

CAPITOLO II

Il *respeaking*

2.1 Introduzione

Nell'esplorare la vasta disciplina della traduzione, uno dei maggiori contributi fu apportato dal teorico James S. Holmes che per primo propose il nome di *Translation Studies* e, nel 1972, ne tracciò una mappa al fine di esemplificare i settori inerenti la ricerca traduttologica pratica e teorica. Tale mappa fu poi ripresa e modificata da diversi studiosi, tra cui Gideon Toury, mantenendo l'idea concettuale di Holmes:

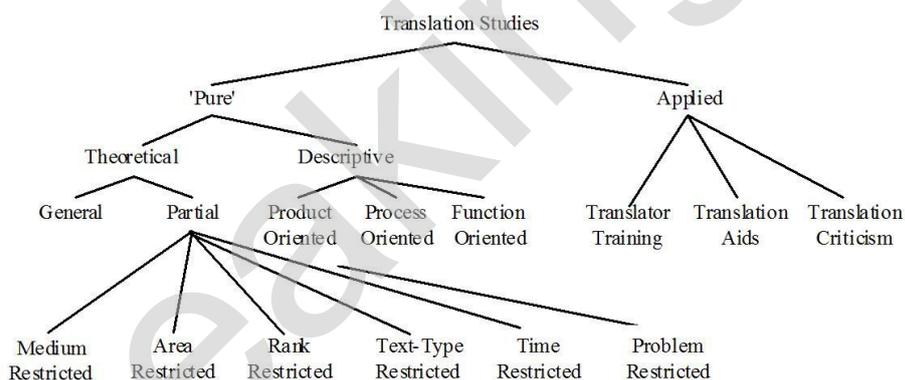


Figura 3 – Mappa dei *Translation Studies*, Holmes e Toury⁵⁶.

Se è vero che nell'elaborare questa mappa Holmes e Toury hanno preso in considerazione come oggetto di studio la traduzione letteraria, è pur vero che la mappa può essere adattata per approfondire lo studio del *respeaking* in tutti i suoi aspetti, secondo la riformulazione che si propone di seguito:

⁵⁶ Cfr. Zagar Galvão 2007.

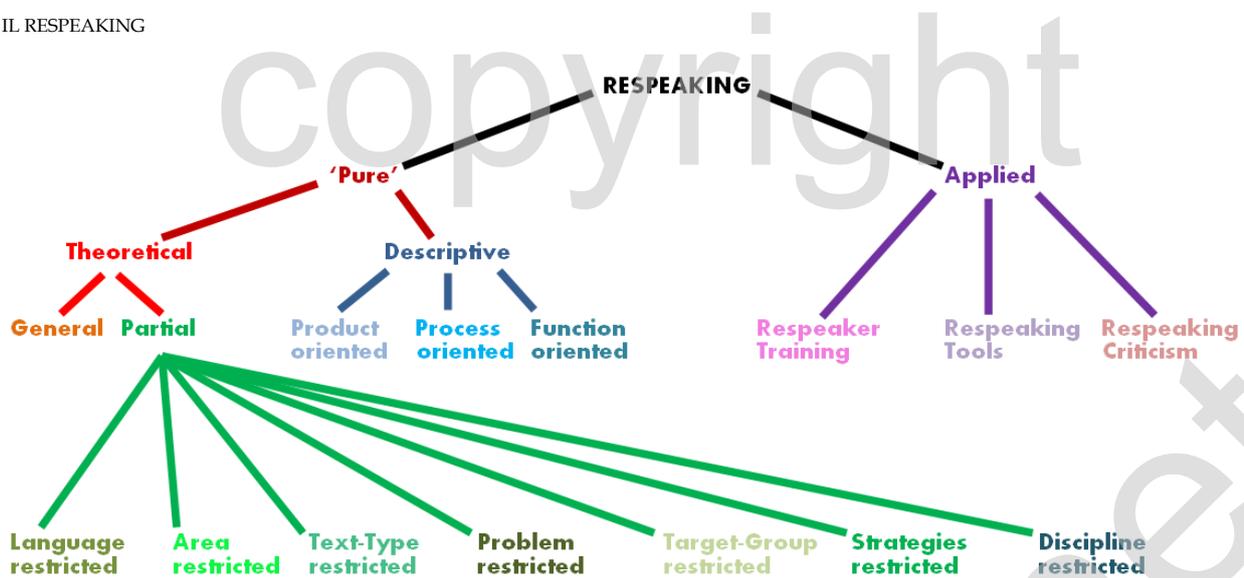


Figura 4 – Mappa per uno studio teorico e pratico del *respeaking*.

Secondo tale schema, vi sono due principali aree di ricerca, *pure* e *applied*:

- la ricerca pura ha due obiettivi legati l'uno all'altro:
 - I. descrittivo: descrivere il fenomeno così come si manifesta nell'esperienza pratica (*Descriptive Translation Studies* – DTS → ‘Studi Descrittivi sul *Respeaking*’ – paragrafo 2.2);
 - II. teorico: stabilire dei principi generali per mezzo dei quali tale fenomeno può essere spiegato e praticato (*Translation Theory* – TTh → ‘Teoria del *Respeaking*’ – capitolo II).

Questi due obiettivi sono a loro volta suddivisi in settori di studio specifici. I DTS in:

- *product oriented*: studio dei prodotti tramite analisi e confronto dei risultati ottenuti sulla base di uno o più parametri prestabiliti - stesso genere televisivo e/o stessa tipologia di *target* e/o stesse condizioni di lavoro e/o lingua di lavoro (studio comparato – sottoparagrafo 2.2.1);
- *process oriented*: studio di ciò che avviene, a livello pratico e psicologico, durante il processo del *respeaking* (psicologia e svolgimento del *respeaking* – sottoparagrafo 2.2.2);
- *function oriented*: descrizione della funzione del prodotto e della funzione del processo nel contesto socioculturale ricevente (sociologia del *respeaking* – sottoparagrafi 2.2.1 e 2.2.2).

La TTh in:

- teoria generale: mira a spiegare e considerare l'atto del *respeaking* nei suoi elementi basilari, rilevanti per il *respeaking* nel suo complesso (è un obiettivo ambizioso in quanto vuole creare delle basi valide per qualsiasi genere e prodotto televisivo – capitolo II);
- teorie parziali: vogliono spiegare e considerare l'atto del *respeaking* secondo parametri ben precisi: lingua (*respeaking* intralinguistico e interlinguistico – capitoli II e IV); area (in una data area culturale, sociale, linguistica – paragrafo 2.4); tipo di testo o meglio genere televisivo da rispeakerare (sport, informazione, intrattenimento, ecc. – paragrafo 3.2 e sottoparagrafo 3.5.2); problema (in base ai problemi specifici del *respeaking*, come l'interazione uomo-macchina, la correzione istantanea degli errori, ecc. – sottoparagrafo 2.3.3, capitolo IV); *target* (audiolesi e normoudenti – capitoli I e II); strategie (utilizzate per effettuare il rispeakeraggio, come omissione, espansione, ecc. – paragrafo 3.4, sottoparagrafo 3.5.4 e capitolo IV); ambito di applicazione (convegnistica, sottotitolazione TV, resocontazione, ecc. – paragrafo 2.2).
- la ricerca applicata si divide in tre ramificazioni a loro volta suddivise in settori specifici:
 - I. formazione del *respeaker*: preparazione ed esperienza (sottoparagrafo 2.3.2);
 - II. strumenti del *respeaking*: uso delle nuove tecnologie (*software* di riconoscimento del parlato, *software* di sottotitolazione, ecc. – sottoparagrafo 2.3.1);
 - III. critica del *respeaking*: valutazione dei risultati da parte di esperti e pubblico di destinazione (sottoparagrafo 2.3.3).

Come si può evincere dalla presente introduzione, teoria e pratica, in uno studio traduttologico di qualunque genere, non possono essere scisse perché inevitabilmente interconnesse. Come sostiene Newmark (2005:9), infatti:

Translation theory is pointless and sterile if it does not arise from the problems of translation practice, from the need to stand back and reflect, to consider all the factors, within the text and outside it, before coming to a decision.

Appare quindi ovvio che dalla sperimentazione e dalla ricerca vengono tratte delle linee teoriche che, a loro volta, possono dare spunto a nuove ricerche pratiche e così via, assicurando uno sviluppo sempre crescente alla disciplina. Si è deciso dunque di analizzare in questo capitolo il *respeaking* in base alle voci della mappa⁵⁷ e seguendo le linee teoriche

⁵⁷ L'ordine seguito per strutturare e spiegare la mappa non è vincolante ai fini dell'ordinamento interno del capitolo.

tracciate dall'esperienza di più studiosi della materia, per poi aprire la strada alla ricerca pratica e a nuovi contributi nei prossimi capitoli.

