presente elaborato: la prospettiva didattica. Per mettere in luce alcuni aspetti, verranno ripercorse le necessità didattiche emerse nel corso delle attività di retrospezione e relative al confronto di interpreti in formazione con le nuove tecnologie. P1 ritiene di aver avuto una forte tendenza a eseguire una traduzione a vista della trascrizione automatica nella condizione GM, pur non controllando l'intero andamento semantico del discorso. Ritiene infatti di aver chiuso alcune frasi senza l'adozione di strategie consapevoli, a causa della mancata comprensione di porzioni testuali delle quali stava eseguendo una traduzione a vista superficiale e poco selettiva. Per la condizione ST, la stessa partecipante ha affermato di non essere invece stata in grado di gestire tutti gli input del CAI, principalmente per la costante tentazione di utilizzare tutti i suggerimenti senza invece

focalizzarsi solo su quelli di cui avesse bisogno. P1 ritiene infatti che la gestione di un input complesso, come quello di SmarTerp, debba essere affrontata durante la formazione delle interpreti; tuttavia, ritiene che questo argomento debba essere affrontato o all'inizio della formazione delle interpreti oppure alla fine, ossia quando l'interprete ha già consolidato le tecniche di simultanea, ma non in una fase intermedia. Anche per la condizione GM, P3 ha sottolineato la difficoltà nella gestione di una molteplicità di input (trascrizione in movimento e input auditivo) che si è sovrapposta a una velocità d'eloquio talvolta non pienamente gestibile dalla partecipante. La partecipante afferma infatti di dover ancora allenarsi su velocità più sostenute e che la stanchezza dovuta alla velocità ha avuto un impatto sull'intera gestione degli input visivi. Per quanto riguarda P3, ritiene di non essere ancora pronta per cogliere l'utilità di SmarTerp, che presenta un input visivo non familiare. P4 riscontra una propria difficoltà nella gestione di due input visivi e afferma di interpretare spesso con gli occhi chiusi oppure guardando il muro, probabile sintomo di una tecnica di interpretazione simultanea ancora in fase di formazione. P5 ritiene che un'attività propedeutica utile per la gestione di più input durante l'IS sia lo svolgimento di simultanee con il testo, con buona punteggiatura e fedele al discorso originale. P10 dice invece di avere un *décalage* troppo corto e poco elastico per poter sfruttare al meglio gli strumenti proposti durante le due IS. Anche P11 ribadisce che il suo *décalage* troppo corto non le ha permesso di apprezzare i benefici del CAI *tool* di SmarTerp: la partecipante afferma infatti di aver avuto un *décalage* estremamente ravvicinato a causa della velocità d'eloquio e sostiene di poter sfruttare a pieno il CAI di SmarTerp solo quando riuscirà a elasticizzare il proprio *décalage*.

4.2.15 Aspettative

Sia per la condizione GM che per la condizione ST, nel corso delle retrospezioni le partecipanti hanno osservato che le diverse aspettative circa l'output visivo-verbale degli strumenti hanno avuto un impatto talora negativo sulla percezione della loro usabilità.

Per quanto riguarda la trascrizione automatica di Google Meet, P1, P6, P10 e P11 hanno sottolineato alcuni elementi interessanti. P1 si attendeva una diversa prestazione dello strumento soprattutto nelle parti del discorso in cui la velocità d'eloquio è aumentata, in quanto secondo la partecipante in quei momenti la trascrizione automatica, anziché essere di supporto, ha presentato una frequenza più elevata di errori. P6 ritiene invece estremamente rischiosa la

presenza di errori di trascrizione (ad esempio, “resto” anziché “reto”, “gigabyte” anziché “gigawatt”), in quanto hanno portato la partecipante a uno stress dovuto all’incongruenza tra input auditivo e visivo-verbale. P10 afferma di essere stata messa in difficoltà prevalentemente a causa della discontinuità della punteggiatura, che in alcune parti del discorso ha organizzato il discorso e in altre ha confuso l’interprete a causa della sua mancanza. Anche P11 ha ribadito di essere stata messa in difficoltà da alcuni errori nella trascrizione che hanno portato a un’elevata incongruenza tra input auditivo e visivo-verbale (ad esempio, “que ven la influencia de…” è stato trascritto con “que vende influencia”). P11 ritiene inoltre di avere trovato particolarmente problematica la trascrizione dei numeri, che si aspettava fosse utile ma molte volte si è rivelata non chiarificatrice.

Per quanto riguarda il CAI *tool* di SmarTerp, P2, P6, P7, P8, P10 e P11 hanno condiviso alcune riflessioni in merito ai momenti in cui il sistema le ha messe in difficoltà. P2 ha sottolineato di avere trovato particolarmente difficile da gestire la colonna relativa alle entità denominate, in quanto molto spesso le sigle (come “MRR”, “AEH2”) non sono state suggerite dal sistema. Anche P6 si attendeva la comparsa dei nomi interi delle leggi menzionate e delle sigle (come “CMNUCC”) nell’interfaccia del CAI. Anche P7 e P11 hanno affermato di aspettarsi la sigla “CMNUCC” nell’input visivo-verbale e di essere state portate a non fidarsi più del sistema, dopo aver preso coscienza della mancanza di molti elementi potenzialmente utili. P8 e P10 hanno invece affermato che non si aspettavano alcune omissioni del sistema relative alla terminologia specialistica (ad esempio, del lessema “calificación energética” o “circulación rodada”). P10 ha inoltre detto di essere stata delusa dalla mancanza del nome proprio “Francisco José Contreras” nell’output del CAI e quest’omissione ha portato l’interprete a capire prima tutto il discorso e a consultare il CAI solo successivamente, quindi nella fase di monitoraggio dell’output. P11 ritiene invece di essere stata confusa dall’organizzazione spaziale dell’input, in quanto si aspettava alcuni suggerimenti nella colonna relativa alle entità denominate, ma sono invece comparsi nella colonna riguardante la terminologia specialistica (ad esempio, questo è stato il caso di “stati membri dell’Unione Europea”).

In sintesi, è necessario rendere consapevoli le interpreti in formazione circa i limiti degli strumenti potenzialmente applicabili in contesti di IS reali, per far sì che eventuali inaccuratezze e omissioni dei sistemi di ASR non comportino un maggior carico cognitivo durante l’incarico.

5. Conclusioni, limiti e prospettive future

Questo studio è mosso dall'esigenza di offrire una riflessione più ampia relativamente al punto di contatto tra intelligenza artificiale e interpreti in formazione. In particolare, nella ricerca si è desiderato mettere in luce le differenze nell'usabilità di due sistemi che potrebbero essere utilizzati in una condizione realistica di IS (il CAI *tool* di SmarTerp e la trascrizione automatica di Google Meet), cercando di evidenziarne la diversa utilità e i possibili impatti sull'accuratezza e la fluidità delle rese di interpreti in formazione. A questo proposito è necessario ribadire che i sottotitoli di Google Meet non sono stati sviluppati per assistere l'interprete durante l'IS, ma rappresentano uno strumento che potrebbe essere utilizzato in questo contesto come CAI *tool*. La ricerca si è dunque sviluppata sui seguenti quesiti: in che modo questi sistemi influiscono sull'accuratezza della resa di interpreti in formazione? Quanto risultano (o non) soggettivamente efficienti e soddisfacenti? In che modo vengono utilizzati durante l'IS? Quali prospettive didattiche emergono dalle opinioni di interpreti in formazione? Qual è l'impatto di queste tecnologie sulla distribuzione dell'attenzione sull'input visivo?

Per rispondere a queste domande di ricerca, nel primo capitolo, si è partiti da una panoramica sulle tecnologie utilizzate durante l'interpretazione simultanea (o CAI *tool*), ponendo particolare enfasi sulla necessità di progettare e sviluppare CAI *tool* che favoriscano un rapporto uomo-macchina sempre più ergonomico (§1.1.2). Si è poi proposto un breve *excursus* teorico in cui sono stati descritti alcuni modelli cognitivi che mirano a illustrare la complessità dei processi cognitivi coinvolti durante l'IS (§ 1.2), con particolare enfasi sull'applicazione del modello del carico cognitivo di Seeber (2011) (§ 1.2.2) in contesti di IS con testo o trascrizione automatica (§1.2.2.1) e di IS con CAI (§1.2.2.2). Successivamente, sono stati forniti alcuni elementi teorici di base riguardanti i sistemi di riconoscimento vocale automatico, sottolineando la loro integrazione nei CAI *tool* utilizzati durante l'IS (§1.3.3) ed evidenziandone possibili limiti e sfide (1.3.4). In seguito, sono stati descritti i dettagli delle due tecnologie utilizzate nello studio sperimentale: sono stati prima descritti i modelli acustici (§1.4.1), i modelli di linguaggio (§1.4.2), il sistema di interpretazione semantica (§1.4.3) e l'architettura di base (1.4.4) di SmarTerp e sono poi stati trattati alcuni elementi della trascrizione automatica di Google Meet ritenuti rilevanti ai fini dello studio (§1.5.1). È stata poi proposta la trattazione degli elementi caratterizzanti due tecnologie di interpretazione automatica che ben sintetizzano l'esponenziale ricerca di sistemi *speech-to-speech* che mirano alla sostituzione dell'interprete (§1.6.1; §1.6.2). Per fornire alle lettrici dell'elaborato alcune chiavi di lettura circa i dati di *eye-tracking*, è stato

successivamente proposto un breve riassunto relativamente all'utilizzo della tecnologia di *eye-tracking* negli *interpreting studies* (§1.7) in cui sono state proposte alcune definizioni delle unità di analisi utilizzate nello studio. Successivamente, dato il campione coinvolto (di sole interpreti in formazione), è stata proposta una breve riflessione relativamente alla didattica delle tecnologie nella formazione delle interpreti (§1.8), in cui si sottolinea la necessità della presenza della competenza tecnologica come componente cruciale nei *curricula* di formazione delle interpreti. Inoltre, considerato che lo studio è stato applicato sulla combinazione linguistica italiano-spagnolo, è stata proposta una trattazione della particolarità di questa combinazione (§1.9), con particolare enfasi sull'interferenza delle strutture dissimmetriche durante l'IS tra queste due lingue e sulla necessità di anticipare strategicamente il discorso per risolvere tali dissimmetrie linguistiche.

Una volta poste le necessarie fondamenta teoriche, si è proceduto con lo studio sperimentale vero e proprio, i cui dettagli metodologici sono stati descritti nell'intero secondo capitolo. Dopo aver chiarito gli obiettivi e i quesiti di ricerca (§2.1), è stato descritto il *design* quasi-sperimentale dello studio (§2.3). Infatti, ad ogni partecipante sono state proposte due prove di interpretazione simultanea, nel corso delle quali è stato tracciato il movimento oculare tramite *eye tracker*, per valutare la gestione dell'input visivo. Dopo ogni prova, ha avuto luogo un'attività di retrospezione (con un *retrieval cue* costituito dal discorso originale sovrapposto al movimento dell'occhio della partecipante durante la prova). Al termine di questa attività è stato proposto un questionario UEQ (Laugwitz et al., 2008) per avere un dato quantitativo circa la percezione dell'usabilità dello strumento. Seguendo i principi metodologici avanzati da Frittella (2023) per l'inserimento strategico degli stimoli all'interno del discorso, ogni discorso conteneva 17 *task* (di complessità variabile) inseriti strategicamente (§2.4.1) e ogni discorso è stato preregistrato in modo da garantire l'uniformità delle condizioni per tutte le partecipanti (§2.4.2). Dopo aver reclutato 11 partecipanti (§2.5.1), è stata loro erogata una formazione asincrona (§2.5.2) suddivisa in due sessioni e in ogni sessione le partecipanti hanno ricevuto alcuni glossari e video per esercitarsi con la trascrizione automatica di Google Meet e con il CAI *tool* di SmarTerp, in modo da acquisire una maggior familiarità con queste tecnologie e con il lessico relativo al Ministero spagnolo per la transizione ecologica e la sfida demografica. Terminata la formazione, ogni partecipante ha svolto l'attività sperimentale. Per tutte le partecipanti è stato seguito uno stesso protocollo sperimentale (§2.6), opportunamente pilotato su una dodicesima partecipante fuori dal campione analizzato (§2.8).

Terminate tutte le batterie di esperimenti, i dati sono stati preparati (§2.7), in modo da organizzare l'analisi. Per quantificare il tasso di successo nella resa dei *task* inseriti nei discorsi (o *task success rate*), per ogni *task* è stato calcolato il valore medio, dato dalla media dei *success rate* raggiunti in ogni singolo stimolo, seguendo le basi metodologiche proposte da Frittella (2023). Per avere invece una valutazione globale dell'accuratezza delle rese, seguendo le scale proposte da Tiselius (2009), ogni resa è stata suddivisa in 35 *cluster* e per ogni *cluster* sono stati determinati due punteggi: uno relativo alla sua intelligibilità e uno relativo alla sua informatività rispetto al discorso originale (§2.7.2.2). Per quel che riguarda l'analisi dei questionari riguardanti la percezione dell'usabilità, i dati sono stati inseriti su un file .xlsx per poter realizzare le opportune analisi statistiche del confronto tra le due tecnologie (§2.7.3). Le retrospezioni sono state invece trascritte e importate sul *software* Taguette e dall'analisi tematica sono emerse 21 categorie di analisi. Per quanto riguarda i dati ottenuti tramite il dispositivo di *eye-tracking* (§2.7.4), questi sono stati analizzati in due modalità: da una parte è stato osservato il *dwell time* in aree di interesse predefinite per ogni strumento (§2.7.4.2) e, dall'altra parte, è stato osservato per ogni partecipante se 51 stimoli venivano osservati in fase di comprensione o in fase di monitoraggio della resa, ossia se venivano osservati prima o dopo aver iniziato l'output (§.2.7.4.1).

Alla luce della metodologia utilizzata e dell'inquadramento teorico sottostante, si giunge ora a trarre le conclusioni del nostro lavoro, sulla base dei dati emersi.