2.2 Il *respeaking* da un punto di vista descrittivo

Il *respeaking* è una tecnica che utilizza la tecnologia del riconoscimento del parlato per la trascrizione automatica e può effettuarsi o direttamente dall'audio originale o tramite un operatore (*respeaker*). Tramite un microfono collegato direttamente al computer, un *respeaker* detta al *software* ciò che vorrebbe mettere per iscritto nel TA, in contemporanea alla trasmissione e all'ascolto del TP da parte del *respeaker* stesso. Il *software* elabora l'*input* vocale, trasforma le sue componenti in codici binari e lo restituisce nella forma visiva più vicina a quella desiderata, nel nostro caso in sottotitoli. Questa tecnologia, ideata inizialmente per la resocontazione e la trascrizione in tutti quei settori in cui è necessario avere nell'immediato un testo scritto delle comunicazioni orali e non è possibile utilizzare le mani (medico, meccanico, ecc.) o disporre di una figura professionale (dattilografo, stenotipista, ecc.) che segua il ritmo dell'eloquio, viene oggi utilizzata anche in ambito televisivo per sottotitolare programmi in tempo reale⁵⁸.

Premesso ciò, si propone un'analisi approfondita di tale tecnica che, secondo Eugeni (2008c:17):

[...] est constitué de quatre composantes fondamentales:

- processus: texte source > respeaker > texte 'moyen' > machines > texte cible;
- fonction du processus: le texte 'moyen' doit être conforme aux contraintes techniques des logiciels employés [...];
- produit: les sous-titres en temps réel d'un programme en direct;
- fonction du produit: assurer la compréhension du texte source au public ciblé dans le respect des normes.

2.2.1 Il prodotto e la sua funzione

Poiché la presente tesi si limita ad analizzare il rispeaking in ambito televisivo, è evidente che il prodotto in questione è costituito dai sottotitoli in tempo reale di un prodotto audiovisivo in diretta, in semidiretta oppure preregistrato. Per quanto riguarda il programma preregistrato, questo può essere sottotitolato durante la sua messa in onda oppure prima. In quest'ultimo caso, il *respeaker* lavora in una situazione più rilassata. Questo perché, pur ascoltando il TP e producendo i sottotitoli allo stesso tempo, può apportare le necessarie

⁵⁸ Cfr. Eugeni 2006.

correzioni, aggiornare il *software* con termini sconosciuti, tornare indietro per capire meglio una frase e non si deve preoccupare dell'aumento del ritardo in rapporto all'enunciazione del TP in quanto i sottotitoli saranno sincronizzati successivamente.

Il prodotto è, secondo Eugeni, una traduzione intersemiotica perché dal canale orale si passa a quello scritto; supersemiotica perché aggiunge la componente scritto-verbale al TP senza sostituirla le altre componenti (nel caso di sordi oralisti e normoudenti), ma diasemiotica quando sostituisce completamente la componente audio (nel caso di Sordi segnanti, i quali non sentono nessun suono e non leggono le labbra in quanto la loro lingua materna è la LIS), per cui i canali di comunicazione sono sempre due (video-non verbale e scritto-verbale) ma diversi rispetto al TP; intralinguistica o interlinguistica; contenente alcuni elementi audio verbali e non verbali del TP; asincrona perché trasmessa sullo schermo pochi secondi dopo l'enunciazione; realizzata secondo alcune convenzioni⁵⁹. Come si evince dal primo capitolo, la realizzazione dei sottotitoli è già *per se* un compito non semplice dato che si devono tenere in considerazione varie restrizioni (tempo, spazio) e costrizioni (convenzioni imposte dal cliente o dall'azienda fornitrice del servizio di sottotitolazione), oltre a fattori quali la cultura e le esigenze del pubblico di destinazione⁶⁰. Restringendo il tutto al campo del *respeaking* intralinguistico e interlinguistico, considerati i tempi stretti in quanto il lavoro durante la messa in onda non permette al sottotitolatore di pensare comodamente alla migliore soluzione possibile, il *respeaker* e/o l'*editor* cura(no):

- la struttura e il contenuto dei sottotitoli: per facilitare la leggibilità, i sottotitoli dovrebbero seguire un senso logico ed essere lineari (soggetto-verbo-oggetto), evitando il più possibile la subordinazione; non dovrebbero contenere tratti tipici dell'oralità, come esitazioni, false partenze, ripetizioni, ecc., ed elementi ridondanti perché impliciti o deducibili dalle immagini; dovrebbero contenere il numero minore possibile di errori⁶¹;
- la punteggiatura: un testo scritto deve essere necessariamente regolato da segnali interpuntivi che veicolino in modo corretto e immediato il messaggio senza dare adito a incomprensioni e fraintendimenti. Inoltre, nel caso di sottotitolazione per S/sordi, la punteggiatura permette di comunicare informazioni paralinguistiche;
- i colori (soprattutto nel caso di sottotitolazione per S/sordi): utili per differenziare i diversi locutori. A tal fine, si può anche dettare o richiamare direttamente con un tasto il nome di chi parla (*tag* o etichetta). Tali accorgimenti possono rivelarsi utili anche nella sottotitolazione in tempo reale interlinguistica. Visto che il ritardo dovuto

⁵⁹ Cfr. Eugeni 2008c:14.

⁶⁰ Cfr. par. 1.3.

⁶¹ Cfr. capitoli III e IV.

all'elaborazione dei sottotitoli e al loro invio (da parte del rispeaker, della macchina e dell'emittente) sfasa la comparsa del testo scritto rispetto alle immagini corrispondenti, questo può creare confusione anche nello spettatore udente che, pur sentendo i diversi toni di voce e accorgendosi del cambio turno, non si rende conto chi dice cosa all'interno del testo scritto. Si pensi, per esempio, a un dibattito politico in cui i turni di parola sono brevi e/o possono sovrapporsi: le parole pronunciate da un interlocutore potrebbero comparire sotto l'immagine del suo 'rivale'. Il comando per l'alternanza dei colori e lo spostamento dei sottotitoli sullo schermo è disponibile anche manualmente grazie a un'apposita periferica;

- i segni grafici (soprattutto nel caso di sottotitolazione per S/sordi): i trattini per differenziare due parlanti all'interno di uno stesso sottotitolo, '#' per indicare testi cantati sulla scena, ecc.;
- i tempi di permanenza dei sottotitoli sullo schermo, nel caso di sottotitoli in blocco; il ritmo di proiezione del testo scritto sullo schermo parola per parola, nel caso di sottotitoli che scorrono.

A ciò si aggiungono aspetti tecnici prestabiliti indipendenti dalla volontà del rispeaker:

- la modalità di proiezione dei sottotitoli: *roll-up* o *pop-on*⁶²;
- dimensioni e tipo del carattere ed eventuale *background* (*box*, contorno o ombra), importanti per una buona leggibilità.

Quanto alla funzione del prodotto, questo deve permettere, a un pubblico di destinazione ben definito e tramite gli accorgimenti appena elencati, la comprensione del TP. Nel caso di rispeakeraggio interlinguistico, pressoché inesistente, i sottotitoli devono consentire la comprensione del TP prodotto in diretta televisiva a un pubblico di lingua e cultura diverse rispetto a quelle cui è originariamente destinato. Nel caso di rispeakeraggio intralinguistico, ormai sempre più praticato in vari Paesi, i sottotitoli devono rendere accessibile un TP prodotto in diretta a un pubblico di audiolesi. Ovviamente, diversi *target* richiedono diversi accorgimenti ed è perciò necessario partire sempre da un'analisi approfondita del bacino di utenza per poter poi riuscire nella realizzazione di un prodotto funzionale.

2.2.2 Il processo e la sua funzione

La tecnica del *respeaking* si distingue da tutte le altre tecniche di sottotitolazione perché prevede la formulazione orale, da parte del rispeaker, di un testo 'di mezzo' (TM) che sarà poi

⁶² Cfr. sottopar. 1.3.6.

trascritto dal *software* TAL (Trattamento Automatico del Linguaggio). Perciò, il rispeaker ascolta e guarda il TP, lo riformula, lo traduce o lo ripete oralmente, creando un TM1 per la macchina che, tramite appositi *software*, elabora e trasforma l'*input* vocale in sottotitoli elettronici (TM2) da verificare, eventualmente correggere e trasmettere direttamente sugli schermi televisivi (TA)⁶³. Ovviamente, tutto ciò ha un prezzo: la percezione, la ricezione e la comprensione del TP, l'elaborazione mentale da parte del rispeaker per il passaggio dal TP al TM1, l'interazione uomo-macchina, la trasmissione da un *software* all'altro, la verifica e, infine, l'eventuale correzione della trascrizione e l'invio dei sottotitoli richiedono del tempo che oscilla dai 2 (senza correzione) agli 8 secondi creando un ritardo tra la produzione del TP e la comparsa dei sottotitoli corrispondenti.

Da un punto di vista pratico, il rispeaking segue questa procedura:

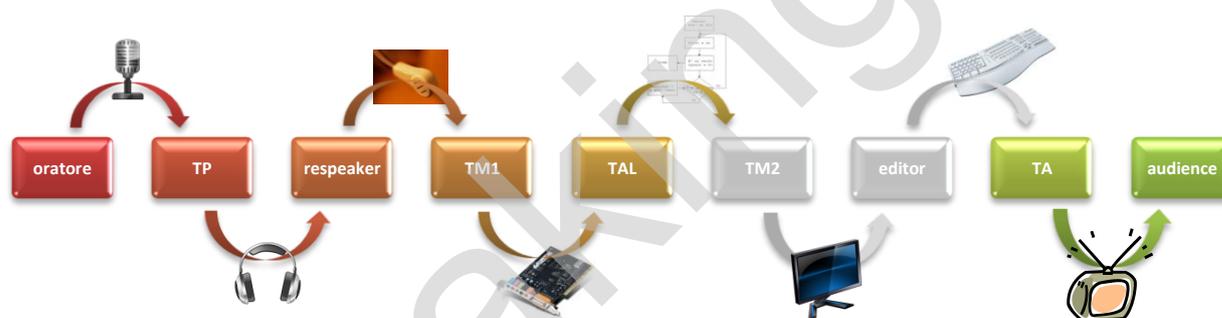


Figura 5 – Il processo del *respeaking*.⁶⁴

Il rispeaker lavora generalmente in cabine insonorizzate e deve essere pronto a sottotitolare ogni genere di programma a diverse velocità di eloquio e a diversi livelli di tecnicità. Se il programma da sottotitolare dura più di un turno di lavoro (dai 15 ai 40 minuti, dipende dall'esperienza del rispeaker e dal genere di programma) o è particolarmente complicato (dibattiti, discussioni parlamentari, ecc.), i rispeaker lavorano in coppia, alternandosi o 'spartendosi' i locutori. Grazie alle moderne tecnologie, il rispeaker ha, inoltre, la possibilità di lavorare in diverse sedi dell'emissione televisiva o addirittura da casa:

Una semplice rete interna permette di collegare tutte le postazioni operanti sullo stesso programma; il cambio di turno viene effettuato tramite un segnale standardizzato che

⁶³ Cfr. sottopar. 2.3.1.

⁶⁴ Il colore grigio delle caselle 'TM2' e 'editor' segnala il carattere opzionale delle rispettive fasi.

avverte chi deve subentrare o chi deve essere sostituito dell'imminente cambio o dell'avvenuto avvicendamento.⁶⁵

Da un punto di vista psico-cognitivo, il *respeaking* richiede una profonda concentrazione, una grande abilità a lavorare sotto condizioni di duro stress e la giusta determinazione per non arrendersi davanti a risultati inevitabilmente imperfetti. Infatti, il rispeaker deve contemporaneamente:

- monitorare continuamente il video per accertarsi che l'eventuale ritardo nella trasmissione del sottotitolo non comprometta la comunicatività del TA. In questo caso, deve rimediare tramite strategie particolari, come l'esplicitazione;
- ascoltare l'audio e capire un testo mai sentito (ci sono casi in cui il rispeaker può visionare il programma con un anticipo non sufficiente alla preparazione dei sottotitoli preregistrati) e non sempre prevedibile;
- dettare alla macchina in modo chiaro e riconoscibile un TM1, frutto della riformulazione, ripetizione o traduzione del TP, in vista di un TA pienamente accessibile al pubblico e possibilmente privo di errori;
- controllare ed eventualmente correggere il TM2 prima dell'invio (ciò non è necessario se il rispeaker è affiancato da un *editor* o se l'azienda di sottotitolazione prevede la messa in onda immediata del sottotitolo così come viene riconosciuto dalla macchina);
- trasmettere i sottotitoli (non necessario se è presente l'*editor* o se si seleziona l'opzione di invio automatico).

Più nello specifico, gli sforzi psico-cognitivi compiuti dal rispeaker comprendono:

- la conversione nell'immediato di un testo orale in uno che dovrà essere scritto: i limiti di tempo e le richieste delle associazioni di S/sordi spingono per la realizzazione di sottotitoli *verbatim* (per quanto riguarda la sottotitolazione intralinguistica). Tuttavia, il passaggio dall'orale allo scritto implica necessariamente un adattamento del testo, tenendo conto degli accorgimenti sopra elencati per la realizzazione di quello che si vorrebbe come prodotto finale. Il tutto continuando a seguire l'eloquio originale, monitorando i video e rispettando i limiti di elaborazione della macchina per non sovraccaricare il sistema⁶⁶;
- il buon utilizzo del *software* e delle periferiche (microfono, cuffie, tastiera in caso di correzione del testo prima della sua trasmissione, ecc.);

⁶⁵ Cfr. Eugeni 2006.

⁶⁶ “[...] the limit of 300 words per minute is not derived by a research among the target audience. It is simply ‘a human limit’ (Marsh, 2005) due to the fact that after the threshold of 300 wpm, the pronunciation is no longer ‘recognizable’ by the software.” (Eugeni 2008a:377)

- la regolazione del volume e del tono di voce, una buona pronuncia, un'attenta e chiara articolazione delle parole, la scansione dei confini tra i vari termini, l'utilizzo di brevi pause, ecc. per una buona interazione uomo-macchina;
- l'uso di termini facilmente riconoscibili e il *pre-editing* di errori che possono essere causati dai limiti del *software* (omonimi, omofoni, parole simili e/o sconosciute, ecc.);
- il controllo dello stress, della frustrazione per aver fatto un errore e/o aver trovato una soluzione scadente;
- la coordinazione con altre figure professionali, come per esempio l'*editor* o colleghi *rispeaker* con cui ci si alterna nel lavoro;
- nel *respeaking* interlinguistico, il passaggio da una lingua all'altra e da una cultura all'altra.

Se a queste attività 'di routine' si dovessero aggiungere il continuo controllo del testo trascritto, la correzione degli errori in *post-editing*, l'invio manuale dei sottotitoli e il controllo del prodotto finale, il carico psico-cognitivo sarebbe enorme e porterebbe a un potenziale aumento degli errori da parte del *respeaker*. Inoltre, più errori si fanno, più lo stress e la frustrazione per i risultati ottenuti aumentano facendo diminuire la concentrazione e generando ulteriori errori:

What is difficult here is that this is the only skill that cannot multitasked with the others. You can respeak as you listen, you can monitor your voice as you listen and respeak and you can even watch images and text as you listen, respeak and monitor your voice, but you cannot respeak as you correct.⁶⁷

Il processo, secondo Eugeni (2006), è perciò una traduzione interlinguistica o intralinguistica, simultanea, isosemiotica perché utilizza lo stesso numero di canali semiotici del TP per il TM1 e intrasemiotica perché utilizza lo stesso canale di produzione, cioè quello orale. Nel processo si seguono delle convenzioni e si verbalizzano alcune componenti para- ed extralinguistiche (prosodia, tono e timbro di voce, effetti speciali, cambio di oratore, ecc.) tramite l'uso della voce. Il TM1 contiene quindi elementi verbali e non verbali (punteggiatura, uso dei colori, didascalie esplicative, ecc.)⁶⁸.

Per quanto riguarda la funzione del processo, questa consiste nel rendere accessibile, a livello tecnico, un TM1 per il *software* di riconoscimento del parlato e per quello di proiezione dei sottotitoli. In questo modo si evitano errori di *layout* del sottotitolo (su tre righe anziché su due, mal frammentato, parole diverse rispetto al TM1, ecc.) e, di conseguenza, si

⁶⁷ Cfr. Romero-Fresco 2011:99.

⁶⁸ Cfr. Eugeni 2008c:15.

fornisce un buon servizio al pubblico di destinazione. Come afferma Eugeni, “[...] il rispeaker parla agli utenti finali attraverso il *software*.”⁶⁹

Altre finalità di secondo piano potrebbero essere quelle di rispettare le leggi a livello nazionale e internazionale sull’accessibilità ai *media* e sulle pari opportunità delle persone disabili e, infine, di risparmio in termini economici per la produzione di sottotitoli in tempo reale. Infatti, come è stato accennato nel primo capitolo, il *respeaking* non è l’unica tecnica disponibile, ma è la più recente ed economica.

2.3 La ricerca applicata

La ricerca applicata si concentra sugli aspetti più concreti e pratici del rispeakeraggio, quali gli strumenti utilizzati, la formazione professionale e le critiche mosse dalla reale applicazione di questa tecnica, entrando nel vivo dell’argomento e nella realtà quotidiana di questo particolare mondo.

2.3.1 Gli strumenti del *respeaking*

Strumenti imprescindibili per fare rispeakeraggio in una situazione ottimale sono: un paio di cuffie, un microfono, uno o più monitor, un computer, il pacchetto *software* di sottotitolazione contenente i programmi di riconoscimento del parlato, di *editing* e di sottotitolazione.