Osservando l'analisi dei dati relativi ai *task success rate*, si osserva che per i *task* a media complessità la condizione con SmarTerp (o ST) sembra aver aiutato maggiormente le interpreti nella resa degli stimoli contenenti numerali, entità denominate e tecnicismi isolati. In particolare, nonostante l'esiguità del campione coinvolto, sembra essere interessante la resa dei *task* contenenti strutture dissimmetriche isolate, che hanno avuto rese più accurate nella condizione contenente il CAI di SmarTerp. Anche per i *task* a media complessità, sembra rafforzarsi la tendenza riscontrata nei *task* di bassa complessità, in quanto gli stimoli contenenti strutture dissimmetriche (superficiali e profonde) e quelli contenenti referenti complessi con numerali hanno avuto una resa migliore nella condizione ST. Tuttavia, nonostante i risultati ottenuti per i *task* di media e bassa complessità, che sembrano portare a risultati percentuali più elevati per la condizione ST, nell'analisi comparativa per *task* di elevata complessità è stato osservato un apparente miglior andamento per quasi tutti i *task* proposti nella condizione contenente la trascrizione automatica di Google Meet (o GM). È pertanto ipotizzabile, sebbene i risultati necessitino di un ampliamento del campione, che l'abbondante flusso informativo dei

task ad elevata complessità sia stato gestito meglio nella condizione GM. Nonostante il miglior andamento in GM, curiosamente sono stati raggiunti valori più elevati in ST per quei *task* ad elevata complessità che contenevano strutture dissimmetriche superficiali o profonde, probabile sintomo di una maggior capacità di elaborazione testuale nella condizione ST. Seguendo la procedura del test dei ranghi con segno di Wilcoxon, è stato condotto un test statistico di tipo non parametrico per valutare la significatività statistica delle differenze relative alle medie ottenute nelle due condizioni sperimentali. I risultati del test sembrano suggerire che le differenze sono statisticamente significative e il CAI *tool* di SmarTerp sembra avere un impatto significativo sui risultati misurati relativamente ai *task success rate*. Tali risultati sono stati rafforzati dall'applicazione della correzione di Bonferroni (Dunn, 1961).

Terminata l'analisi dei *task success rate*, per una valutazione più globale della *performance*, sono state valutate anche l'intelligibilità e l'informatività delle rese. Le rese sono risultate mediamente più intelligibili nella condizione ST, con una differenza statisticamente rilevante rispetto ai dati ottenuti per la condizione GM, come emerso dal test di Wilcoxon rafforzato mediante l'applicazione della correzione di Bonferroni. Questo dato sembra suggerire, nonostante il campione limitato, che il CAI *tool* di SmarTerp favorisce un maggior controllo delle interpreti sul proprio output, con un'elaborazione del discorso che permette una miglior fluidità e comprensibilità della resa. Nella valutazione dell'informatività media delle rese, sembra invece emergere una differenza tra i valori ottenuti tra GM e ST non statisticamente significativa. Questi risultati sembrano suggerire che entrambi gli strumenti possono essere un aiuto, ma è necessario insegnare come usarli.

Conclusa l'analisi dell'intelligibilità e dell'informatività delle rese, sono stati analizzati i risultati dell'*User Experience Questionnaire.* Anche se un questionario relativo alla percezione dell'usabilità di un prodotto richiederebbe un campione ben più ampio rispetto a quello coinvolto, sembrano emergere alcuni dati significativi. Infatti, nonostante il miglior andamento generale nella valutazione dell'accuratezza delle rese per la condizione ST, il CAI *tool* di SmarTerp è risultato meno attraente, apprendibile, efficiente, controllabile, stimolante e originale rispetto alla trascrizione automatica di Google Meet (con valori negativi per le scale dell'efficienza e della controllabilità). Paragonando i dati di interpreti in formazione (raccolti in questo studio) per le categorie dell'apprendibilità, dell'efficienza e della controllabilità con quelli raccolti da Frittella (2023) relativamente a interpreti professionisti, sembra emergere che le interpreti professioniste ritengono il CAI di SmarTerp molto più efficiente e controllabile. Infatti, pur riportando un valore molto simile per la scala dell'apprendibilità, le interpreti

professioniste sembrano in grado di adattarsi più facilmente al nuovo strumento, grazie alla maggior esperienza.

Una volta completata l'analisi dei dati dell'UEQ, sono stati analizzati i dati di *eye-tracking*. Per quanto riguarda l'analisi quali-quantitativa, si osserva che nella condizione ST sono stati più frequenti i casi di dati non disponibili, probabile sintomo di una maggior frequenza di situazioni in cui le stesse partecipanti hanno allontanato lo sguardo dallo schermo. Inoltre, in un numero maggiore di casi rispetto alla condizione GM, le partecipanti non hanno guardato molti stimoli nella condizione ST, possibile segnale di una maggior capacità critica nel corso dell'IS da parte delle partecipanti per decidere circa l'utilità e la necessità di guardare alcuni stimoli. Nei casi in cui le partecipanti hanno guardato gli stimoli dopo aver iniziato la traduzione, sembra emergere che ci siano state più riformulazioni con peggioramento dell'output nella condizione GM e, parallelamente, si osserva che in un numero maggiore di casi ci sono state riformulazioni con miglioramento della propria resa nella condizione ST. Inoltre, in un maggior numero di casi, le partecipanti hanno guardato lo stimolo dopo aver iniziato la traduzione e senza riformularsi nella condizione ST, principalmente a causa del ridotto *décalage* che non ha permesso in molti casi di attendere i suggerimenti del sistema (che hanno avuto una latenza maggiore nel CAI *tool* di SmarTerp). Con una differenza di quasi il doppio, si osserva invece che in un numero superiore di casi sono stati guardati gli stimoli prima di iniziare l'output nella condizione GM: è infatti stata osservata una tendenza a eseguire una vera e propria traduzione a vista, poco selettiva e critica, probabile indice di un minor sforzo di elaborazione semantica nella condizione GM.

Sono stati poi analizzati i dati relativi al tempo trascorso dalle singole partecipanti sulle AOI preconfigurate, per valutare come gli strumenti venissero utilizzati. Pur consapevoli della limitatezza del campione, osserviamo che l'area di interesse contenente l'oratrice su sfondo bianco ha riportato un *dwell time* nettamente superiore per la condizione ST, mentre l'AOI contenente i suggerimenti ha riportato un tempo trascorso medio più elevato nella condizione GM. Pur non avendo indagato la correlazione statistica tra questi dati e quelli relativi all'accuratezza delle rese, sembra possibile affermare che il maggior allontanamento dall'input visivo-scritto nella condizione ST abbia favorito un maggior distaccamento dalla struttura del testo originale, portando inevitabilmente l'interprete a una maggior concentrazione nella fase di comprensione del testo originale, non potendo ricorrere a una traduzione a vista. Osservando i dati relativi alle tre colonne contenenti i suggerimenti nella condizione ST, si osserva che la

colonna più utilizzata è stata quella dei termini specialistici e la colonna meno utilizzata è stata quella contenente le entità denominate.

Per quanto riguarda le retrospezioni, nonostante le diverse categorie emerse siano di difficile generalizzazione in una trattazione sintetica, sono emersi alcuni elementi ribaditi da più partecipanti. In particolare, molte partecipanti sono d'accordo nel ribadire la maggiore utilità del CAI *tool* di SmarTerp rispetto alla trascrizione automatica di Google Meet nella resa di entità denominate e terminologia specialistica, in quanto lo strumento propone direttamente i traducenti e non si limita ad aiutare l'interprete nella fase di comprensione del discorso. Tuttavia, al contempo, molte interpreti hanno sottolineato l'impatto negativo della latenza dell'ASR di SmarTerp, che in molti casi non ha permesso alle interpreti di utilizzare lo strumento, a causa della percezione di un ritardo notevole nella resa dei suggerimenti da parte di SmarTerp. Tuttavia, proprio in queste opinioni emerge la necessità di una formazione delle interpreti che le aiuti a rendere più elastico il proprio *décalage* (spesso troppo corto) in modo da poter sfruttare a proprio vantaggio quei suggerimenti che giungono con la loro inevitabile latenza. Molte partecipanti hanno infatti evidenziato di aver utilizzato il CAI di SmarTerp prevalentemente in fase di monitoraggio della resa. Molte studentesse hanno inoltre sottolineato che la difficoltà è stata in molti casi legata ad un *décalage* poco flessibile. Inoltre, alcune partecipanti hanno sottolineato la maggior possibilità di ricorrere a strategie di riformulazione e sintesi nella condizione ST, in quanto ritengono che la trascrizione in GM abbia talvolta indotto in calchi, sintomo evidente dell'inesperienza delle studentesse che rimanda inevitabilmente a un'ulteriore esigenza didattica. Oltre a questo, nonostante le due modalità risultino di difficile paragone dato il diverso obiettivo, si osserva che nella condizione GM alcune partecipanti hanno ritenuto opportuno allontanare il proprio sguardo dagli stimoli per poter concentrarsi meglio sulla concretezza del discorso, in quanto il testo, a loro avviso, ne impediva un'elaborazione sufficientemente approfondita. Curiosamente, alcune partecipanti hanno percepito il proprio output come più "robotico" nella condizione GM proprio a causa della trascrizione che sembra aver amplificato, in molti casi, la densità informativa dei discorsi. Nelle riflessioni relative alla gestione dell'input visivo e dell'input auditivo, risulta particolarmente illustrativa del campione coinvolto e delle sue necessità la tendenza di alcune partecipanti a interrompere l'ascolto attivo durante l'IS nella condizione ST, in quanto l'estrema fiducia nell'accuratezza della trascrizione sembra aver avuto un impatto negativo sull'elaborazione dei discorsi. Per quanto riguarda il CAI di SmarTerp, molte partecipanti hanno riportato una difficoltà significativa nel gestire la complessità dell'input visivo e, al contempo, hanno posto particolare rilievo sulla ridondanza informativa talora presente nel sistema. Per

quanto riguarda la trascrizione di Google Meet, molte partecipanti hanno sottolineato l'impatto positivo della punteggiatura e la tendenza a eseguire una traduzione a vista della trascrizione automatica. Per entrambe le condizioni, le partecipanti hanno riscontrato difficoltà notevoli dovute alle omissioni dei due rispettivi sistemi. Relativamente alle necessità didattiche, le partecipanti hanno più volte sottolineato la difficoltà nel rendere più flessibile il proprio *décalage*, in modo da poter sfruttare a proprio vantaggio i sistemi proposti (in particolare, il CAI di SmarTerp); al contempo, alcune studentesse hanno evidenziato la difficoltà nel gestire tutti gli input dell'interfaccia di SmarTerp.

Sulla base di quanto emerso, è necessario sottolineare che il potenziale dell'interazione interprete-CAI *tool* può emergere solo se l'interprete è in grado di gestire attivamente la complessità dell'input visivo.

Naturalmente, questo progetto di ricerca non è esente da limiti. Anzitutto, le conclusioni sono state tratte da un campione esiguo per poterle estendere a considerazioni più generali; in particolare, i dati relativi alla percezione dell'usabilità degli strumenti - raccolti tramite UEQ (Laugwitz et al., 2008) - richiederebbero un campione più ampio per poter convalidare l'attendibilità e la rappresentatività dei risultati. Inoltre, la valutazione delle *performance* è stata condotta dal ricercatore autore dello studio: pur avendo prestato particolare attenzione all'imparzialità nella valutazione, le scale utilizzate richiederebbero una valutazione condotta da almeno due valutatrici esterne (che devono raggiungere un valore relativo all'*inter-rater reliability* tale da poter garantire l'omogeneità del giudizio). Oltre a ciò, il *software* SmarTerp è stato utilizzato nella sua versione Beta, in quanto ancora in fase di sperimentazione; pertanto, sarebbe interessante replicare lo studio valutando le *performance* in una condizione sperimentale contenente la versione definitiva del CAI. Inoltre, le retrospezioni sono state organizzate seguendo uno schema descrittivo che talvolta necessiterebbe di una maggior gerarchizzazione delle opinioni sulla base della frequenza e dell'utilità per l'identificazione dei problemi più significativi. A questo proposito, per eventuali studi futuri si raccomanda di seguire la matrice proposta da Davis Travis (2009), ideata proprio per l'organizzazione dei dati qualitativi durante un *usability testing*.

Un ulteriore sviluppo della ricerca potrebbe contemplare un confronto tra due gruppi di partecipanti con diversa formazione nel campo delle nuove tecnologie applicate all'IS. Sarebbe inoltre interessante replicare lo stesso studio per una combinazione linguistica che richiede un *décalage* tendenzialmente più esteso (ad esempio, la combinazione italiano-tedesco), in modo da valutare in quel caso l'impatto della latenza dei sistemi proposti. Inoltre, sarebbe interessante

verificare le necessità che potrebbero emergere dalla proposta di video non monologici a interpreti in formazione (ad esempio, Q&A), sempre mettendo a confronto una trascrizione automatica di buona qualità con il CAI di SmarTerp. Per ampliare il campione coinvolto sarebbe necessaria una replica sperimentale, ampliando il numero di partecipanti si giungerebbe a una maggior validità dei dati. Sarebbe inoltre interessante approfondire la correlazione tra la percezione dell'usabilità degli strumenti tecnologici (quantitativamente misurata tramite UEQ) e l'attitudine verso l'adozione delle nuove tecnologie da parte delle interpreti, come proposto da Mellinger & Hanson (2018).

In potenziali studi futuri, sarebbe interessante analizzare aspetti come l'ansia e lo stress durante l'IS con CAI *tool.* Per perseguire tale fine sarebbe opportuna una triangolazione dei dati di *eye-tracking* con quelli relativi all'accuratezza delle rese e con quelli ottenuti tramite il questionario NASA-TLX (Hart, 2006), in modo da avere un approccio multidimensionale al costrutto del carico di lavoro nel contesto dell'interazione interprete-CAI *tool*.

Infine, quanto messo in luce in questo studio (e quanto potrebbe essere ulteriormente approfondito attraverso studi futuri) potrebbe essere applicato attraverso la promozione e l'organizzazione di corsi mirati al rafforzamento della competenza tecnologica in interpreti in formazione. In questo modo, si potrebbe consolidare un rapporto interprete-macchina, in cui le interpreti siano consapevoli dell'importanza della gestione di un input visivo complesso per poter migliorare una *performance* in cui il lato umano dell'interprete non potrà mai essere superato da una macchina.

Bibliografia

Agrawal, P., & Ganapathy, S. (2019). Modulation filter learning using deep variational networks for robust speech recognition. *IEEE J. Sel. Topics Signal Process, 13*(2), 244 - 253. doi:10.1109/JSTSP.2019.2913965

Amos, R., Seeber, K., & Pickering, M. (2022). Prediction during simultaneous interpreting: Evidence from the visual-world paradigm. *Cognition*, 1 - 16.

Baddeley, A. D. (2000). Short-term and working memory. In E. Tulving, & F. Craik, *The Oxford handbook of memory* (p. 77 - 92). New York: Oxford University Press.

Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower, *Psychology of learning and motivation* (p. 47 - 89). New York: Academic Press.