Innanzitutto, è necessario che il sistema sia collegato con il servizio *teletext*, specificando il tipo di protocollo di comunicazione (TCP/IP) e il computer remoto sul quale è in esecuzione il servizio. Si deve poi impostare la forma dei sottotitoli (*font*, tipo e dimensioni del carattere, ecc.) e decidere se inviarli in modo automatico o manuale al *teletext*⁷⁰. In più, è necessario calibrare il microfono e regolare le cuffie. Esistono diversi tipi di microfono più o meno adatti a questo tipo di lavoro: cuffia con microfono, microfono da tavolo, ‘gelato’, con collegamento USB o *Wireless* o *Bluetooth*, *array*⁷¹, *stenomask*⁷². La scelta del microfono da utilizzare non deve essere secondaria perché una buona qualità dell’*input* agevola il riconoscimento da parte del *software*.

Dopodiché, il sistema è così strutturato:

⁶⁹ Cfr. Eugeni 2008b:73.

⁷⁰ Cfr. Aliprandi & Verruso 2006.

⁷¹ Cfr. Romero-Fresco 2011:74.

⁷² Una maschera dotata di microfono che capta solo i suoni pronunciati dal locutore, isolandoli da quelli dell’ambiente circostante, compreso il respiro del rispeaker.

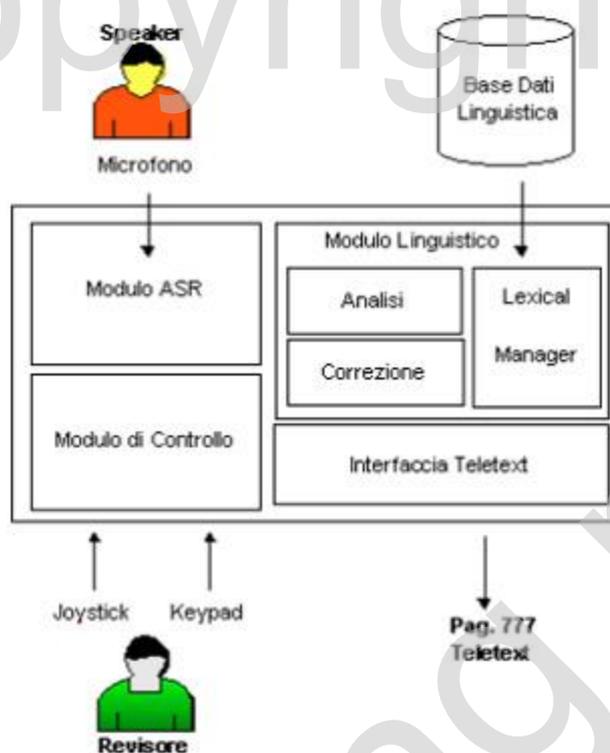


Figura 6 – Architettura del sistema⁷³.

Il rispeaker (denominato *speaker* nella figura) riceve l'audio tramite cuffie e il video tramite monitor. A seconda dei sistemi utilizzati, il rispeaker riceve la trasmissione direttamente dalla televisione oppure in anticipo rispetto alla sua effettiva visualizzazione da parte degli utenti finali. In quest'ultimo caso, il segnale viene catturato prima di arrivare al satellite che lo trasmette sugli schermi, perciò il rispeaker ha qualche secondo di tempo per creare i sottotitoli e inviarli affinché questi siano mandati in onda in sincronia col programma o con un ritardo minimo (2-4 secondi circa, contro i 4-6 di media nel primo caso). Simultaneamente alla continua ricezione di audio e video, l'operatore rispeaker il programma dettando il TM1 al *software* di riconoscimento del parlato con apposito microfono collegato al computer. Il Modulo *ASR* (*Automatic Speech Recognition*), grazie agli ausili fonetico-fonologici, morfologici, lessicali e sintattici (vocabolario di base e vocabolari specifici), riconosce il naturale eloquio del rispeaker e lo trascrive in formato testo. Tuttavia, il *software* di sottotitolazione converte in automatico il formato per generare sottotitoli. I sottotitoli passano a uno a uno al Modulo Linguistico, dove il sottomodulo *Analisi* fa l'analisi morfologica e sintattica, mentre il sottomodulo *Correzione*, grazie alle indicazioni ricevute dal

⁷³ Cfr. Aliprandi & Verruso 2006.

modulo gestore della Base Dati Linguistica (*Lexical Manager*), fa le correzioni automatiche. Dopodiché, il sottotitolo viene messo in coda nel *buffer* e finalmente visualizzato nel Modulo di Controllo. Se l'invio del sottotitolo avviene in automatico, allora questo viene direttamente trasferito, nel rispetto di un lasso di tempo predefinito (Ritardo di Invio⁷⁴), al servizio *teletext* che provvede alla sua messa in onda. Se, invece, l'invio è manuale, il rispeaker o l'*editor* verifica il sottotitolo, eventualmente lo corregge e solo dopo ne conferma l'invio. In una situazione di lavoro ottimale, il rispeaker, così come l'*editor*, ha a disposizione tre *input* audiovisivi: il video del TP così come viene trasmesso dall'emittente, l'area di dettatura, dove si visualizza il TM2 ed è possibile apportare modifiche manuali, e un'area *teletext*, dove viene visualizzato il prodotto finale. Solitamente, tutti i dati ottenuti dai vari passaggi per la creazione di sottotitoli vengono registrati⁷⁵.

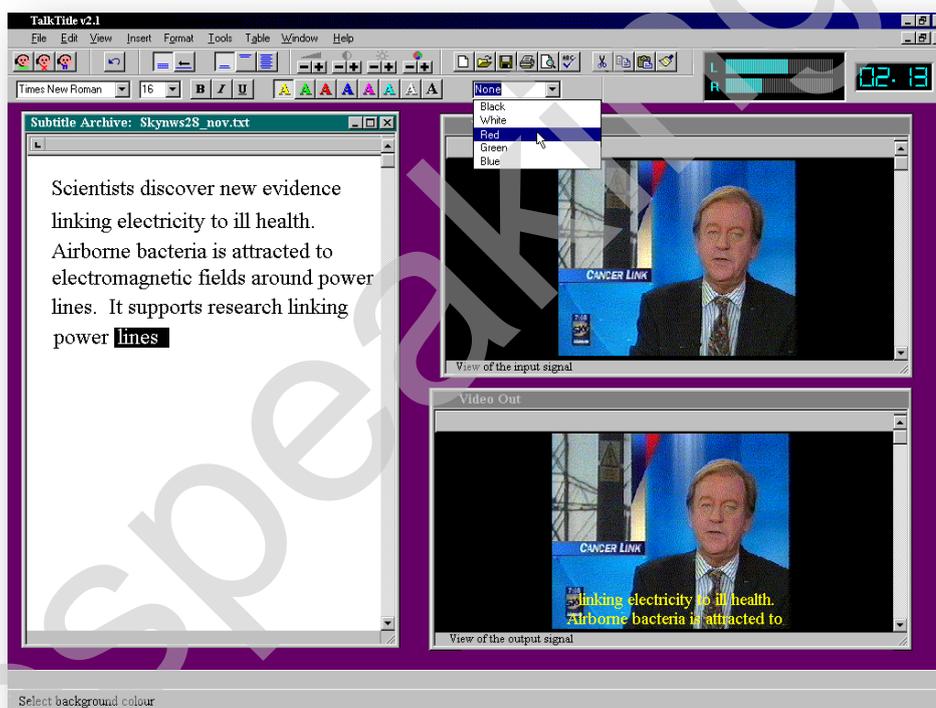


Figura 7 – Interfaccia per la visualizzazione dei sottotitoli in fase di produzione. © JRC VOICE project.⁷⁶

⁷⁴ Se i sottotitoli sono inviati in blocco, il *software* riconosce la pausa naturale del rispeaker come la fine del sottotitolo e lo invia in automatico, al posto del sottotitolo precedente. Nel caso di un sottotitolo breve, si rischia di far scomparire il sottotitolo precedente prima che lo spettatore riesca a leggerlo. Perciò, questa opzione si rivela molto importante perché permette di ritardare la trasmissione dei sottotitoli pronti per la messa in onda, rispettando i tempi di proiezione e di permanenza dei sottotitoli sullo schermo.

⁷⁵ Cfr. Aliprandi & Verruso 2006.

⁷⁶ Cfr. Eugeni 2008b:38.

La chiave di volta del rispeakeraggio è l'utilizzo del *software* di riconoscimento del parlato, il quale, più che uno strumento, è un vero e proprio collaboratore del rispeaker: per la prima volta, è la macchina e non l'uomo a trasformare il testo orale in testo scritto elettronico.

Il funzionamento dei diversi *software* è molto simile, le differenze riguardano perlopiù il livello di accuratezza e, di conseguenza, la velocità di elaborazione e trascrizione: più il sistema è accurato, più è lento. È importante, per un rispeaker, sapere come funziona il *software* e come questo processa i dati acustici, in modo da anticiparne i possibili errori ed evitarli tramite l'uso di strumenti messi appositamente a sua disposizione. Mentre l'uomo utilizza un *top-down approach* per riconoscere il parlato, cioè in base ai concetti e alle conoscenze possedute, i computer utilizzano un *bottom-up approach*, cioè analizzano le strutture del suono partendo dai singoli fonemi⁷⁷. Nello specifico, l'attività svolta dal *software* si divide in sette fasi:

- registrazione del suono;
- riconoscimento delle singole unità lessicali da elaborare grazie al modello acustico, alle grammatiche, ai vocabolari, al lessico, ai dizionari e al modello/i di linguaggio contenuti nel *software*. Il modello acustico contiene una raccolta di materiale audio cui corrispondono le rispettive trascrizioni contenute nei vocabolari sotto forma di liste e che vengono selezionate in base alle grammatiche e ai modelli di linguaggio. Il riconoscimento più utilizzato è quello basato sui *pattern*, per cui il parlato è confrontato con stringhe già note, determinando così delle corrispondenze. Perché questo sia possibile, il *software* deve riconoscere il punto di inizio e di fine di ogni singola unità. Mentre il primo è facilmente identificabile perché si confrontano i livelli audio dell'ambiente circostante con il campione appena registrato, il secondo è più difficile perché il rispeaker non fa pause tra una parola e l'altra o emette suoni non lessicali, come respiri, schiocchi con la lingua, ecc. Di conseguenza, il *software*, progettato per riconoscere il parlato naturale, deve capire dove finisce ogni parola a prescindere dalle pause. Se si sbaglia, riconoscerà male l'enunciato, come per esempio nel caso di: 'Giulia non va a casa' riconosciuto come *Giuliano va a casa*;
- pre-filtraggio (pre-amplificazione, normalizzazione, spostamento di banda, ecc.);
- suddivisione dei dati in un formato 'pulito', utilizzabile nella fase successiva di elaborazione dell'*input*;
- eventuale ulteriore filtraggio di ciascun dato durante il quale si effettuano gli ultimi aggiustamenti del campione;

⁷⁷ Cfr. Arumí Ribas & Romero-Fresco 2008.

- confronto con le possibili combinazioni tra fonemi e grafemi e *matching* dell'*input* con le parole contenute nei vocabolari.
- trascrizione.

Tutti questi passaggi avvengono in tempi molto brevi (meno di un secondo)⁷⁸.

Ogni programma ha un vocabolario di base creato a partire da *corpora* generalisti formati da notizie giornalistiche, documenti e testi di vario genere, a cui viene associato un modello del linguaggio addestrato sui *corpora* generalisti⁷⁹. Il *software* può essere dipendente o indipendente da chi parla (rispettivamente *speaker dependent* e *speaker independent*). Nel primo caso, il *software* è stato sviluppato per le esigenze di ogni singolo utente e ha prestazioni inferiori se viene usato da un altro utente senza che questo abbia realizzato un proprio profilo vocale. Nel secondo caso, qualunque utente può utilizzare il *software* senza creare un profilo vocale, ma l'accuratezza è inferiore rispetto a quella di un utente che usa il proprio profilo in un *software* dipendente dal parlante. La creazione di un profilo utente è necessaria affinché il computer comprenda:

[...] la manière dont le respeaker prononce chaque phonème et enchaîne ses mots. Pour ce faire, au fur et à mesure que le respeaker lira des textes plus ou moins longs, le logiciel enregistrera des informations physiques concernant la voix du respeaker (timbre, prosodie, volume, rythme, etc.), adaptant ainsi les algorithmes de reconnaissance vocale au parler du respeaker. Une fois le profil vocal créé, le logiciel nécessite un réglage du son ambiant, de manière à bien faire la distinction entre les bruits de fonds (réverbération, parasitage acoustique, caractéristiques vocales non-linguistiques du respeaker, etc.) et les phonèmes produits par le respeaker.⁸⁰

Se una parola è pronunciata male o non esiste nel vocabolario, il sistema ricerca il termine acusticamente più simile. In caso di omofonia, il sistema sceglie, tramite un'analisi semantica, il termine più coerente col testo o, tramite un'analisi statistica, quello più usato di frequente nei *corpora* con cui è stato addestrato (ogni voce del vocabolario ha un indice di frequenza che viene aggiornato man mano che il termine viene utilizzato). È per questi motivi che, prima del lavoro *online*, è importante addestrare il *software* con i termini tecnici e/o stranieri, i nomi propri e i neologismi che probabilmente saranno menzionati nel programma. Il sistema consente, infatti, di ampliare il vocabolario tramite l'inserimento di nuove parole con rispettive pronunce e trascrizioni fonetiche. Nonostante le entrate del vocabolario siano potenzialmente illimitate, più parole vi sono presenti, più tempo il *software* impiegherà per effettuare il *matching*. È perciò possibile creare vocabolari specifici attivabili a comando per non sovraccaricare il vocabolario di base con termini normalmente poco utilizzati. Per

⁷⁸ Cfr. Eugeni 2008b:16-17.

⁷⁹ Cfr. Aliprandi & Verruso 2006.

⁸⁰ Cfr. Eugeni 2008c:18.

esempio, per la sottotitolazione di una partita di calcio Milan-Inter, si attivano i vocabolari tecnici ‘Milan’ e ‘Inter’, contenenti presumibilmente i nomi dei rispettivi giocatori. Questa opzione non solo consente l’inserimento di nuove parole, tecnicismi e nomi propri altrimenti irriconoscibili, ma, grazie al modello del linguaggio, permette anche di risolvere eventuali ambiguità distinguendo sempre la forma corretta in base ai vocabolari in uso. Il modello di linguaggio, infatti, è un meccanismo che calcola la possibilità di occorrenza di una stringa di parole, cioè analizza la parola riconosciuta dal modello acustico e calcola le probabilità per cui questa può seguire o precedere una o più parole in una frase⁸¹. Infine, esistono le grammatiche contestuali, le quali decidono, in base al contesto, che cosa è più probabile che debba essere riconosciuto (‘il cane abbaia’ anziché ‘il cane abbaiano’). Una delle funzioni delle grammatiche contestuali consente al rispeaker di pronunciare poche parole o frasi sintetiche che il modulo *ASR* riconosce ed estende automaticamente in strutture prestabilite. Le grammatiche contengono parti fisse personalizzabili e parti variabili che permettono di velocizzare i tempi di produzione del sottotitolo e standardizzare il testo. Le parole pronunciate nelle parti variabili funzionano come dei ‘segnalibri’ e sono contenute all’interno di vocabolari specifici chiamati vocabolari *NLU* (*Natural Language Understanding*)⁸². Un’altra funzione delle grammatiche è quella delle ‘macro’, cioè un comando vocale che serve a distinguere un termine da omofoni (per distinguere *Rossi* da *rossi*, si pronuncerà la macro ‘macro-Rossi’ per ottenere il primo e semplicemente ‘rossi’ per ottenere il secondo) o per essere più sintetici (‘macro-Regina’ per ‘sua Maestà la Regina Elisabetta II’)⁸³.

Il *software* di riconoscimento del parlato mette a disposizione anche altri strumenti, come i comandi vocali, con i quali il rispeaker può dettare punteggiatura e forma dei sottotitoli (colore, stile, ecc.) senza dover ricorrere alla tastiera, e lo strumento di analisi di documenti al fine di rintracciare le parole sconosciute all’interno degli stessi e permetterne l’inserimento con relative pronunce. Infine, si ha la funzione *house-style* che permette di correggere e scrivere in maniera opportuna frequenti errori ortografici o di trascrizione. Si tratta di liste di parole che vorrebbero essere scritte in modo diverso da come le scrive solitamente il *software* (*Presidente* invece di *presidente*). Molto utilizzate per le sigle e i nomi propri, possono essere sia di base, sia specialistiche. La lista creata in precedenza, così come i vocabolari specialistici, deve essere attivata prima della sottotitolazione del programma. Al suo interno è necessario inserire solo le parole che non hanno omofoni, altrimenti il *software* riconoscerà e utilizzerà in automatico la parola presente nella lista selezionata. In caso di

⁸¹ Cfr. Romero-Fresco 2011:58.

⁸² Cfr. Aliprandi & Verruso 2006.

⁸³ Cfr. Arumí Ribas & Romero-Fresco 2008.

errore di dettatura e/o trascrizione, il rispeaker può correggere subito il testo o in seguito. In ogni caso, il *software* impara i nuovi termini e li inserisce nel sistema⁸⁴.