Baigorri Jalón, J. (2004). *Interpreters at the United Nations: A History.* Salamanca: Universidad de Salamanca.

Barbero, J., Bermejo, F., & San Vicente, F. (2018). *Contrastiva. Grammatica della lingua spagnola.* Bologna: CLUEB.

Barrault, L., Chung, Y.A., Meglioli, M., Dale, D., Dong, N., Duppenthaler, M., Kulikov, I. (2023). *Seamless: Multilingual Expressive and Streaming Speech Translation.* Seamless Communication.

Bendazzoli, C. (2010). *Testi e contesti dell'interpretazione di conferenza: uno studio etnografico.* Bologna: Asterisco.

Bertozzi, M., González Rodríguez, M., & Russo, M. (2021). Interpretare tra spagnolo e italiano. In M. Russo, *Interpretare da e verso l'italiano. Didattica e innovazione per la formazione dell'interprete.* (p. 289 - 312). Bononia. University Press.

Bevan, N., Carter, J., Earthy, J., Geis, T., & Harker, S. (2016). New ISO standards for usability, usability reports and usability measures. In M. Kurosu, *Human-computer interaction: Theory, design, development and practice, 18th international conference on Human-Computer Interaction (HCI 2016)* (p. 266 - 278). Cham: Springer.

Bevan, N., Kirakowskib, J., & Maissela, J. (1991). What is usability. *Proceedings of the 4th international conference on HCI.*

Braun, S. (2019). Technology and interpreting. In M. O'Hagan, *The Routledge handbook of translation and technology* (p. 271-288). Oxford: Pergamon Press.

Brooke, J. (1996). Sus: A "quick and dirty" usability scale. In P. Jordan, T. B., I. McClelland, & B. Weerdmeester, *Usability evaluation in industry* (p. 189 - 194). London: CRC Press.

Budiu, R., & Moran, K. (2021, July 25). How Many Participants for Quantitative Usability Studies: A Summary of Sample-Size Recommendations. *Nielsen Norman Group*.

Cai, C., Xu, Y., Ke, D., & Su, K. (2015). A fast learning method for multi-layer perceptrons in automatic speech recognition systems. *J. Robot, 2015*, 1 - 7. doi:10.1155/2015/797083

Calvi, M. (2003). La lexicografía bilingüe de español e italiano. In F, F. San Vicente, & M. Calvi, *Didáctica del léxico y nuevas tecnologías.* (p. 39 - 53). Mauro Baroni.

Carl, M., & Braun, S. (2018). Translation, interpreting and new technologies. In K. Malmkjaer, *The Routledge Handbook of Translation Studies and Linguistics.* (p. 374 - 390). London: Routledge. doi:10.4324/9781315692845-25

Carroll, J. (1966). An Experiment in Evaluating the Quality of Translations. *Mechanical Translations and Computational Linguistics, 9*(3 - 4), 55 - 66.

Cauwenberghe, G. V. (2020). *La reconnaissance automatique de la parole en interprétation simultanée.* Gent: Universiteit Gent.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format od instruction. *Cognition and Instruction, 8*(4), 239 - 332. doi:10.1207/s1532690xci0804_2

Chen, Z., Droppo, J., Li, J., & Xiong, W. (2018). Progressive joint modeling in unsupervised single-channel overlapped speech recognition. *IEEE/ACM Trans. Audio, Speech, Language Process, 26*(1), 184 - 196. doi:10.1109/TASLP.2017.2765834

Chmiel, A. (2021). Eye-tracking studies in conference interpreting. In M. Albl-Mikasa, & E. Tiselius, *The Routledge Handbook of Conference Interpreting* (p. 457 - 470). Routledge.

Chmiel, A., & Lijewska, A. (2019). Syntactic processing in sight translation by professional and trainee interpreters: Professionals are more time-efficient while trainees view the source text less. *Target, 31*(3), 378 - 397.

Chmiel, A., Lijewska, A., & Janikowski, P. (2020). Multimodal processing in simultaneous interpreting with text: Interpreters focus more on the visual than the auditory modality.

Target International Journal of Translation Studies, 32(1), 37 - 58. doi:10.1075/target.18157.chm

Christian, B. (2011). *The Most Human Human: A Defence of Humanity in the Age of the Computer.* Londra: Viking.

Cohen, R., Swerdlik, M., & Phillips, S. (1996). *Psychological Testing and Assessment: An Introduction to Tests and Measurement.* Mayfield: Mountain View.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104*(2), 163 - 191. doi:10.1037/0033-2909.104.2.163

Creswell, J., & Creswell, D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage: Thousand Oaks.

Defrancq, B., & Fantinuoli, C. (2021). Automatic speech recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers. *Target, 33*(1), 73 - 102.

Dietrich, K., & Peters, J. (2002). Testing the correlation of word error rate and perplexity. *Speech Communication*, 19 - 28. doi:10.1016/S0167-6393(01)00041-3

Dunn, O. (1961). Multiple Comparisons among Means. *Journal of the American Statistical Association, 56*(293), 52-64. Tratto da: https://doi.org/10.1080/01621459.1961.10482090

EABM. (2018). *CAI. What is computer-assisted interpreting?* Tratto da Ghent University: https://www.eabm.ugent.be/CAI

Ehlich, K., & Rehbein, J. (1976). Halbinterpretative Arbeitstranskriptionen (HIAT). *Linguistische Berichte, 45*, 21-41.

El Shafey, L., Soltau, H., & Shafran, I. (2019). *Joint Speech Recognition and Speaker Diarization via Sequence Transduction.* ArXiv. Tratto da https://doi.org/10.48550/arXiv.1907.05337

Englund Dimitrova, B. (2005). *Expertise and Explicitation in the Translation Process.* Amsterdam/Philadelphia: Benjamins.

Englund Dimitrova, B., & Tiselius, E. (2009). Exploring retrospection as a research method for studying the translation process and the interpreting process. In I. Mees, F. Alves, & S.

Göpferich, *Methodology, Technology and Innovation in Translation Process Research. A Tribute to Arnt Lykke Jakobsen.* (p. 109 - 134). Frederiksberg: Samfundslitteratur Press.

Englund Dimitrova, B., & Tiselius, E. (2014). Retrospection in interpreting and translation: explaining the process? *MonTi Monografías de Traducción e Interpretación*, 177 - 200. doi:10.6035/MonTI.2014.ne1.5

Enríquez Raído, V. (2011). Developing web searching skills in translator training. *redit: Revista electrónica de didáctica de la traducción y la interpretación.*, 60 - 80. Tratto da https://riuma.uma.es/xmlui/handle/10630/11378

Ericsson, K., & Simon, H. (1987). Verbal reports on thinking. In C. Faerch, & G. Kasper, *Introspection in second language research* (p. 24 - 53). Multilingual Matters.

Ericsson, K., & Simon, H. (1993). *Protocol Analysis. Verbal Reports as Data.* Cambridge: Mass: MIT.

Færch , C., & Kasper, G. (1987). From product to process - Introspective methods in second language research. In C. Færch, & G. Kasper, *Introspection in Second Language Research* (p. 5 - 23). Clevedon/Philadelphia: Multilingual Matters.

Fantinuoli, C. (2016). InterpretBank: Redefining computer-assisted interpreting tools. In J. Esteves-Ferreira, J. Macan, R. Mitkov, & O.M. Stefanov, *Proceedings of Translating and the Computer 38* (p. 42 - 52). Londra: AsLing. Tratto da https://aclanthology.org/2016.tc-1.5

Fantinuoli, C. (2017). Speech recognition in the interpreter workstation. In J. Steves-Ferreira, J. Macan, R. Mitkov, & O.M. Stefanov, *Proceedings of the Translating and the Computer 39* (p. 25 - 34). Londra: AsLing.

Fantinuoli, C. (2018a). Interpreting and technology: The upcoming technological turn. In C. Fantinuoli, *Interpreting and technology* (p. 1 - 12). Berlino: Language Science Press.

Fantinuoli, C. (2018b). Computer-assisted Interpreting: Challenges and Future Perspectives. In G. C. Pastor, & I. Durán-Muñoz, *Trends in E-Tools and Resources for Translators and Interpreters* (p. 153 - 176).

Fantinuoli, C., Marchesini, G., Landan, D., & Horak, L. (2022). *KUDO Interpreter Assist: Automated Real-time Support for Remote Intepretation.* arXiv.

Feng, J. (2018). *译术||机器口译：口译教育面临的挑战与对策. In: 翻译.* Tratto da http://www.sohu.com/a/247079088_176673

Frittella, F. (2023). *Usability research for interpreter-centred technology: The case study of SmarTerp.* Berlin: Language Science Press. doi:10.5281/zenodo.7376351

Frittella, F., & Rodríguez, S. (2022). Putting SmartTerp to Test: A tool for the challenges of remote interpreting. *INContext Studies in Translation and Interculturalism*, 137 - 166. Tratto da https://doi.org/10.54754/incontext.v2i2.21

Fusco, M. (1990). Quality in Conference Interpreting between Cognate Languages: A Preliminary approach to the Spanish-Italian case. *The Interpreters' Newsletter*, 93 - 97.

Gaido, M., Rodríguez, S., Negri, M., Bentivogli, L., & Turchi, M. (2021). Is "Moby Dick" a whale or a bird? Named entities and terminology in speech translation. *arXiv*. doi:10.48550/arXiv.2109.07439

Ganapathy, S. (2017). Multivariate autorefressive spectrogram modeling for noisy speech recognition. *IEEE Signal Process. Lett., 24*(9), 1373 - 1377. doi:10.1109/LSP.2017.2724561

Gile, D. (1988). Le partage de l'attention et le "Modèle d'efforts" en interprétation simultanée. *The Interpreters' Newsletter 1*, 4 - 22. Tratto il giorno 26 Gennaio 2024 da http://hdl.handle.net/10077/2156

Gile, D. (1995). *Regards sur la recherche en interprétation de conférence.* Lille: Presses universitaires de Lille.

Gile, D. (1997). Conference interpeting as a cognitive management problem. In F. Pöchhacker, & M. Shlesinger, *The interpreting studies reader* (p. 163 - 176). London: Routledge.

Gile, D. (1999). Testing the effort models' tightrope hypothesis in simultaneous interpeting: A contribution. *HERMES: Journal of Language and Communication in Business, 23*, 153 - 172. doi:10.7146/hjlcb.v12i23.25553

Gile, D. (2008). Local cognitive load in simultaneous interpreting and its implications for empirical research. *FORUM: Revue internationale d'interprétation et de traduction / International Journal of Interpretation and Translation, 6*(2), 59 - 77. doi:10.1075/forum.6.2.04gil

Gile, D. (2009). *Basic concepts and Models for Interpreter and Translator Training.* John Benjamins Publishing Company. doi:https://doi.org/10.1075/btl.8

Google Cloud (2024). *Misurare e migliorare la precisione della voce*. Tratto da https://cloud.google.com/speech-to-text/docs/speech-accuracy?hl=it.

Gosztolya, G., & Grósz, T. (2016). Domain adaptation of deep neural networks for automatic speech recognition via wireless sensors. *J. Electr. Eng., 67*(2), 124 - 130. doi:10.1515/jee-2016-0017

Guo, M., Anacleto, M., & Han, L. (2023). Computer-Assisted Interpreting Tools: Status Quo and Future Trends. *Theory and Practice in Language Studies, 13*(1), 89 - 99. doi:https://doi.org/10.17507/tpls.1301.11

Hart, S. (2006). NASA-task load index (NASA - TLX); 20 years later. *Proceedings of the human factors and ergonomics society annual meeting. 50*, p. 904 - 908. Los Angeles: Sage Publications.

Hart, S., & Staveland, L. (1988). Devolpment of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock, & N. Meshkati, *Human mental workload* (p. 139-183). Amsterdam: North-Holland.

Hassenzahl, M. (2003). The thing and I: Understanding the relationship between user and product. In M. A. Blythe, K. Overbeeke, A. F. Monk, & P. C. Wright, *Funology* (p. 31 - 42). Dordrecht: Springer. doi:10.1007/1-4020-2967-5_4

Hee-Jo, Y. (2021). *Translatotron 2*. Tratto da Smilegate AI: https://smilegate.ai/en/2021/09/02/translatotron-2/

Hidalgo-Barnes, M., & Massaro, D. (2007). Read my lips: An animated face helps communicate musical lyrics. *Psychomusicology: A Journal of Research in Music Cognition, 19*(2), 3 - 12.

Hirayama, N., Yoshino, K., Itoyama, K., Mori, S., & Okuno, H. (2015). Automatic speech recognition for mixed dialect utterances by mixing dialect language models. *IEEE/ACM Trans. Audio, Speech, Lanugage Process, 23*(2), 373 - 382. doi:10.1109/TASLP.2014.2387414

Holmqvist, K., Nystrom, M., Anderson, R., Dewhurst, R., Jarodzka, H., & van de Weiter, J. (2015). *Eye Tracking: A Comprehensive Guide to Methods and Measures.* Oxford: Oxford University Press.

Hvelplund, K. (2014). Eye tracking and the translation process: Reflections on the analysis and interpretation of eye-tracking data. *MonTI: Monografías de traducción e interpretación 1, 1*, 201 - 224.

Hwang, W., & Salvendy, G. (2010). Number of people required for usability evaluation: The 10±2 rule. *Communications of the ACM, 53*(5), 130 - 133. Tratto da https://doi.org/10.1145/1735223.1735255

Hyscaler. (2023, 12 1). *Seamless Communication: Meta's breakthrough in universal speech translation.* Tratto da Hyscaler: https://hyscaler.com/insights/seamless-communication-meta-breakthrough/?p=cg9zddoxndmymg

Ilg, G. (1978). De l'allemand vers le français: L'apprendissage de l'interprétation simultanée. *Parallèles*, 69 - 99.

ISO. (2010). Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems. *Standard ISO 9241-210*. Geneva.

ISO. (2018). Ergonomics of human-system interaction - Part 11: usability: definitions and concepts. *Standard ISO 9241-11*. Geneva.

Jacquemot, C., Dupoux, E., & Bachoud-Lévi, A.-C. (2011). Is the word-length effect linked to subvocal rehearsal? *Cortex, 47*(4), 484 - 494. doi:10.1016/j.cortex.2010.07.007

Jia, Y., Ramanovich, M., Remez, T., & Pomerantz, R. (2021). *Translatotron 2: High-quality direct speech-to-speech translation with voice preservation.* Tratto da https://doi.org/10.48550/arXiv.2107.08661

Jones, R. (2014). *Conference interpreting explained.* Londra e New York: Routledge.