2.3.2 La formazione e le competenze del rispeaker

Come si può ben capire dai paragrafi precedenti, il lavoro del rispeaker non è semplice e richiede impegno, dedizione e costanza. Tuttavia, la strada da percorrere per diventare un rispeaker professionista non è particolarmente difficile, soprattutto se si posseggono già delle competenze che possiamo definire ‘di *background*’. Il rispeaker è un operatore che si avvicina un po’ alla figura dell’interprete di simultanea e un po’ a quella del sottotitolatore⁸⁵, con l’aggiunta di ulteriori abilità che non appartengono né all’una né all’altra categoria e che sono necessarie per l’interazione uomo-macchina.

La formazione professionale di un rispeaker si basa su diverse tecniche ed esercizi pratici che permettono, prima di tutto, l’apprendimento della tecnica, poi, un aumento graduale della velocità e dell’accuratezza nella produzione dei sottotitoli. Prendendo l’esempio di un discente privo di competenze da simultaneista, Jie-Chen propone una formazione iniziale simile a quella per l’interpretazione simultanea⁸⁶. Innanzitutto, propone lo *shadowing*, cioè il discente ascolta un TP e ripete parola per parola, nella stessa lingua, quello che viene detto. Questo esercizio permette di abituarsi ad ascoltare e parlare allo stesso tempo senza perdere la concentrazione e il filo del discorso. Norman⁸⁷ fa una distinzione tra *phonemic shadowing*, cioè ripetere ogni parola senza il bisogno di capirne il significato, e *phrase shadowing*, cioè ripetere il testo aspettando prima il completamento di una *idea unit* (unità di senso in cui si segmenta il testo). Nicholson⁸⁸ aggiunge una terza distinzione, ‘*adjusted lag shadowing*, cioè ripetere il testo con uno scarto di parole prestabilito, fino a un massimo di dieci. Se lo scarto diventa maggiore e supera una proposizione, si crea troppa confusione e il discente potrebbe non distinguere più una proposizione dall’altra. L’esercizio è utile per sviluppare la memoria a breve termine, per imparare a seguire il discorso anche quando si ha un ritardo nel ripetere il testo e per cercare di recuperare. Per Jie-Chen, il passo successivo è il *clozing*, cioè il discente ascolta un discorso a un livello di competenza al di sopra delle proprie capacità e inizia a fare lo *shadowing*, cercando di completare le frasi senza

⁸⁴ Cfr. Eugeni 2008b:21, 183-184.

⁸⁵ Il sottotitolatore di prodotti preregistrati e il rispeaker si trovano spesso ad affrontare gli stessi problemi per quanto riguarda la sovrapposizione dei turni di parola, la velocità nel cambio turno e l’uso di *realia* (nomi propri, riferimenti geografici, ecc.). Cfr. Romero-Fresco 2011:48.

⁸⁶ Cfr. Jie-Chen 2006.

⁸⁷ Norman 1976, *op. cit.* in Eugeni 2008a: 359-360.

⁸⁸ Nicholson 1990:34, *op. cit.* in Eugeni 2008a: 359-360.

fermarsi. Tale esercizio permette di sviluppare l'intuito e di trovare sempre una soluzione anche quando non si capiscono o non si sentono delle parole. Così, nel caso in cui, per un problema di audio o di concentrazione, una o più parole non siano percepite chiaramente, il rispeaker non si ferma, ma continua nel suo lavoro trovando delle soluzioni accettabili. Gli esercizi per la ricerca di sinonimi hanno pressappoco lo stesso fine, ma sono più utili per il rispeaking interlinguistico in quanto il rispeaker può avere difficoltà nel trovare immediatamente la traduzione esatta di un termine pur capendo all'istante di cosa si sta parlando. L'esercizio consiste nel mostrare alcune parole chiave di cui il discente deve trovare il maggior numero possibile di sinonimi o parole correlate. Il passo successivo è l'*abstracting* che consiste nell'ascoltare un messaggio molto veloce (in cui possono essere state aggiunte parti ridondanti, false partenze, esitazioni, ecc.) e sintetizzarlo in poche frasi oppure rispondere ad alcune domande mirate sul testo. Questo esercizio è utile per analizzare il TP e carpirne le unità di senso più importanti da restituire in sintesi. Si sviluppano così le capacità di sintesi che si rivelano necessarie soprattutto quando il locutore parla molto velocemente. Il *paraphrasing* consiste nell'ascoltare un messaggio e ripeterlo in modo logico e professionale utilizzando le proprie parole. L'esercizio è utile per acquisire abilità di *editing* e riformulazione del messaggio, serve a riorganizzare un discorso di partenza per poterlo rendere in senso logico quando il TP non è totalmente comprensibile, il locutore usa un lessico povero, il locutore dice cose insensate oppure il discorso non è comprensibile se tradotto e/o ripetuto parola per parola. Un esercizio utile allo stesso fine è proposto da Gillies⁸⁹: ascoltare un testo e ripeterlo cambiando la struttura sintattica, invertendo le frasi o rendendolo ironico senza cambiarne il significato. Tutti questi esercizi possono essere fatti prima nella lingua materna del rispeaker, poi, solo dopo aver acquisito le varie competenze, dalla lingua straniera alla lingua materna e, se necessario, dalla lingua materna alla lingua straniera.

Esercizi simili sono proposti da Arumí Ribas e Romero-Fresco (2008):

- ascoltare un testo e identificare le idee principali, fare una lista di parole chiave e creare una mappa concettuale delle idee (gerarchia e relazioni interne);
- ascoltare un testo con parti mancanti e ripeterlo cercando di completare il discorso in modo da sviluppare abilità di anticipazione;
- ascoltare una narrazione mentre si conta o si recita una poesia, per sviluppare l'abilità di fare due cose contemporaneamente (*multitasking capacity*).

Altri esercizi per la formazione di un rispeaker riguardano ovviamente l'interazione uomo-macchina e la realizzazione dei sottotitoli. Innanzitutto, sono necessari esercizi per

⁸⁹ Gillies 2001, *op. cit.* in Arumí Ribas & Romero-Fresco 2008.

acquisire competenze fonetiche: riscaldamento, respirazione, modulazione della voce e articolazione. Per capire se si sono acquisite tali competenze, si possono proporre degli esercizi di lettura, prima senza, poi con il computer. Questi servono per imparare a modulare il tono e il volume della voce, capire il giusto ritmo da tenere, abituarsi alla distinta articolazione delle parole e, in seguito, familiarizzare con la macchina, imparare a parlarci evitando inutili produzioni verbali (riformulazioni, autocritiche, ecc.) e non verbali (colpi di tosse, schiocchi con la lingua, ecc.), imparando a conoscere il suo funzionamento e i suoi limiti. Degli esercizi di dizione, inoltre, aiuterebbero a migliorare la qualità dell'*input* per il riconoscimento del parlato e quindi a ridurre il numero di errori⁹⁰. Contemporaneamente, gli esercizi di lettura servono per l'aggiornamento del proprio profilo vocale, l'eventuale aggiunta di macro e l'addestramento del *software*, che è la parte più importante nel lavoro di un rispeaker. A tal fine, si può anche chiedere ai discenti di preparare una lista di parole chiave da dettare al *software* facendo particolare attenzione alla pronuncia di termini sconosciuti perché il *software*, successivamente, riconoscerà sempre quella stessa pronuncia. In secondo luogo, sono necessari esercizi di segmentazione per dividere il TP in unità concettuali da trasporre poi, più facilmente, nei sottotitoli. Il discente può guardare un programma, preparare la segmentazione e riscrivere il testo su un foglio o sul PC per poi rileggerlo con l'aggiunta della punteggiatura. Un'alternativa è quella di preparare i sottotitoli dalla segmentazione e inserirli successivamente nel TP.

Il gradino successivo è quello di vedere più volte un programma per poi cercare di farne il rispeakeraggio. In un secondo momento, il discente potrà iniziare a rispeakere brevi passaggi di programmi in diretta, allungando sempre di più il turno di lavoro man mano che fa pratica. Il formatore può chiedere di evitare l'uso di certe parole o inserire nel testo termini non presenti nel vocabolario per fare in modo che i discenti trovino delle soluzioni alternative. Per evitare di fare sottotitoli troppo lunghi e per abituare i discenti a guardare l'*output* si può, per esempio, imporre loro di fare sottotitoli di una sola linea⁹¹.

Vale la pena spendere due parole sul materiale da usare durante gli esercizi⁹². Sarebbe meglio utilizzare materiale che si avvicina ai programmi televisivi che il futuro rispeaker si ritroverà a dover sottotitolare. Perciò si può cominciare con i programmi sportivi, dove le immagini sono più importanti del testo, passando poi ai discorsi pubblici in forma monologica (dove non c'è bisogno di cambiare forma o posizione al sottotitolo perché il locutore è sempre lo stesso) fino ad arrivare al telegiornale, dove il testo è più importante delle immagini, ai

⁹⁰ Cfr. Eugeni 2008c:140-144.