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition.* (2nd ed., Vol. Prentic Hall series in artificial intelligence). Pearson Prentic Hall, Upper Saddle River.

Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329 - 354. Tratto da https://doi.org/10.1037/0033-295X.87.4.329

Karpagavalli, S., & Chandra, E. (2016). A Review on Automatic Speech Recognition Architecture and Approaches. *International Journal of Signal Processing, Image Processing and Pattern Recognition, 9*(4), 393 - 404. Tratto da http://dx.doi.org/10.14257/ijsip.2016.9.4.34

Kekule, M., Spanova, I., & Viiri, J. (2019). Benefits of using the eye-tracking method for qualitative observation of students' multiple-choice physics tasks solution process. *Pedagogická orientace, 29*(4), 424 - 465. doi:10.5817/PedOr2019-4-424

Korpal, P., & Stachowiak-Szymczak, K. (2018). The whole picture: Processing of numbers and their context in simultaneous interpreting. *Poznan Studies in Contemporary Linguistics, 54*(3), 335 - 354.

Kruger, J.-L., & Steyn, F. (2013). Subtitles and Eye Tracking: Reading and Performance. *Reading Research Quarterly, 49*(1), 105 - 120. Tratto da https://doi.org/10.1002/rrq.59

Lakhotia, K., Kharitonov, E., Hsu, W.N., Adi, Y., Polyak, A., Bolte, B., Dupoux, E. (2021). *Generative Spoken Language Modeling from Raw Audio.* arXiv.

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In A. Holzinger, *HCI and usability for education and work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society (USAB 2008)* (p. 63 - 76). Berlin: Springer. doi:10.1007/978-3-540-89350-9_6

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451-468. doi:10.1037/0096-1523.21.3.451

Lazar, J., Jinjuan, H., & Hochheiser, H. (2017). *Research methods in human-computer interaction.* . Cambridge: MA: Morgan Kaufmann.

Lederer, M. (1982). Le processus de la traduction simultantée. *Multilingua*, 149 - 158.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*, 325 - 343. doi:10.3758/s13428-011-0146-0

Lewis, J. (2012). Usability testing. In G. Salvendy, *Handbook of human factors and ergonomics* (p. 1267 - 1312). Hoboken, NJ: John Wiley.

Liversedge, S., Paterson, K., & Pickering, M. (1998). Eye movements and measures of reading time. In G. Underwood, *Eye guidance in reading and scene perception* (p. 55 - 75). Elsevier Science Ltd. doi:https://doi.org/10.1016/B978-008043361-5/50004-3

Logie, R. H. (1995). *Visuo-spatial working memory.* London: Psychology Press.

Longhui, Z., Carl, M., & Feng, J. (2022). Patterns of Attention and Quality in English-Chinese Simultaneous Interpreting with Text. *International Journal of Chinese and English Translation & Interpreting*, 1 - 23. doi:https://doi.org/10.56395/ijceti.v2i2.50

Lu, S., Wickens, C., Sarter, N., Thomas, L., Nikolic, M., & Sebok, A. (2012). Redundancy Gains in Communication Tasks: A Comparison of Auditory, Visual, and Redundant Auditory-Visual Informaion Presentation on NextGen Flight Decks. *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting*, (p. 1476 - 1480). Tratto da https://doi.org/10.1177/1071181312561413

Mayer, R., & Fiorella, L. (2014). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. Mayer, *The Cambridge handbook of multimedia learning* (p. 279 - 315). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139547369.015

Mayhew, D. J. (2007). Requirements specifications within the usability engineering lifecycle. In A. Sears, & J. Jacko, *The human-computer interaction handbook* (2 ed., p. 913-921). Boca Raton: CRC Press.

Meade, M., & Roediger, H. (2002). Explorations in the social contagion of memory. *Memory and Cognition, 30*(7), 995 - 1009.

Mellinger, C., & Hanson, T. (2017). *Quantitative Research Methods in Translation and Interpreting Studies.* Routledge.

Mellinger, C., & Hanson, T. (2018). Interpreter traits and the relationship with technology and visibility. (N. Pokorn, & C. Mellinger, A cura di) *Community Interpreting, Translation and Technology*, 366 - 392. doi:https://doi.org/10.1075/tis.00021.mel

Merrienboer, J. J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147 - 177. doi:10.1007/s10648-005-3951-0

Ming, J., & Crookes, D. (2017). Speech enhancement based on full-sentence correlation and clean speech recognition. *IEEE/ACM Trans. Audio, Speech, Language Process, 25*(3), 531 - 543. doi:10.1109/TASLP.2017.2651406

Montecchio, M. (2021). Defining maximum acceptable latency of AST-enhanced CAI tools. Quantitative and qualitative assessment of the impact of ASR latency on interpreters'performance. Germersheim: University of Mainz.

Morelli, M. (2010). *La interpretación español-italiano: planos de ambigüedad y estrategias.* Granada: Comares.

Müller, M., Nguyen, T., Niehues, J., Cho, E., Krüger, B., Ha, T.-L., Waibel, A. (2016). Lecture Translator - Speech translation framework for simultaneous lecture translation. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations.* (p. 82 - 86). San Diego, California: Association for Computational Linguistics. doi:10.18653/v1/N16-3017

Nielsen, J. (1993). *Usability engineering.* San Francisco, CA: Morgan Kaufmann.

Nielsen, J. (2010). What is usability? In C. Wilson, *User experience re-mastered: Your guide to getting the right design* (p. 3 - 22). Burlington: MA: Morgan Kaufamnn.

Nielsen, J. (2012). *Usability 101: Introduction to usability.* Nielsen Norman Group. Tratto da http://www.nngroup.com/articles/usability-101-introduction-to-usability/

Nielsen, J., & Landauer, T. (1993). A mathematical model of the finding of usability problems. In B. Arnold, G. van der Veer, & T. White, *Proceedings of the INTERACT'93 and CHI'93 Conference on Human factors in Computing Systems* (p. 206 - 213). New York: Association for Computing Machinery. doi:10.1145/169059.169166

Norm, D. (2018). Ergonomics of human-system interaction - part 11: Usability: Definitions and concepts. *ISO 9241 11* .

Orlando, M. (2010). Digital pen technology and consecutive interpreting: Another dimension in notetaking training and assessment. *The Interpreters' Newsletter, 15*, 71 - 86. Tratto da http://hdl.handle.net/10077/4750

Palmerini, M., & Savy, R. (2014). Gli errori di un sistema di riconoscimento auomatico del parlato. Analisi linguistica e primi risultati di una ricerca interdisciplinare. *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & and of the*

Fourth International Workshop EVALITA 2014 : 9-11 December 2014, Pisa - Pisa : Pisa University Press, 2014- Casalini id: 3003965, (p. 281 - 285). Pisa.

Paradis, M. (1994). Toward a neurolinguistic theory of simultaneous translation: The framework. *International Journal of Psycholinguistics, 10*(3 [29]), 319 - 335.

Pöchhacker, F. (2016). *Introducing interpreting studies* (2 ed.). Routledge.

Prandi, B. (2017). Designing a multimethod study on the use of CAI tools during simultaneous interpreting. In J. Esteves-Ferreira, J. Macan, R. Mitkov, & O.M. Stefanov, *Proceedings of Translating and the Computer 39* (p. 76 - 113). London: AsLing.

Prandi, B. (2023). Computer-assisted simultaneous interpreting. A cognitive-experimental study on terminology. *Translation and Multilingual Natural Language Processing, 22*.

Pratap, V., Xu, Q., Kahn, J., Avidov, G., Likhomanenko, T., Hannun, A., Collobert, R. (2020). Scaling Up Online Speech Recognition Using ConvNets. *arXiv*.

Pym, A. (2011). Training Translators. In K. Malmkjaer, & K. Windle, *The Oxford Handbook of Translation Studies.* doi:10.1093/oxfordhb/9780199239306.013.0032

Pym, A. (2011). What technology does to translating. *Translation & Interpreting, 3*(1), 1 - 9.

Ramadhika, B., Yosintha, R., & Yunianti, S. (2022). Automatic caption features on Google Meet as a pronunciation assessment tool. *IJEE (Indonesian Journal of English Education), 9*(2), 396 - 410. doi:10.15408/ijee.v9i2.22482

Rampin, R., & Rampin, V. (2021). Taguette: open-source qualitative data analysis. *Journal of Open Source Software, 6*(68). doi:https://doi.org/10.21105/joss.03522

Rodríguez, S., Gretter, R., Matassoni, M., Falavigna, D., Alonso, Á., Corcho, O., & Rico, M. (2021, Luglio 05). SmarTerp: A CAI System to Support Simultaneous Interpreters in Real-Time. *Translation and Interpreting Technology Online*, 102 - 109. Tratto da https://doi.org/10.26615/978-954-452-071-7_012

Rosenwein, T. (2023, March 30). *Overcoming Automatic Speech Recognition Challenges: The Next Frontier. Advancements, Opportunities, and Impacts of Automatic Speech Recognition Technology in Various Domains.* Tratto da Medium: https://towardsdatascience.com/overcoming-automatic-speech-recognition-challenges-the-next-frontier-e26c31d643cc

Ross, B., Pechenika, E., Aeschliman, C., & Chase, A.M. (2017). Print versus figital texts: understanding the experimental research and challenging the dichotomies. *Research in Learning Technology*. doi:https://doi.org/10.25304/rlt.v25.1976

Russo, M. (2012). *Interpretare lo spagnolo. L'effetto delle dissimmetrie morfosintattiche nella simultanea.* Bologna: CLUEB.

Russo, M. (2019). Discursos jurídicos en interpretación simultánea. Un reto concreto: los falsos amigos. *INTERLINGUA*, 57 - 82.

Rütten, A., Eberhardt , A., & Sand, P. (2020, 11 30). *The Unfinished Handbook for Remote Simultaneous Interpreters.* Tratto da Dolmetscher-wissen-alles.de: https://blog.sprachmanagement.net/rsi-handbook/

Schrepp, M. (2023). *User Experience Questionnaire Handbook. All you need to know to apply the UEQ successfully in your projects.* User Experience Questionnaire.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. *Third International Conference of Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience.* (p. 383 - 392). Heraklion: Springer International Publishing. doi:10.1007/978-3-319-07668-3_37

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence, 4*(4), 40 - 44. doi:10.9781/ijimai.2017.445

Seeber, K. (2011). Cognitive load in simultaneous interpreting: Existing theories-new models. *Interpreting: International Journal of Research and Practice in Interpreting, 13*(2), 176 - 204. doi:10.1075/intp.13.2.02see

Seeber, K. (2012). Multimodal input in simultaneous interpreting: an eye-tracking experiment. *Proceedings of the 1st International Conference TRANSLATA. Translation and Interpreting Research: Yesterday-Today-Tomorrow.* Peter Lang.

Seeber, K. (2015). Eye tracking. In F. Pöchhacker, *Routledge Encyclopedia of Interpreting* (p. 157). New York: Routledge.

Seeber, K. (2017). Multimodal processing in simultaneous interpreting. In J. W. Schwieter, & A. Ferreira, *The handbook of translation and cognition* (p. 461 - 475). Hoboken: Wiley. doi:10.1002/9781119241485.ch25

Seeber, K. G. (2007). Thinking outside the cube: Modeling language processing tasks in a multiple resource paradigm. *Interspeech 2007*.

Seeber, K. G., Keller, L., & Hervais-Adelman, A. (2020). When the ear leads the eye: The use of text during simultaneous interpretation. *Language, Cognition and Neuroscience, 35*(10), 1480 - 1494. doi:10.1080/23273798.2020.1799045

Seeber, K., & Kerzel, D. (2012). Cognitive load in simultaneous interpreting: Model meets data. *International Journal of Bilingualism, 16*(2), 228 - 242.

Seeber, K., Keller, L., & Hervais-Adelman, A. (2020). When the ear leads the eye – the use of text during simultaneous interpretation. *Language, Cognition and Neuroscience, 35*(10), 1480 - 1494. doi:10.1080/23273798.2020.1799045

Seewald, F., & Hassenzahl, M. (2004). Vom kritischen Ereignis zum Nutzungsproblem: Die qualitative Analyse in diagnostischen Usability Tests. In M. Hassenzahl, & M. Peissner , *Tagungsband Usability Professionals 2004* (p. 142 - 148). Stuttgart: Fraunhofer Verlag.

Seligman, M., Waibel, A., Joscelyne, A., & Stoker, A. (2017). *TAUS Speech-to-Speech Translation Technology Report.* TAUS BV. Tratto da https://isl.anthropomatik.kit.edu/downloads/S2STranslationTechnologyReport.final.pdf

Seubert, S. (2019). *Visuelle Informationen beim Simultandolmetschen: Eine Eyetracking-Studie.* Berlin: Frank & Timme.

Spinolo, N. (2018). *Tra il dire e il significare: Il linguaggio figurato nell'interpretazione simultanea fra italiano e spagnolo.* Pisa: ETS Edizioni.

SR Research Eyelink Ltd. (2009). EyeLink® 1000 User Manual. Tower, Desktop, LCD Arm, Primate and Long Range Mounts. Remote, 2000 hz and Fiber Optic Camera Upgrades. Mississauga, Ontario, Canada.

Stachowiak-Szymczak, K., & Korpal, P. (2019). Interpreting Accuracy and Visual Processing of Numbers in Professional and Student Interpreters: an Eye-tracking Study. *Across Languages and Cultures, 20*(2), 235 - 251. doi:10.1556/084.2019.20.2.5

Tarasenko, R., Amelina, S., & Semerikov, S. (2021). Conceptual Aspects of Interpreter Training Using Modern Simultaneous Interpretation Technologies. *ICTERI 2021: 17th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer.*, (p. 1 - 13). Kherson.

Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*(4), 424 - 436. doi:10.1037/00220663.80.4.424

The jamovi project (2024). jamovi (Version 2.5) [Computer Software]. Tratto da https://www.jamovi.org

Tiselius, E. (2009). Revisiting Carroll's scales. In C. V. Angelelli, & H. E. Jacobson, *Testing and Assessment in Translation and Interpreting Studies* (p. 95 - 121). Amsterdam/Philadelphia: John Benjamins.

Travis, D. (2009). *How to prioritise usability problems.*

Trovato, G. (2014). *AGON, 2*, 37 - 70.

Wang, D., Wang, X., & Lv, S. (2019). An overview of end-to-end auomatic speech recognition. *Symmetry, 11*(8), 1018. doi:10.3390/sym11081018

Whiteside , J., Bennett, J., & Karen Holtzblatt. (1988). Usability engineering: Our experience and evolution. In M. G. Helander, *Handbook of human-computer interaction.* (p. 791 - 817). Amsterdam: North-Holland.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasumaran, & D. Davies, *Varieties of attention* (p. 63 - 101). New York: Academic Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159 - 177. doi:10.1080/14639220210123806

Yuan, L., & Wang, B. (2023). Cognitive processing of the extra visual layer of live captioning in simultaneous interpreting. Triangulation of eye-tracked process and performance data. *Ampersand. An International Journal of General and Applied Linguistics., 11*, 1 - 9. Tratto da https://doi.org/10.1016/j.amper.2023.100131

Zavras, A., & Finn, E. (2017). [Recensione di What algorithms want: Imagination in the age of computing]. *Icon, 23*, 201 - 203. Tratto da www.jstor.org/stable/26454995

Zeghidour, N., Luebs, A., Omran, A., Skoglund, J., & Tagliasacchi, M. (2021). *SoundStream: An End-to-End Neural Audio Codec.* ArXiv.

Ziegler, K., & Gigliobianco, S. (2018). Present? Remote? Remotely present! New technological approaches to remote simultaneous conference interpreting. In C. Fantinuoli, *Interpreting and technology.* (p. 119 - 139). Berlin: Language Science Press. doi:10.5281/zenodo.1493299

Sitografia

https://interplex.com/ (Ultimo accesso: 29 aprile 2024)

https://interpretershelp.com/ (Ultimo accesso: 29 aprile 2024)

https://www.lookup-web.de/index.php (Ultimo accesso: 29 aprile 2024)

https://interpretbank.com/site/ (Ultimo accesso: 29 aprile 2024)

https://kudoway.com (Ultimo accesso: 29 aprile 2024)

https://smarterp.me (Ultimo accesso: 29 aprile 2024)

https://commonvoice.mozilla.org/it (Ultimo accesso: 24 aprile 2024)

https://www.euronews.com/ (Ultimo accesso: 24 aprile 2024)

https://ai.meta.com/research/seamless-communication/ (Ultimo accesso: 14 giugno 2024)

https://ai.meta.com/research/publications/seamless-multilingual-expressive-and-streaming-speech-translation/ (Ultimo accesso: 19 giugno 2024)

https://dit.unibo.it/it (Ultimo accesso: 29 aprile 2024)

https://www.unibo.it/it/studiare/dottorati-master-specializzazioni-e-altra-formazione/insegnamenti/insegnamento/2023/493148 (Ultimo accesso: 24 aprile 2024)

https://www.unibo.it/it/studiare/dottorati-master-specializzazioni-e-altra-formazione/insegnamenti/insegnamento/2023/433465 (Ultimo accesso: 24 aprile 2024)

https://dittafono.ditlab.it/ (Ultimo accesso: 24 aprile 2024)

http://www.aispi.it/ (Ultimo accesso: 24 aprile 2024)

https://eur-lex.europa.eu/homepage.html?locale=es (Ultimo accesso: 14 aprile 2024)

https://es.wikipedia.org/ (Ultimo accesso: 18 aprile 2024)

https://aiic.org/site/world/about/profession/abc (Ultimo accesso: 30 aprile 2024)

https://www.lextale.com/validity.html (Ultimo accesso: 30 aprile 2024)

https://www.congreso.es/es/busqueda-de-intervenciones (Ultimo accesso: 14 aprile 2024)

https://filmora.wondershare.es/ (Ultimo accesso: 13 aprile 2024)

https://www.audacityteam.org/download/ (Ultimo accesso: 13 aprile 2024)

https://openai.com/research/whisper (Ultimo accesso: 13 aprile 2024)

https://www.exmaralda.org/pdf/HIAT_EN.pdf (Ultimo accesso: 13 aprile 2024)

https://www.ueq-online.org/ (Ultimo accesso: 28 aprile 2024)

https://app.taguette.org/ (Ultimo accesso: 28 aprile 2024)

https://www.jamovi.org/ (Ultimo accesso: 6 giugno 2024)

Appendice I. Discorso A.1 con istruzioni per la lettura.

Istruzioni per la lettura:

1. Leggi a una velocità di circa 127 parole al minuto.

2. Fai una pausa di 3 secondi circa [equivalente a un respiro non troppo prolungato] quando vedi questo simbolo nel discorso: [pause].

3. Utilizza un andamento prosodico naturale e spontaneo, cercando di utilizzare una pausa naturale alla fine di un'idea o di un paragrafo. La velocità deve essere rilassata e spontanea, senza rallentamenti o accelerazioni.

4. Il simbolo // segna il passaggio da un minuto all'altro. Tendenzialmente, ogni minuto dovresti arrivare a un // (NON EQUIVALE A UNA PAUSA).

5. Inquadratura dell'oratrice a mezzo busto, non solo faccia. Sfondo totalmente bianco (muro). L'oratrice non deve fare gesti durante la lettura. L'oratrice deve preferibilmente utilizzare una cuffia con microfono.

Durata: 8 minuti e 50 secondi circa

Parole al minuto: 127

Parole totali: 1122

Discorso:

Hoy subo a esta tribuna para defender el proyecto de ley de Presupuestos Generales del Estado en lo relativo a la sección del Ministerio para la Transición Ecológica y el Reto Demográfico. Los presupuestos que presentamos se alinean con la urgente necesidad de reactivar la economía.

El presupuesto asciende en su conjunto a 10 195 millones de euros, de los que 5817 corresponden a recursos del presupuesto nacional y 4378 corresponden a recursos del plan España Puede, tal y como les explicaré extensa y detalladamente. Este presupuesto supone una movilización sin precedentes en relación con tres de los grandes retos que tiene el país: la

transformación del sistema energético, la respuesta a la emergencia climática y la cohesión territorial. ***[pause]***

El primer gran reto, relativo a la transición energética, // cuenta con un presupuesto de 7397 millones de euros. Con este presupuesto, el Gobierno de España quiere mostrar su firme determinación en su voluntad de asumir un papel de liderazgo en esta transición a nivel europeo. ***[pause]***

El presupuesto destinado al Instituto para la Transición Justa crecerá hasta 442 millones, gracias al Mecanismo de Recuperación, que no deja de ser trascendental en la política de nuestro Gobierno. En el presupuesto de este instituto resalta el programa de 53 millones para proyectos generadores de empleo, que siempre hemos defendido los socialistas, en nuestro constante compromiso por la solidaridad social. Tampoco nos podemos olvidar de que la recuperación debe ser justa con los consumidores, sobre todo con los consumidores más vulnerables: los que tienen que luchar por llegar a fin de mes.// ***[pause]***

Nuestro compromiso frente a la pobreza energética queremos reforzarlo en este año complicado. Quiero poner en valor el bono social térmico, el cual se ha sometido a una comisión de expertos. Además, las medidas para la eficiencia energética y la promoción de comunidades energéticas son el pilar, la base de la transformación, sirven para reducir el consumo de energía. ***[pause]***

Por eso, estas políticas contarán con 1233 millones de euros que se pueden financiar con fondos FEDER. Como pueden observar, el esfuerzo en materia energética de estos presupuestos es enormemente relevante; servirá para afrontar el reto de la transición hacia una energía limpia, eficiente, sostenible. ***[pause]***

De hecho, cabe destacar el gran éxito que ha tenido durante el periodo de apertura iniciado apenas hace diez días y // que sigue abierto del Programa DUS 5000. El segundo gran reto al que nos enfrentamos es el de la preservación de nuestros ecosistemas, de la biodiversidad, de los recursos naturales y del territorio que ven la influencia de la acción del ser humano. ***[pause]***

La Secretaría de Estado liderada por Morán Fernández cuenta con un presupuesto de 579 millones de euros, cifra que supera con creces la de 2021, la cual representa un 2,85 % más. En este contexto, España aspira a lograr los 22 gigavatios de capacidad antes de 2032. Además, el capítulo para la preservación de biodiversidad marina tiene un presupuesto de casi medio millón de euros. El mar debe preocuparnos, principalmente porque es uno de los principales sumideros de CO2. ***[pause]***

El mar ejerce de termostato fundamental // y las cuestiones relacionadas con el mar son más de voluntad política que de carencias de recursos. También se realizarán programas de inversión para la mejora del conocimiento del medio natural, para la restauración de los ecosistemas de algunos de los lugares más representativos de nuestro territorio. ***[pause]***

Con respecto a la eficiencia energética, según el IDEA, más de la mitad de las viviendas de España reciben una calificación energética demasiado baja. Como saben, el aumento del riesgo de incendios determinado por el cambio climático es exponencial. Nuestros medios, la anticipación, el reforzamiento de los servicios de extinción y Protección Civil es una prioridad indiscutible. ***[pause]***

Los expertos achacan a la falta de regulación, el descontrol que se ha producido en los servicios de extinción de los incendios forestales. // En el ámbito de los recursos hídricos, queremos apostar decididamente por los procedimientos que nos exige la Unión Europea de modernización, gestión, vigilancia y sanción. ***[pause]***

Sin embargo, desde el realismo, somos conscientes de las limitaciones de los procesos de depuración de nuestro país. Por ello estamos trabajando con científicos de gran nivel para ir modernizando las fases de percolación y floculación. Optimizar el consumo del agua en toda España resulta esencial si se quiere garantizar el suministro futuro en un territorio tan expuesto a las sequías. ***[pause]***

Las instalaciones urbanas comprenden un complejo sistema de plantas potabilizadoras que no vemos, pero que necesitan inversión y mantenimiento. En España, el gobierno del agua depende de los municipios, que, en su mayoría son muy pequeños y disponen de sistemas anticuados // en los que no se está midiendo la pérdida real de agua. ***[pause]***

Insistimos en que es necesario renovar las infraestructuras y desplegar la tecnología necesaria para calcular de manera fiable el consumo de agua, sin dejar nada por tocar. Otro reto al que hacía referencia al inicio de mi intervención, no menos importante, es el denominado reto demográfico. Creemos que se trata de una cuestión de calidad de la democracia. ***[pause]***

Con toda modestia, pienso que es fundamental impulsar políticas para la cohesión territorial y social: desde la rebaja de las facturas eléctricas hasta el aumento de las nóminas. Con estos presupuestos, el Gobierno se compromete de manera decidida para transformar el modelo energético en España y conseguir la neutralidad climática. Será una oportunidad para la creación de empleo verde. // ***[pause]***

Señorías, con independencia de nuestra orientación política, ¡qué ilusión sería ver a nuestros sobrinos, a nuestros hijos y a nuestros nietos vivir en un mundo más respirable! Señorías, queremos un país verde, un país sostenible y muestra de ello es lo que hemos llevado a cabo a lo largo de la pasada legislatura, nosotros queremos logros posibles. ***[pause]***

El Gobierno socialista propone instalar 102.500 megavatios nuevos de energías renovables en 7 años solo en la Comunidad Valenciana. Solo en esta Comunidad el Gobierno socialista propone instalar 14.650 megavatios nuevos al año durante siete años. Es decir, el Gobierno socialista propone aumentar un 23% la producción de energías renovables en solo siete años en la Comunidad Valenciana**.** A lo largo de la legislatura socialista hemos definido los objetivos de reducción de emisiones // de gases de efecto invernadero, de penetración de energías renovables y de eficiencia energética. ***[pause]***

Para el año 2030, en España, se prevé una potencia total instalada en el sector eléctrico de 157 gigavatios. El sector de movilidad y transporte reducirá sus emisiones en 28 millones de toneladas de CO2 equivalente antes de 2027, en toda nuestra península. Antes de que acabe 2019, solo en la región de Murcia, habrá un ahorro de energía acumulado en una década de más de 1.300 kilotoneladas equivalentes de petróleo.

En este recorrido, tenemos al conjunto de la ciudadanía a nuestro lado. Muchas gracias.

Il discorso A.1 è stato redatto seguendo la metodologia illustrata in §2.4.1; pertanto, il testo è stato scritto dal ricercatore in funzione degli obiettivi della ricerca. Per la redazione del discorso sono state utilizzate strutture tipiche del genere comunicativo dei testi orali pronunciati presso il *Congreso de los diputados* e alcuni elementi contenutistici del discorso sono stati tratti da un discorso della Ministra Teresa Ribera, disponibile al seguente link: https://rb.gy/ff6mjg. Nel discorso sono tuttavia presenti informazioni tratte da altre fonti e da altri discorsi.

Appendice II. Discorso B.1 con istruzioni per la lettura.

Istruzioni per la lettura:

1. Leggi a una velocità di circa 127 parole al minuto.

2. Fai una pausa di 3 secondi circa [equivalente a un respiro non troppo prolungato] quando vedi questo simbolo nel discorso: [pause].

3. Utilizza un andamento prosodico naturale e spontaneo, cercando di utilizzare una pausa naturale alla fine di un'idea o di un paragrafo. La velocità deve essere rilassata e spontanea, senza rallentamenti o accelerazioni.

4. Il simbolo // segna il passaggio da un minuto all'altro. Tendenzialmente, ogni minuto dovresti arrivare a un //. (// NON EQUIVALE A UNA PAUSA).

5. Inquadratura dell'oratrice a mezzo busto, non solo faccia. Sfondo totalmente bianco (muro). L'oratrice non deve fare gesti durante la lettura. L'oratrice deve preferibilmente utilizzare una cuffia con microfono.

Durata: 9 minuti e 10 secondi circa

Parole al minuto: 127

Parole totali: 1168

Discorso:

Muchas gracias, presidenta, buenas tardes, señorías. La casualidad ha querido que hoy sea también el aniversario de la toma de la Bastilla, hito esencial de la Revolución francesa, considerada el inicio del Estado Moderno, el triunfo del pluralismo, los principios republicanos de libertad, igualdad y fraternidad.