⁹¹ Cfr. Arumí Ribas & Romero-Fresco 2008.

⁹² Cfr. Arumí Ribas & Romero-Fresco 2008 e Remael 2006.

dibattiti e ai *talk show*, dove i diversi turni di parola spesso e volentieri si sovrappongono ed è impossibile sottotitolare tutto.

La prima università al mondo a fornire una formazione accademica *online* per il rispeakeraggio è stata l'*Universitat Autònoma de Barcelona (UAB)*. L'università offre un master in traduzione audiovisiva e un master europeo in traduzione audiovisiva. Entrambi comprendono un modulo di sottotitolazione per non udenti. L'insegnamento comprende una sessione introduttiva in cui i discenti devono imparare a creare il proprio profilo usando il programma *Dragon NaturallySpeaking* e come usarlo. Durante la formazione si studia anche la teoria e la sottotitolazione in preregistrato, dopodiché i discenti apprendono le abilità pratiche del *respeaking*. La *Roehampton University*, invece, è l'unica università al mondo a offrire una formazione sul rispeakeraggio in diverse lingue, cioè inglese, spagnolo, francese, italiano e tedesco. Si tratta del modulo compreso nel master multilingue in traduzione audiovisiva. I programmi utilizzati sono *Dragon NaturallySpeaking* e *ViaVoice* nelle loro versioni multilingue. I discenti imparano a creare il proprio profilo vocale e ad addestrare il *software*, apprendono le capacità pratiche del rispeaker per poter lavorare in un contesto professionale e si esercitano con diversi generi audiovisivi ed eventi pubblici. La formazione prevede anche delle esercitazioni sul *respeaking* interlinguistico. L'*Artesis University College* offre, all'interno del proprio master in interpretazione, un corso di *respeaking* in olandese e utilizza *Dragon NaturallySpeaking*⁹³. Infine, la SSLMIT di Forlì è stata la prima facoltà al mondo a proporre l'insegnamento accademico del *respeaking* in un corso di laurea magistrale e l'Università di Winterthur prepara i rispeaker in italiano, francese e tedesco per la TV svizzera.

Per quanto riguarda la formazione in azienda, basandoci sull'esempio di *Red Bee Media (RBM)*⁹⁴, mentre in passato i rispeaker erano semplici sottotitolatori o interpreti di simultanea minimamente addestrati, con l'evolversi della professione si sono creati corsi *ad hoc*. Oggi il rispeaker per lavorare deve seguire una formazione ben organizzata e strutturata, alla fine della quale si devono superare dei test di prova. La formazione fornita dall'azienda dura solitamente tre mesi. Ogni rispeaker è costantemente monitorato per testare, in ogni momento, il livello di preparazione tenendo conto che la qualità da raggiungere è come minimo quella del 97%, ossia 3 parole su 100 possono essere sbagliate. *Red Bee Media* usa i rispeaker sia per la sottotitolazione in tempo reale, sia per il cosiddetto '*scripting*', cioè la realizzazione degli *script* di programmi preregistrati che saranno poi utilizzati per creare i sottotitoli, alternandoli

⁹³ Cfr. Romero-Fresco 2011:40-42.

⁹⁴ *Red Bee Media* ha una sede centrale a Londra e uffici in Europa, Asia e Australia. Offre servizi di sottotitolazione, di audio-descrizione, di interpretazione in lingua dei segni e di trascrizione. Tra i suoi clienti, annovera la *BBC* e *Canal+*. Produce circa 20.000-25.000 ore di sottotitolazione in tempo reale all'anno, due terzi dei quali sono realizzate da rispeaker. Cfr. Marsh 2006 e Romero-Fresco 2011:23-27.

nelle due pratiche per evitare il sovraccarico di stress dovuto alla diretta. Oggi *Red Bee Media* conta più di 100 rispeaker, molti dei quali lavorano da casa.

2.3.3 *Le critiche al respeaking intralinguistico*

Benché siano stati registrati, specialmente negli ultimi anni, notevoli sviluppi sia in termini di evoluzione tecnologica, sia in termini di formazione professionale dei rispeaker, ci sono ancora molte critiche nei confronti del *respeaking* intralinguistico. In particolare, le associazioni di non udenti⁹⁵ e i critici si lamentano della scarsa qualità dei sottotitoli che contengono, a loro avviso, un alto numero di errori e non riproducono il testo parola per parola.

Effettivamente, nonostante tutti gli strumenti tecnologici a disposizione del rispeaker, la sottotitolazione tramite *respeaking* raramente raggiunge la perfezione del 100% e la presenza di alcuni errori è ancora inevitabile. Gli errori possono essere imputati al rispeaker (cattiva pronuncia, fraintendimenti, ecc.⁹⁶) o al computer. In quest'ultimo caso, gli errori più comuni sono la non elaborazione di monosillabi e/o bisillabi (e, con, non, dove), parole simili al posto di altre (*porta diversa ai* invece di 'Porta di Versaille'), parole in più (*ha fatto* invece di 'affatto'), errori di concordanza grammaticale (*delle perturbazione*), ecc. Molti errori, con il giusto addestramento, una buona formazione professionale e soprattutto l'esperienza, si possono prevedere ed evitare grazie al *pre-editing*. In altre parole, il rispeaker cerca di limitare i possibili errori imputabili alla macchina non usando omofoni e frasi complesse, articolando bene le parole più brevi, rallentando l'eloquio, ecc. Ai fini di migliorare il *pre-editing*, sono sicuramente utili gli studi comparati di cui si è parlato nell'introduzione alla tesi. Infatti, se si conoscono gli errori più comuni, si possono trovare delle soluzioni per evitarli. Tuttavia, mantenere un controllo impeccabile anche su questo fronte è difficile per il rispeaker, al quale può sfuggire sempre qualcosa. Allora, una buona soluzione sarebbe quella di affiancargli un *editor*. L'*editing* è praticato in Italia e in Francia, ma non, per esempio, in Gran Bretagna, dove i *software* sono talmente avanzati e precisi che questa pratica non è ritenuta necessaria, malgrado la presenza di errori. In effetti, aggiungere un *editor* significa aumentare sia i costi, sia i tempi di trasmissione dei sottotitoli. Il sottotitolo realizzato in tempo reale per compiere il suo iter, dalla percezione del TP all'invio del sottotitolo, impiega almeno 2 secondi ritardando la sincronizzazione con il TP. Se a questi aggiungiamo il tempo di revisione, il ritardo rispetto alla trasmissione del programma aumenta. Il problema è sempre

⁹⁵ Cfr. Di Silvio 2006 e Mereghetti 2006.

⁹⁶ Cfr. capitoli III e IV.

lo stesso: più precisione richiede più tempo. Tuttavia, se il segnale viene captato prima del suo invio al satellite e agli schermi, si recupera il tempo necessario per fornire sottotitoli più precisi con un minimo di ritardo nella trasmissione o addirittura in sincronia col TP. L'*editor* deve:

- ascoltare il TP;
- verificare il TM2, identificare gli eventuali errori nei sottotitoli e correggerli il più velocemente possibile;
- cestinare i sottotitoli o confermarne l'invio al *teletext*.

Nel momento in cui l'*editor* individua un errore e lo corregge, inevitabilmente la sua attenzione sul TP e sul TM2 cala. Il rispeaker deve quindi applicare una strategia di compensazione, stando più attento alla formulazione del TM1 in modo da non generare nuovi errori che non possono essere immediatamente notati dall'*editor* e cercando di diminuire il divario che si crea tra TP e sottotitoli. Dato che l'*editor* è inevitabilmente più lento rispetto al rispeaker, i sottotitoli, una volta pronti, si accumulano nel *buffer* e appaiono uno dopo l'altro nel riquadro di revisione del *software* di *editing*. La presenza di un *editor* 'allevia le sofferenze' del rispeaker, che può concentrarsi esclusivamente sul TM1 senza preoccuparsi di eventuali errori. Pare ovvio che una soluzione di questo tipo, benché più costosa e lenta, riduca drasticamente gli errori e il carico di stress: il rispeaker fa meno errori perché si focalizza solo sul TM1 e gli errori in uscita vengono corretti, nel limite del possibile, dall'*editor*.

Per quanto riguarda i sottotitoli 'incompleti', si pone il problema dell'accessibilità. Accessibilità significa leggere tutti i termini che un locutore dice oppure ricevere tutte le sue idee? I punti da prendere in considerazione sono i seguenti: seguire di pari passo un eloquio rapido, come quello di un locutore che legge un testo preparato o che parla molto velocemente, richiederebbe il rallentamento del TP; il rispeaker, oltre a ripetere il discorso parola per parola, dovrebbe anche aggiungere la punteggiatura, le maiuscole e cambiare il colore dei sottotitoli in base al locutore; la riformulazione del testo rende i sottotitoli più facilmente accessibili al pubblico rispetto al mero susseguirsi di parole il cui significato nello scritto può cambiare rispetto all'orale; sottotitoli più lunghi e più frequenti non lasciano il tempo di guardare il resto del video al pubblico. La risposta pare ovvia, tuttavia, si critica il fatto che una riformulazione implica sempre una minima modifica del messaggio, alterando il testo da trasmettere ed 'escludendo' parte delle informazioni⁹⁷.