Señorías, les insto a que trabajemos juntos para llegar a la aprobación de la ley de cambio climático y transición energética, es una ley que aspira a transformar nuestra sociedad. Con esta ley, queremos asentar las bases de un futuro más justo e inclusivo. Reivindicamos más libertad para innovar sobre la base del conocimiento. ***[pause]***

En los últimos años las desigualdades no han hecho sino aumentar; ahora, queremos más igualdad de oportunidades y en el acceso a beneficios ambientales, a través de una// política todo lo arriesgada que se quiera, pero necesaria. Queremos más fraternidad y solidaridad, queremos una revolución no cruenta que ha de permitir sentar las bases para la España del siglo XXI que piden nuestros ciudadanos. ***[pause]***

Todos nosotros hemos visto las protestas de estos últimos días de los vecinos de la Ciudad de Madrid y no podemos quedar en silencio ante tanto enfado. Aunque, afortunadamente, son minoría, todavía en nuestro país hay quienes no escuchan a la ciencia, que desde hace décadas viene advirtiendo de las causas y las consecuencias del cambio climático. ***[pause]***

Señorías, España ha importado combustibles fósiles por un valor de 46.570 millones de euros solo el año pasado, una cantidad demasiado elevada. Es sorprendente que no escuchen tampoco a los ciudadanos, ya sea // a los jóvenes en las calles o a la sociedad civil organizada, a los sindicatos o al consenso internacional. ***[pause]***

La electrificación es la clave para la sostenibilidad en el tiempo. Se trata de que el Congreso de los diputados coopere en esta transformación aprobando las medidas inevitables hacia la descarbonización. No, señorías, no hay alarmismo climático ni se trata de una nueva religión. Es pura física y química, ajenas a las ideologías y a las religiones. ***[pause]***

No, señorías, las previsiones de la CMNUCC ni son alarmistas ni son mentirosas. Al contrario, lamentablemente se ven confirmadas por los hechos. Tenemos que tomarnos en serio la amenaza climática, aún más en un periodo crítico para nuestro país en el que estamos debatiendo juntos qué futuro queremos. ***[pause]***

La fase de // adaptación y mitigación en la que estamos representa un momento fundamental. Y no está de más hablar de las claves necesarias para contestar a las grandes preguntas que nos plantea la transición ecológica. Este proyecto de ley responde a las alertas de la ciencia y a las demandas de la sociedad; se trata de un proyecto de ley muy participado, sólido, completo y equilibrado. ***[pause]***

Señor Francisco José Contreras, no podemos ignorar los datos fruto del conocimiento científico, como usted hace con su negacionismo; al contrario, debemos integrarlas en la toma de decisión a tiempo y de manera eficaz. Es fundamental debatir asuntos concretos, con políticas concretas, con límites y obligaciones. La calidad del aire está empeorando, por esto proponemos una reducción drástica en el número de vuelos. // ***[pause]***

Desde el PSOE, proponemos en 2 años reducir el número de vuelos eliminando 25.000 vuelos cortos en toda España. Con ello, en España, durante 2 años queremos trabajar desde el Gobierno socialista, para eliminar la emisión de 200.000 toneladas de CO2; lo que en 2 años nos permitirá ahorrar cerca de 38 millones de euros en España, gracias al Gobierno socialista.

No podemos separar las decisiones que repercuten en nuestra salud y bienestar de las que afectan al medio ambiente, no podemos dejarnos vencer por opciones que ignoren costes medioambientales. ***[pause]***

Señorías, les voy a dar un ejemplo: justamente a lo largo de la pasada legislatura se emitieron más de 56,5 millones de toneladas de CO2 al año, y no se adoptaron medidas adecuadas más que en algunos sectores. // Además, debido a la importación de combustibles fósiles, se gastaron más de 120 millones de euros diarios. Ahora tenemos que aferrarnos a nuestros objetivos sostenibles. Por eso, queremos apostar por un cambio de modelo productivo, energético. Queremos un modelo productivo que genere empleo y reduzca las desigualdades, que nos permita vivir en ciudades más habitables. ***[pause]***

Recientemente me he estado ocupando de la búsqueda de soluciones para que la acción del Gobierno tenga un verdadero papel transformador en esta fase de mitigación, siguiendo los pronósticos de los expertos. Apostar por la acción climática es apostar por nuestra economía y por el progreso social. Distintos actores económicos, empresas, organismos internacionales resaltan que es imperativa una recuperación económica verde e inclusiva. ***[pause]***

El fondo del MRR es un instrumento fundamental para fomentar reformas // y efectuar inversiones en los Estados miembros de la UE por un valor total de 723 800 millones de euros. Ejemplo de que las propuestas europeas trabajan en nuestra misma dirección. Dejar atrás los combustibles fósiles y sustituirlos por energías renovables genera empleo cualificado y fortalece nuestro tejido empresarial e industrial. ***[pause]***

Es fundamental, promover la ecologización de la industria, para que haya desarrollo sostenible y trabajo decente a la vez. Las energías renovables ayudarán además a hacer frente a los retos de la España vaciada y a la despoblación de las zonas rurales; la correcta gestión del agua contribuirá en la generación de riqueza. ***[pause]***

La inversión en infraestructuras que requiere la movilidad sostenible, la inversión en los vehículos compartidos y en la reducción de la circulación rodada,// son cuestiones que se tienen que abordar a la mayor brevedad posible. Gracias a la transición energética y a la expansión de las energías renovables nuestra economía será más competitiva, al disminuir el coste energético generará ahorro a las familias. ***[pause]***

Como saben, la Comisión Europea ha aprobado nuestro plan de recuperación y resiliencia, que ahora asciende a 163.200 millones de euros, compuesto por 83.000 millones en préstamos y 80.200 millones en lo que se refiere a subvenciones para la transición energética. Además, los planes de sostenibilidad compatibles con la lucha contra el cambio climático crean más empleo. Esta ley nos ayudará a salir de la crisis, mejorando el futuro de los jóvenes. ***[pause]***

En 2030 proponemos que España haya reducido sus emisiones al menos un 20 % // con respecto a las de 1990, alcanzando los 40 gramos de dióxido de carbono equivalentes por megajulio, tal y como nos recomienda la AIE, tenemos 7 años. Señorías, con esta ley se quiere apoyar la senda de desarrollo e implementación del hidrógeno verde. ¡Queremos ser un líder emblemático y un referente europeo en hidrógeno! ***[pause]***

Antes de 2028, el 20% de los proyectos en hidrógeno renovable a escala mundial se ubicarán en España. La Asociación AEH2 acaba de afirmar que programas de, al menos, 21.000 millones de euros han sido anunciados en Andalucía en los últimos 4 meses.

Señorías, ya somos un referente a nivel planetario: los investigadores del CSIC de nuestro país han desarrollado en el mes de agosto un método para producir esta energía limpia // usando 10 veces menos iridio del que era necesario hasta ahora.

Señorías, esta es la revolución que queremos desde el Gobierno socialista. Muchas gracias.

Anche il discorso B.1 è stato redatto seguendo la metodologia illustrata in §2.4.1; pertanto, il testo è stato scritto dal ricercatore in funzione degli obiettivi della ricerca. Per la redazione del discorso sono state utilizzate strutture tipiche del genere comunicativo dei testi orali pronunciati presso il *Congreso de los diputados* e alcuni elementi contenutistici del discorso sono stati tratti da un discorso della Ministra Teresa Ribera, disponibile al seguente link: https://rb.gy/2zjb1r. Nel discorso sono tuttavia presenti informazioni tratte da altre fonti e da altri discorsi.

Appendice III. Organizzazione dei *task* nel discorso A.1.

Funzione del cluster	Task	Nome Task	Definizione Task	Testo	Stimoli Isolati	Grado di complessità (1/2/3). (1=facile)
Introduzione di 45 parole *(warm-up)*				Hoy subo a esta tribuna para defender el proyecto de ley de Presupuestos Generales del Estado en lo relativo a la sección del Ministerio para la Transición Ecológica y el Reto Demográfico. Los presupuestos que presentamos se alinean con la urgente necesidad de reactivar la economía. *[pause]*		
Target Sentence	NRS	Numerali + referente complesso + dissimmetria superficiale	Task complesso con 3 numeri complessi, referente complesso (acronimo, entità nominale o termine specialistico) e struttura che richiede dominanza di elaborazione superficiale.	El presupuesto asciende en su conjunto a 10 195 millones de euros, de los que 5817 corresponden a recursos del presupuesto nacional y 4378 corresponden a recursos del plan España Puede, tal y como les explicaré extensa y detalladamente.	10 195 millones, 5817 millones, 4378 millones//plan España Puede // extensa y detalladamente	3
Control Sentence				Este presupuesto supone una movilización sin precedentes en relación con tres de los grandes retos que tiene el país: la transformación del sistema energético, la respuesta a la emergencia climática y la cohesión territorial. *[pause]*		
Target Sentence	NI	Numerale isolato	Numerale isolato (ordine di grandezza: miliardi, con 4 cifre).	El primer gran reto, relativo a la transición energética, cuenta con un presupuesto de 7397 millones de euros.	7397 millones	1
Control Sentence				Con este presupuesto, el Gobierno de España quiere mostrar su firme determinación en su voluntad de asumir un papel de liderazgo en esta transición a nivel europeo. *[pause]*		
Target Sentence	NLP	Numerali + sfida lessicale + dissimmetria profonda	Task complesso con 2 cluster e ogni cluster deve contenere: (1) 1 numerale (ordine di grandezza: milioni); (2) 1 sfida lessicale e (3) 1 struttura che richiede dominanza di elaborazione profonda.	El presupuesto destinado al Instituto para la Transición Justa crecerá hasta 442 millones, gracias al Mecanismo de Recuperación, que no deja de ser trascendental en la política de nuestro Gobierno. En el presupuesto de este instituto resalta el programa de 53 millones para proyectos generadores de empleo, que siempre hemos defendido los socialistas, en nuestro constante compromiso por la solidaridad social.	Cluster 1: 442 millones//no deja de ser// trascendental; Cluster 2: 53 millones// hemos…los socialistas//compromiso	3
Control Sentence				Tampoco nos podemos olvidar de que la recuperación debe ser justa con los consumidores, sobre todo con los consumidores más vulnerables: los que tienen que luchar por llegar a fin de mes. *[pause]*		

Target Sentence	SI	Dissimmetria superficiale isolata	Task semplice con struttura asimmetrica nella combinazione italiano-spagnolo che richiede dominanza di elaborazione superficiale in una proposizione semplice (20-30 parole max, sintassi semplice, coesione logica esplicita e assenza di ulteriori trigger).	Nuestro compromiso frente a la pobreza energética queremos reforzarlo en este año complicado. Quiero poner en valor el bono social térmico, el cual se ha sometido a una comisión de expertos.	se ha sometido	1
Control Sentence				Además, las medidas para la eficiencia energética y la promoción de comunidades energéticas son el pilar, la base de la transformación, sirven para reducir el consumo de energía. *[pause]*		
Target Sentence	RN	Referente complesso + numerale	Referente complesso (acronimo/entità isolata/termine specialistico) e 1 numerale (5 cifre, ordine di grandezza: miliardi)	Por eso, estas políticas contarán con 1233 millones de euros que se pueden financiar con fondos FEDER.	1233 millones//FEDER	2
Control Sentence				Como pueden observar, el esfuerzo en materia energética de estos presupuestos es enormemente relevante; servirá para afrontar el reto de la transición hacia una energía limpia, eficiente, sostenible. *[pause]*		
Target Sentence	EI	Entità nominale isolata	Entità nominale isolata in proposizione semplice.	De hecho, cabe destacar el gran éxito que ha tenido durante el periodo de apertura iniciado apenas hace diez días y que sigue abierto del Programa DUS 5000.	Programa DUS 5000	1
Control Sentence				El segundo gran reto al que nos enfrentamos es el de la preservación de nuestros ecosistemas, de la biodiversidad, de los recursos naturales y del territorio que ven la influencia de la acción del ser humano. *[pause]*		
Target Sentence	UN	Unità informativa con numerali	Task con unità informativa complesso di circa 30-40 parole. L'unità è composta da: (1) un referente complesso; (2) un'unità di misura complessa, 5 numerali semplici consecutivi (2 anni, 2 numerali semplici e una percentuale).	La Secretaría de Estado liderada por Morán Fernández cuenta con un presupuesto de 579 millones de euros, cifra que supera con creces la de 2021, la cual representa un 2,85 % más. En este contexto, España aspira a lograr los 22 gigavatios de capacidad antes de 2032.	Morán Fernández//gigavatios//2032//2021//22//579 millones//2,85%	3
Control Sentence				Además, el capítulo para la preservación de biodiversidad marina tiene un presupuesto de casi medio millón de euros. El mar debe preocuparnos, principalmente porque es uno de los principales sumideros de CO2. *[pause]*		
Target Sentence	PI	Dissimmetria profonda isolata	Task semplice con struttura che richiede dominanza di elaborazione profonda in proposizione semplice.	El mar ejerce de termostato fundamental y las cuestiones relacionadas con el mar son más de voluntad política que de carencias de recursos.	son más de…que de…	1
Control Sentence				También se realizarán programas de inversión para la mejora del conocimiento del medio natural, para la restauración de los ecosistemas de algunos de los lugares más representativos de nuestro territorio. *[pause]*		

Target Sentence	AI	Acronimo isolato	Acronimo isolato in proposizione semplice.	Con respecto a la eficiencia energética, según el IDAE *[Instituto para la Diversificación y Ahorro de la Energía]* , más de la mitad de las viviendas de España reciben una calificación energética demasiado baja.	IDEA	1
Control Sentence				Como saben, el aumento del riesgo de incendios determinado por el cambio climático es exponencial. Nuestros medios, la anticipación, el reforzamiento de los servicios de extinción y Protección Civil es una prioridad indiscutible. *[pause]*		
Target Sentence	LI	Sfida lessicale isolata	Sfida lessicale isolata in proposizione semplice.	Los expertos achacan a la falta de regulación, el descontrol que se ha producido en los servicios de extinción de los incendios forestales.	achacan	1
Control Sentence				En el ámbito de los recursos hídricos, queremos apostar decididamente por los procedimientos que nos exige la Unión Europea de modernización, gestión, vigilancia y sanción. *[pause]*		
Target Sentence	TP	Tecnicismi + dissimmetria profonda	Due tecnicismi in una frase semanticamente complessa con una struttura che richiede elaboraziona profonda del testo.	Sin embargo, desde el realismo, somos conscientes de las limitaciones de los procesos de depuración de nuestro país. Por ello estamos trabajando con científicos de gran nivel para ir modernizando las fases de percolación y floculación.	desde el realismo//percolación//floculación	2
Control Sentence				Optimizar el consumo del agua en toda España resulta esencial si se quiere garantizar el suministro futuro en un territorio tan expuesto a las sequías. *[pause]*		
Target Sentence	TI	Tecnicismo isolato	Tecnicismo isolato in proposizione semplice.	Las instalaciones urbanas comprenden un complejo sistema de plantas potabilizadoras que no vemos, pero que necesitan inversión y mantenimiento.	plantas potabilizadoras	1
Control Sentence				En España, el gobierno del agua depende de los municipios, que, en su mayoría son muy pequeños y disponen de sistemas anticuados en los que no se está midiendo la pérdida real de agua. *[pause]*		
Target Sentence	SP	Dissimmetria superficiale + dissimmetria profonda	Task complesso che con struttura che richiede dominanza di elaborazione superficiale e con struttura che richiede dominanza di elaborazione profonda in proposizione semplice.	Insistimos en que es necesario renovar las infraestructuras y desplegar la tecnología necesaria para calcular de manera fiable el consumo de agua, sin dejar nada por tocar.	Insistimos en que…//sin dejar nada por tocar	2

Control Sentence				Otro reto al que hacía referencia al inicio de mi intervención, no menos importante, es el denominado reto demográfico. Creemos que se trata de una cuestión de calidad de la democracia. *[pause]*		
Target Sentence	LS	Sfida lessicale + dissimmetria superficiale	Task con sfida lessicale e struttura che richiede dominanza di elaborazione superficiale in proposizione semplice.	Con toda modestia, pienso que es fundamental impulsar políticas para la cohesión territorial y social: desde la rebaja de las facturas eléctricas hasta el aumento de las nóminas.	Con toda modestia…//nóminas	2
Control Sentence				Con estos presupuestos, el Gobierno se compromete de manera decidida para transformar el modelo energético en España y conseguir la neutralidad climática. Será una oportunidad para la creación de empleo verde. *[pause]*		
Target Sentence	LP	Sfida lessicale + dissimmetria profonda	Task con sfida lessicale e con struttura che richiede dominanza di elaborazione profonda in proposizione semplice.	Señorías, con independencia de nuestra orientación política, ¡qué ilusión sería ver a nuestros sobrinos, a nuestros hijos y a nuestros nietos vivir en un mundo más respirable!	con independencia de…//ilusión	2
Control Sentence				Señorías, queremos un país verde, un país sostenible y muestra de ello es lo que hemos llevado a cabo a lo largo de la pasada legislatura, nosotros queremos logros posibles. *[pause]*		
Target Sentence	NRI	Cluster con numerali ridondanti	Una parte del discorso numericamente densa costituita da 3 microunità informative. Ogni microunità contiene: (1) un deittico spaziale e un deittico temporale (che rimangono gli stessi in tutte e tre le unità e si ripetono); (2) almeno un'unità di misura; (3) un numerale che cambia in ciascuna delle tre unità; (4) un referente che rimane lo stesso in tutti e tre i cluster.	El Gobierno socialista propone instalar 102.500 megavatios nuevos de energías renovables en 7 años solo en la Comunidad Valenciana. Solo en esta Comunidad el Gobierno socialista propone instalar 14.650 megavatios nuevos al año durante siete años. Es decir, el Gobierno socialista propone aumentar un 23% la producción de energías renovables en solo siete años en la Comunidad Valenciana.	Cluster 1: Gobierno socialista//102.500 megavatios//en 7 años//Comunidad Valenciana. Cluster 2: esta Comunidad//Gobierno socialista//14.650 megavatios//durante siete años. Cluster 3: Gobierno socialista//23%//en un periodo de solo 7 años//Comunidad Valenciana.	3
Control Sentence				A lo largo de la legislatura socialista hemos definido los objetivos de reducción de emisiones de gases de efecto invernadero, de penetración de energías renovables y de eficiencia energética. *[pause]*		

Target Sentence	NNR	Cluster con numerali non ridondanti	Task complesso e numericamente denso con le seguenti caratteristiche: (1) 3 microunità informative; (2) tempo, luogo, referente e unità di misura e numerale cambiano in ogni microunità informativa; (3) o il referente o l'unità di misura devono essere complessi in ogni microunità.	Para el año 2030, en España, se prevé una potencia total instalada en el sector eléctrico de 157 gigavatios. El sector de movilidad y transporte reducirá sus emisiones en 28 millones de toneladas de CO2 equivalente [MtCO2-eq] antes de 2027, en toda nuestra península. Antes de que acabe 2019, solo en la región de Murcia, habrá un ahorro de energía acumulado en una década de más de 1.300 kilotoneladas equivalentes de petróleo [ktep].	Cluster 1: para el año 2030//España//potencia total instalada//157 gigavatios Cluster 2: sector de movilidad y transporte//28 MtCO2-eq//antes de 2027//en toda nuestra península Cluster 3: antes de que acabe 2019//Murcia//ahorro de energía acumulado//una década//1.300 ktep	3
Cluster conclusivo				En este recorrido, tenemos al conjunto de la ciudadanía a nuestro lado. Muchas Gracias		

Appendice IV. Organizzazione dei *task* nel discorso B.1.

Funzione del cluster	Task	Nome Task	Definizione Task	Testo	Stimoli Isolati	Grado di complessità (1/2/3). (1=facile)
Introduzione di 45 parole *(warm-up)*				Muchas gracias, presidenta, buenas tardes, señorías. La casualidad ha querido que hoy sea también el aniversario de la toma de la Bastilla, hito esencial de la Revolución francesa, considerada el inicio del Estado Moderno, el triunfo del pluralismo, los principios republicanos de libertad, igualdad y fraternidad.		
Target Sentence	SI	Dissimmetria superficiale isolata	Task semplice con struttura asimmetrica nella combinazione italiano-spagnolo che richiede dominanza di elaborazione superficiale in una proposizione semplice (20-30 parole max, sintassi semplice, coesione logica esplicita e assenza di ulteriori trigger).	Señorías, les insto a que trabajemos juntos para llegar a la aprobación de la ley de cambio climático y transición energética, es una ley que aspira a transformar nuestra sociedad.	les insto a que…	1
Control Sentence				Con esta ley, queremos asentar las bases de un futuro más justo e inclusivo. Reivindicamos más libertad para innovar sobre la base del conocimiento. *[pause]*		
Target Sentence	SP	Dissimmetria superficiale + dissimmetria profonda	Task complesso che con struttura che richiede dominanza di elaborazione superficiale e con struttura che richiede dominanza di elaborazione profonda in proposizione semplice.	En los últimos años las desigualdades no han hecho sino aumentar; ahora, queremos más igualdad de oportunidades y en el acceso a beneficios ambientales, a través de una política todo lo arriesgada que se quiera, pero necesaria.	no han hecho sino aumentar//todo lo arriesgada que…	2

Control Sentence				Queremos más fraternidad y solidaridad, queremos una revolución no cruenta que ha de permitir sentar las bases para la España del siglo XXI que piden nuestros ciudadanos. *[pause]*		
Target Sentence	LI	Sfida lessicale isolata	Sfida lessicale isolata in proposizione semplice.	Todos nosotros hemos visto las protestas de estos últimos días de los vecinos de la Ciudad de Madrid y no podemos quedar en silencio ante tanto enfado.	vecinos	1
Control Sentence				Aunque, afortunadamente, son minoría, todavía en nuestro país hay quienes no escuchan a la ciencia, que desde hace décadas viene advirtiendo de las causas y las consecuencias del cambio climático. *[pause]*		
Target Sentence	NI	Numerale isolato	Numerale isolato (ordine di grandezza: miliardi, con 4 cifre).	Señorías, España ha importado combustibles fósiles por un valor de 46.570 millones de euros solo el año pasado, una cantidad demasiado elevada.	46.570 millones	1
Control Sentence				Es sorprendente que no escuchen tampoco a los ciudadanos, ya sea a los jóvenes en las calles o a la sociedad civil organizada, a los sindicatos o al consenso internacional. *[pause]*		
Target Sentence	PI	Dissimmetria profonda isolata	Task semplice con struttura che richiede dominanza di elaborazione profonda in proposizione semplice.	La electrificación es la clave para la sostenibilidad en el tiempo. Se trata de que el Congreso de los diputados coopere en esta transformación aprobando las medidas inevitables hacia la descarbonización.	se trata de que…	1
Control Sentence				No, señorías, no hay alarmismo climático ni se trata de una nueva religión. Es pura física y química, ajenas a las ideologías y a las religiones. *[pause]*		
Target Sentence	AI	Acronimo isolato	Acronimo isolato in proposizione semplice	No, señorías, las previsiones de la CMNUCC *[Convención Marco de las Naciones Unidas sobre el Cambio Climático]* ni son alarmistas ni son mentirosas. Al contrario, lamentablemente se ven confirmadas por los hechos.	CMNUCC	1
Control Sentence				Tenemos que tomarnos en serio la amenaza climática, aún más en un periodo crítico para nuestro país en el que estamos debatiendo juntos qué futuro queremos. *[pause]*		
Target Sentence	LP	Sfida lessicale + dissimmetria profonda	Task con sfida lessicale e con struttura che richiede dominanza di elaborazione profonda in proposizione semplice.	La fase de adaptación y mitigación en la que estamos representa un momento fundamental. Y no está de más hablar de las claves necesarias para contestar a las grandes preguntas que nos plantea la transición ecológica.	no está de más hablar//contestar	2

Control Sentence				Este proyecto de ley responde a las alertas de la ciencia y a las demandas de la sociedad; se trata de un proyecto de ley muy participado, sólido, completo y equilibrado. *[pause]*		
Target Sentence	EI	Entità nominale isolata	Entità nominale isolata in proposizione semplice.	Señor Francisco José Contreras, no podemos ignorar los datos fruto del conocimiento científico, como usted hace con su negacionismo; al contrario, debemos integrarlas en la toma de decisión a tiempo y de manera eficaz.	Francisco José Contreras	1
Control Sentence				Es fundamental debatir asuntos concretos, con políticas concretas, con límites y obligaciones. La calidad del aire está empeorando, por esto proponemos una reducción drástica en el número de vuelos. *[pause]*		
Target Sentence	NRI	Cluster con numerali ridondanti	Una parte del discorso numericamente densa costituita da 3 microunità informative. Ogni microunità contiene: (1) un deittico spaziale e un deittico temporale (che rimangono gli stessi in tutte e tre le unità e si ripetono); (2) almeno un'unità di misura; (3) un numerale che cambia in ciascuna delle tre unità; (4) un referente che rimane lo stesso in tutti e tre i cluster.	Desde el PSOE, proponemos en 2 años reducir el número de vuelos eliminando 25.000 vuelos cortos en toda España. Con ello, en España, durante 2 años queremos trabajar desde el Gobierno socialista, para eliminar la emisión de 200.000 toneladas de CO2; lo que en 2 años nos permitirá ahorrar cerca de 38 millones de euros en España, gracias al Gobierno socialista.	Cluster 1: PSOE//2 años//25000//vuelos cortos//toda España; Cluster 2: en España//durante 2 años//Gobierno socialista//emisión//200.000 toneladas de CO2; Cluster 3: en 2 años//ahorrar//cerca de 38 millones de euros//en España//Gobierno socialista	3
Control Sentence				No podemos separar las decisiones que repercuten en nuestra salud y bienestar de las que afectan al medio ambiente, no podemos dejarnos vencer por opciones que ignoren costes medioambientales. *[pause]*		
Target Sentence	NLP	Numerali + sfida lessicale + dissimmetria profonda	Task complesso con 2 cluster e ogni cluster deve contenere: (1) 1 numerale (ordine di grandezza: milioni); (2) 1 sfida lessicale e (3) 1 struttura che richiede dominanza di elaborazione profonda.	Señorías, les voy a dar un ejemplo: justamente a lo largo de la pasada legislatura se emitieron más de 56,5 millones de toneladas de CO2 al año, y no se adoptaron medidas adecuadas más que en algunos sectores. Además, debido a la importación de combustibles fósiles, se gastaron más de 120 millones de euros diarios. Ahora tenemos que aferrarnos a nuestros objetivos sostenibles.	Cluster 1: justamente//56,5 millones//no se…más que… Cluster 2: debido a//120 millones//aferrarnos	3
Control Sentence				Por eso, queremos apostar por un cambio de modelo productivo, energético. Queremos un modelo productivo que genere empleo y reduzca las desigualdades, que nos permita vivir en ciudades más habitables. *[pause]*		
Target Sentence	LS	Sfida lessicale + dissimmetria superficiale	Task con sfida lessicale e struttura che richiede dominanza di elaborazione superficiale in proposizione semplice.	Recientemente me he estado ocupando de la búsqueda de soluciones para que la acción del Gobierno tenga un verdadero papel transformador en esta fase de mitigación, siguiendo los pronósticos de los expertos.	me he estado ocupando de…//pronósticos	2

Control Sentence				Apostar por la acción climática es apostar por nuestra economía y por el progreso social. Distintos actores económicos, empresas, organismos internacionales resaltan que es imperativa una recuperación económica verde e inclusiva. *[pause]*		
Target Sentence	RN	Referente complesso + numerale	Referente complesso (acronimo/entità isolata/termine specialistico) e 1 numerale (5 cifre, ordine di grandezza: miliardi)	El fondo del MRR *[Mecanismo de Recuperación y Resiliencia]* es un instrumento fundamental para fomentar reformas y efectuar inversiones en los Estados miembros de la UE por un valor total de 723 800 millones de euros.	MRR//723.800 millones	2
Control Sentence				Ejemplo de que las propuestas europeas trabajan en nuestra misma dirección. Dejar atrás los combustibles fósiles y sustituirlos por energías renovables genera empleo cualificado y fortalece nuestro tejido empresarial e industrial. *[pause]*		
Target Sentence	TI	Tecnicismo isolato	Tecnicismo isolato in proposizione semplice.	Es fundamental, promover la ecologización de la industria, para que haya desarrollo sostenible y trabajo decente a la vez.	ecologización	1
Control Sentence				Las energías renovables ayudarán además a hacer frente a los retos de la España vaciada y a la despoblación de las zonas rurales; la correcta gestión del agua contribuirá en la generación de riqueza. *[pause]*		
Target Sentence	TP	Tecnicismi + dissimmetria profonda	Due tecnicismi in una frase semanticamente complessa con una struttura che richiede elaboraziona profonda del testo.	La inversión en infraestructuras que requiere la movilidad sostenible, la inversión en los vehículos compartidos y en la reducción de la circulación rodada, son cuestiones que se tienen que abordar a la mayor brevedad posible.	vehículos compartidos//circulación rodada//a la mayor brevedad posible	2
Control Sentence				Gracias a la transición energética y a la expansión de las energías renovables nuestra economía será más competitiva, al disminuir el coste energético generará ahorro a las familias. *[pause]*		
Target Sentence	NRS	Numerali + referente complesso + dissimmetria superficiale	Task complesso con 3 numeri complessi, referente complesso (acronimo, entità nominale o termine specialistico) e struttura che richiede dominanza di elaborazione superficiale.	Como saben, la Comisión Europea ha aprobado nuestro plan de recuperación y resiliencia, que ahora asciende a 163.200 millones de euros, compuesto por 83.000 millones en préstamos y 80.200 millones en lo que se refiere a subvenciones para la transición energética.	plan de recuperación y resiliencia//163.200 millones//83.000 millones//80.200 millones//en lo que se refiere a…	3
Control Sentence				Además, los planes de sostenibilidad compatibles con la lucha contra el cambio climático crean más empleo. Esta ley nos ayudará a salir de la crisis, mejorando el futuro de los jóvenes. *[pause]*		

Target Sentence	UN	Unità informativa con numerali	Task con unità informativa complesso di circa 30-40 parole. L'unità è composta da: (1) un referente complesso; (2) un'unità di misura complessa, 5 numerali semplici consecutivi (2 anni, 2 numerali semplici e una percentuale).	En 2030 proponemos que España haya reducido sus emisiones al menos un 20 % con respecto a las de 1990, alcanzando los 40 gramos de dióxido de carbono equivalentes por megajulio *[gCO2eq/MJ]* , tal y como nos recomienda la AIE *[Agencia Internacional de la Energía]* . Tenemos 7 años.	AIE//gCO2eq/MJ//1990//2030//40//20%//7	3
Control Sentence				Señorías, con esta ley se quiere apoyar la senda de desarrollo e implementación del hidrógeno verde. ¡Queremos ser un líder emblemático y un referente europeo en hidrógeno! *[pause]*		
Target Sentence	NNR	Cluster con numerali non ridondanti	Task complesso e numericamente denso con le seguenti caratteristiche: (1) 3 microunità informative; (2) tempo, luogo, referente e unità di misura e numerale cambiano in ogni microunità informativa; (3) o il referente o l'unità di misura devono essere complessi in ogni microunità.	Antes de 2028, el 20% de los proyectos en hidrógeno renovable a escala mundial se ubicarán en España. La Asociación AEH2 *[Asociación Española del Hidrógeno]* acaba de afirmar que programas de, al menos, 21.000 millones de euros han sido anunciados en Andalucía en los últimos 4 meses. Señorías, ya somos un referente a nivel planetario: los investigadores del CSIC *[Consejo Superior de Investigaciones Científicas]* de nuestro país han desarrollado en el mes de agosto un método para producir esta energía limpia usando 10 veces menos iridio del que era necesario hasta ahora.	Cluster 1: antes de 2028//España//20%//proyectos en hidrógeno renovable Cluster 2: Asociación AEH2//programas//20.000 millones//Andalucía//4 meses Cluster 3: CSIC//en el mes de Agosto//10 veces menos//iridio	3
Cluster conclusivo				Señorías, esta es la revolución que queremos desde el Gobierno socialista. Muchas gracias.		

Appendice V. Descrizione della formazione asincrona erogata alle partecipanti.

Messaggio e materiali inviati 10 giorni prima dell'attività sperimentale:

Ciao (NOME DELLA PARTECIPANTE)! 3...2...1...Via! Il (DATA) alle ore (ORARIO) ci vediamo in lab.4 con la prof.ssa (NOME DELLA RELATRICE). Ti verranno proposte due SIM dallo spagnolo all'italiano che avranno come filo conduttore il Ministero Spagnolo per la Transizione Ecologica e la sfida demografica. Ti invio la formazione per la prima sessione asincrona. Ti invio una cartella con tutto il materiale dentro. La prossima sessione te la invierò il (GIORNO). Cosa ci troverai nella cartella?

1 Glossario (vedi se utilizzarlo e come)

2 video da fare in SIM

2 trascrizioni dei video

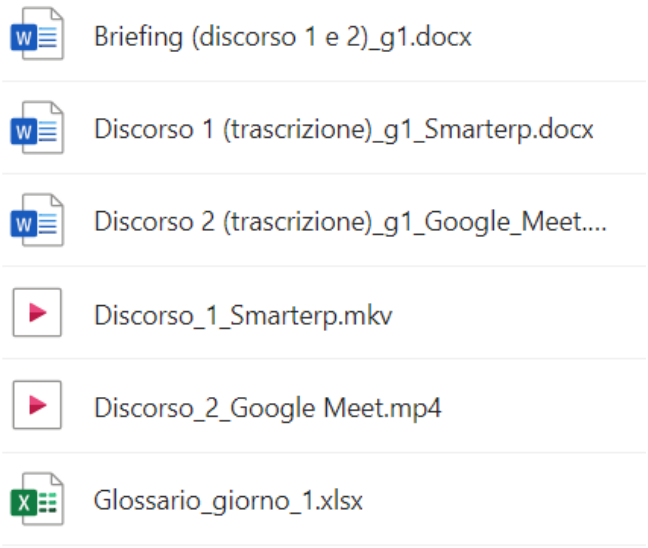
1 doc con i briefing dei video

Piccole precisazioni:

1. purtroppo nelle due formazioni l'oratore sarò io, ma le due SIM dell'esperimento vedranno come oratrice una persona spagnola;

2. durante la formazione e durante l'esperimento non potrai prendere appunti.

Se hai qualsiasi domanda, sono qui! Grazie mille per aver accettato di partecipare, davvero!



Briefing dei discorsi utilizzati nella prima sessione di formazione:

Briefing discorso 1

La ministra Teresa Ribera presenta la Legge di Bilancio per il 2023, focalizzandosi specificamente sulla Sezione 23, che riguarda il Ministero della Transizione Ecologica e la Sfida Demografica.

Link alla trascrizione del discorso: https://bitly.cx/j6Zd

Briefing discorso 2

La Ministra Teresa Ribera presenta il disegno di legge con cui si adottano misure urgenti nel settore dell'energia per promuovere la mobilità elettrica, l'autoconsumo e la diffusione delle energie rinnovabili.

Link alla trascrizione del discorso: https://bitly.cx/YN6

Messaggio e materiali inviati 4 giorni prima dell'attività sperimentale della partecipante:

Ciao (NOME DELLA PARTECIPANTE)! Grazie ancora per il tempo che stai dedicando a questo esperimento. Adesso ti invio la seconda sessione per la formazione asincrona. Il contenuto è simile alla prima sessione. Quindi nella cartella troverai:

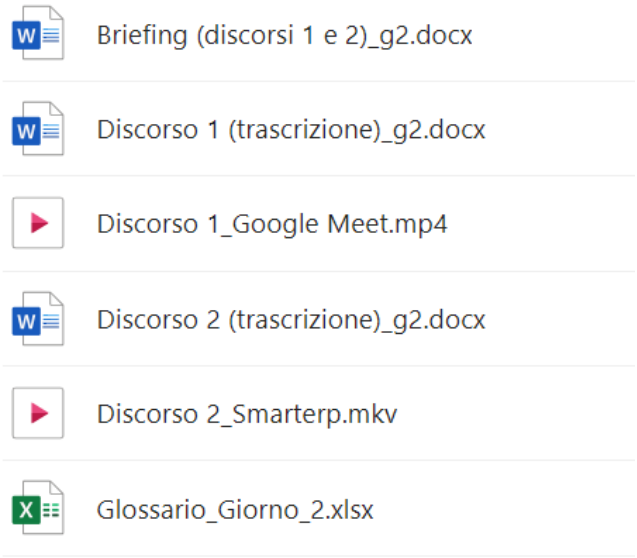
1 Glossario

2 video da fare in SIM

2 trascrizioni dei video

1 doc con i briefing dei video

Mi farò risentire il giorno prima dell'esperimento per darti qualche istruzione sull'esperimento, niente di particolare. Grazie ancora! Se hai qualsiasi domanda, sono qui. Un abbraccio.



Briefing dei discorsi utilizzati nella seconda sessione di formazione:

Briefing discorso 1

La Ministra Teresa Ribera presenta la Legge di Bilancio per il 2021, focalizzandosi specificamente sulla Sezione 23, che riguarda il Ministero della Transizione Ecologica e la Sfida Demografica.

Link alla trascrizione del discorso: https://bitly.cx/TT4

Briefing discorso 2

La Ministra Teresa Ribera presenta il Regio Decreto-Legge 23/2020, del 23 giugno, che approva misure nel campo dell'energia e in altri settori per la riattivazione economica.

Link alla trascrizione del discorso: https://bitly.cx/8RUQt

Messaggio inviato il giorno precedente all'attività sperimentale della partecipante:

Ciao (NOME PARTECIPANTE)! Domani ci vediamo alle ore (ORARIO) in lab.4. Ti contatto solo per comunicarti qualche raccomandazione. Visto che utilizzeremo l'eye-tracker, è meglio che eviti un trucco accentuato; inoltre, se porti sia gli occhiali che le lenti a contatto, ti chiederei di portare entrambi (se porti solo gli occhiali, andrà bene lo stesso). Ti invio anche i due moduli informativi che domani ti chiederò di compilare e firmare prima dell'inizio dell'esperimento. Nei moduli troverai ulteriori dettagli sul cosa comporta la partecipazione all'esperimento. Ti invio i moduli solo per darti il tempo di leggerli con calma, non importa che li stampi, ti porterò io domani due copie già stampate che puoi compilare e riempire al momento. Grazie ancora per quello che stai facendo! E non rimane che augurarti(ci) buona fortuna!

Insieme al messaggio sono stati allegati due documenti:

1. Modulo informativo relativo al trattamento dei dati personali;
2. Modulo informativo per la partecipazione – Consenso Informato.

Ringraziamenti

Ringrazio la prof.ssa Nicoletta Spinolo. Prof, ce l'abbiamo fatta. La ringrazio per aver colto ogni mia titubanza e ogni mia incertezza con la saggezza di chi sa dare un senso a ogni inquietudine. La ringrazio per aver stimolato la mia libertà creativa, nell'ascolto costante e nella concretezza della ricerca. La ringrazio per avermi fatto salire sul vagone delle domande: motore e carburante della meraviglia. La ringrazio per avermi capito nonostante le mie poche parole, per avermi fermato a riflettere ricentrandomi sugli obiettivi, per aver trovato il tempo di trascorrere interi giorni in laboratorio e di collegarsi in videochiamata anche solo per spiegarmi una sua correzione, per aver dato la sua opinione solo dopo un "e tu che faresti?". Come Kavafis, anche noi abbiamo avuto costantemente in mente Itaca, con il pensiero costante di raggiungerla, ma la ringrazio per non aver mai affrettato il viaggio, per la sua pazienza, per essere stata la prima coraggiosa lettrice delle mie bozze imperfette e per tutti i tesori che mi ha fatto vedere lungo la strada. Nonostante gli errori d'ogni sorta, sono sicuro di aver scelto la miglior compagna di viaggio con cui condividere queste miglia. Con lei e grazie a lei, per una volta, ha vinto la bellezza dello spirito universitario in cui credo. Le auguro di conservare la sua unicità.

Ringrazio la prof.ssa Mariachiara Russo per avermi orientato durante l'impostazione dello studio, riportandomi sempre ai quesiti della mia ricerca. La ringrazio per avermi accompagnato e corretto nel corso dei due anni della magistrale, con *feedback* onesti e dritti al punto, e per aver sempre inebriato l'aula con il suo pozzo di conoscenze.

Ringrazio la prof.ssa María Isabel Fernández García per essere stata protagonista anche in questa tesi, mettendo al servizio il suo dono più prezioso: la sua voce. La ringrazio per essere un costante porto sicuro, per avermi incoraggiato a navigare portando la libertà creativa della vita accademica nella mia quotidianità.

Ringrazio tutte le partecipanti, voi sapete chi siete, per aver deciso di essere il cuore pulsante di questo studio mettendo al servizio i vostri doni più preziosi: il vostro tempo e la vostra voce. Grazie per i vostri sorrisi, per la vostra sincerità, per la vostra curiosità e il vostro entusiasmo. Questo studio è un piccolo frutto della vostra bellezza.

Ringrazio Susana, Sukiko e tutta la meravigliosa squadra dietro SmarTerp, per aver trovato sempre il tempo di aiutarmi nella realizzazione dei video per l'esperimento nonostante l'agenda fitta di impegni e le mie mille domande. Vi ringrazio per aver minimizzato ogni ostacolo

burocratico e per esservi fidate nel darmi la possibilità di condurre uno studio che coinvolgesse il vostro *software*.

Ringrazio Achille per aver accompagnato questa tesi itinerante tra Spagna e Portogallo. Grazie per avermi riportato nel qui e ora nei momenti di incertezza, per avermi insegnato a perdermi per poter apprezzare le occasioni inaspettate. Grazie per avermi costantemente ribadito che la vita non è matematica, ma poesia. Grazie per le onde surfate insieme all'alta marea nei momenti di scoraggiamento.

Ringrazio Marta, Antonio e Cate per avermi fatto sentire l'odore di famiglia a Forlì, per la vostra semplice e genuina follia, per la vostra voglia di lasciare un segno nella vita degli altri, per la vostra ricerca di un senso in quello che fate, per avermi insegnato che non esiste solo l'opzione A e l'opzione B dinanzi a una scelta. Che possiate vivere con le vele gonfie di sogni e idee chiare solo quanto basta per dirigere la nave.

Ringrazio Martina e tutto il Servizio di Strada per aver dato un senso alla mia presenza a Forlì. Grazie per i giorni al mare, i viaggi lunghi, le chiacchiere belle, le cene piene di sorrisi, i riequilibri, le serate in furgone, gli scherzi nei boschi e grazie per tutti gli amici che mi avete fatto incontrare sulle panchine più dimenticate.

Ringrazio la Spagna e la sua gente, per avermi insegnato ad essere leggero senza cadere nella superficialità, per avermi regalato uno sguardo nuovo sul mondo.

Grazie a mio zio Mario, contadino silenzioso, stravagante osservatore, navigante della libertà, per avermi incoraggiato a vivere nell'essenzialità, per avermi insegnato a non giudicare, per essere innamorato della vita. Grazie a mia zia Lisa, per aver sostenuto in diversi modi la mia carriera universitaria, per la sua testarda e pura ingenuità. Grazie ai miei zii, Anna e Stefano, per la loro sincera generosità d'animo.

Grazie alle mie nonne, Elia e Angela, per avermi insegnato ad andare in bicicletta inseguendo i più nobili ideali, per amarmi ogni giorno, per avermi fatto vedere l'orizzonte, consapevoli che ci saremmo arrivati insieme, ma distanti. Nonne, mi sono laureato!

Grazie alla mia famiglia. A Giulia e Francesca, coraggiose paladine della libertà, voci giuste, per avermi insegnato ad amare la vita e a non aver paura di sbagliare, per essere state punti di riferimento, per essere un rifugio sicuro. A mamma e papà, scoppiettanti avventurieri della libertà, folli innamorati del dialogo, poeti della vita, per essere stati una porta sempre aperta. Che sia l'amore a guidarvi sempre.