⁹⁷ Gambier, *op. cit.* in Eugeni 2008d:46.

Quali sono i difetti che dopo vent'anni affliggono ancora le sottotitolazioni? Innanzitutto, non c'è fedeltà linguistica, i testi vengono eccessivamente ridotti e sintetizzati e così perdono il livello significativo ed emotivo che la lingua parlata esprime. I sottotitoli non specificano il tipo di lingua parlata, i modi di dire, pensano che i sordi non siano in grado di capire la lingua italiana. Sbagliano!⁹⁸

In definitiva, le aziende di sottotitolazione, le emittenti televisive e le associazioni di S/sordi spingono per una sottotitolazione *verbatim*, mentre accademici e ricercatori sono per la riformulazione⁹⁹. Le linee guida dell'*ITC* raccomandano l'introduzione nei sottotitoli di una 'ragionevole percentuale' di parole pronunciate e una 'buona percentuale' di idee concettuali contenute nel TP¹⁰⁰. Per quanto riguarda la pratica reale, a seconda della velocità di eloquio nel TP e della preponderanza o meno delle immagini rispetto alla componente verbale, variabili prevedibili in base al genere di appartenenza del programma e a un'analisi multimodale dello stesso, il rispeaker deciderà il ritmo del proprio eloquio e le strategie da utilizzare per la produzione dei sottotitoli (ripetizione, espansione, compressione, omissione, ecc.). TP lenti vengono generalmente trascritti *verbatim* e/o tradotti parola per parola per non rallentare eccessivamente il ritmo annoiando il telespettatore, mentre TP veloci richiedono inevitabilmente una riformulazione e una riduzione quantitativa e qualitativa del testo¹⁰¹.

Oltre alla presenza di errori, che comunque sono sempre meno, e ai sottotitoli 'infedeli', ci si lamenta della non simultaneità dei sottotitoli con il TP. Anche in questo caso, lo sviluppo di nuove tecnologie e l'esperienza dei rispeaker fanno in modo di accorciare sempre più i tempi di produzione, revisione e invio dei sottotitoli, riducendo in maniera impressionante il ritardo tra trasmissione del TP e messa in onda del sottotitolo. Queste critiche valgono anche per le altre tecniche di sottotitolazione in tempo reale e sono quindi infondate. Per di più, il *respeaking* ha permesso una velocizzazione del processo innanzitutto perché il rispeaker produce oralmente il testo e non lo deve scrivere, come fa invece un dattilografo. In secondo luogo perché, con l'arrivo del digitale terrestre, la messa in onda sugli schermi degli utenti è rallentata e quindi, se si riesce a captare il segnale prima che questo arrivi al satellite, i tempi di trasmissione dei sottotitoli si riducono ulteriormente¹⁰². Se consideriamo poi i costi e gli anni di formazione necessari per addestrare un bravo stenotipista, la soluzione del rispeakeraggio appare ora la più conveniente e ricercata dalle diverse emittenti televisive.

⁹⁸ Cfr. Mereghetti 2006.

⁹⁹ Cfr. Romero-Fresco 2011:16.

¹⁰⁰ *ITC (Independent Television Commission)* è un'organizzazione britannica di consulenza radiotelevisiva sostituita da *Ofcom* nel 2003. Nel 1999, pubblica l'*ITC Guidance on Standards for Subtitling*, un manuale che riporta delle linee guida per la sottotitolazione per non udenti di qualsiasi programma televisivo e che viene usato attualmente dalla *Ofcom* come strumento indispensabile all'espletamento del proprio lavoro di controllo della qualità della produzione televisiva britannica. Cfr. Eugeni 2008b:143-144 e ITC 1999.

¹⁰¹ Cfr. capitoli III e IV.

¹⁰² Cfr. sottopar. 2.3.1.

Criticare i risultati di un servizio come quello della sottotitolazione in tempo reale tramite *respeaking* è semplice se non si tiene in considerazione il lavoro che c'è dietro e la situazione a livello europeo e mondiale:

[...] la particolarità della sottotitolazione di una trasmissione televisiva in diretta, spesso per motivi organizzativi non preceduta da una scaletta o altro materiale utile alla preparazione, implica una enorme quantità di problematiche, come la presenza di numerosi vocaboli esteri, cognomi, neologismi, dialoghi sovrapposti o concitati, ridondanze, frasi lasciate in sospenso, termini gergali o in dialetto, citazioni, doppi sensi, ecc., che devono essere “risolte” e rese “comprensibili” in quell’istante, senza la minima possibilità di ritardo o correzione.¹⁰³

Ed è ancora più semplice criticare nel momento in cui non si ha idea di come vengono realizzati i sottotitoli in diretta. Da un sondaggio condotto da Romero-Fresco, infatti, risulta che molte persone ritengono che siano realizzati in maniera automatica senza intervento umano¹⁰⁴. Inoltre, gran parte dei problemi risiede nella difficoltà oggettiva di comprendere un ‘parlato spontaneo’ sotto forma di lingua scritta. Le trasmissioni in diretta, quali *talk show* e approfondimenti, non seguono un copione preparato precedentemente, ma lasciano libero spazio ai conduttori e agli ospiti. Un udente capisce il senso del discorso perché oltre a sentire, coglie tutte le sfumature para- ed extralinguistiche che rendono più facile la comprensione, mentre la sola lettura integrale del testo, il più delle volte, diventa di difficile comprensione. Infine, molto dipende anche dall’abitudine a leggere o no: chi legge ha molta più proprietà di linguaggio e un vocabolario più ricco rispetto a chi non legge per cui sarà facilitato nella comprensione e nella velocità di lettura.

Un’ultima critica, questa volta pienamente giustificata, è quella sulla quantità dei programmi rispeakerati e sulla poca attenzione che si dà al pubblico dei S/sordi e alle loro esigenze. Solo ultimamente le varie reti televisive si sono attrezzate e si stanno attrezzando per la sottotitolazione di programmi in diretta e/o in semidiretta, sotto la spinta di normative a livello nazionale (UK, Danimarca, Spagna, ma non in Italia) e internazionale.

L’unica soluzione a queste critiche può venire dalla ricerca, trovando gli errori più comuni e una soluzione per evitarli, migliorando le tecniche di comunicazione e trasmissione tra emittente televisiva, postazione del rispeaker e servizio di messa in onda dei sottotitoli, sviluppando sistemi di trascrizione del parlato sempre più precisi e, infine, analizzando a fondo il pubblico di destinazione in quanto:

¹⁰³ Cfr. Cafarella 2009.

¹⁰⁴ Cfr. Romero-Fresco 2011:173.

[...] speech recognition based subtitling is target-oriented. As a consequence, the subtitler has to decide not only which is the best way possible to satisfy the needs of the end users, but also to identify them.¹⁰⁵

In attesa di nuovi risvolti, si continuano a generare linee guida sulla base di quelle della britannica *Ofcom*, risalenti al 2003: sottotitolare il programma in diretta nel rispetto delle *idea units*, cioè le unità di senso in cui si divide il TP, senza esigere una trascrizione parola per parola. Inoltre: per i sottotitoli in tempo reale è consentito utilizzare lo *scrolling*; il ritardo di sincronizzazione non deve essere maggiore ai 3 secondi; due linee di sottotitolo sono preferibili a tre; i colori devono essere diversi per differenziare i locutori; lo sfondo deve permettere la lettura dei sottotitoli; suoni ed effetti speciali devono essere sottotitolati laddove necessario; le bocche e le scritte non devono essere coperte; i sottotitoli devono trovarsi in basso in una zona ‘libera’; i programmi sottotitolati devono essere menzionati nelle guide TV; tutti i programmi sottotitolati devono contenere un messaggio iniziale che avvisi della presenza del servizio; tutti gli avvisi di emergenza a livello nazionale devono essere sottotitolati; tutte le emittenti televisive devono monitorare la buona qualità dei sottotitoli¹⁰⁶.

Come spesso succede, prima ancora di perfezionare una tecnica, si pensa già all’evoluzione futura, cioè la generazione automatica di sottotitoli. Questa consisterebbe in una prima fase di riconoscimento e sintesi vocale automatica e in una seconda fase di sottotitolazione a partire dal testo generato. Si pensa poi alla traduzione automatica del testo in uno o più lingue (*Global Autonomous Language Exploitation* della *DARPA*¹⁰⁷) e alla relativa costituzione o integrazione del *corpus* di riferimento¹⁰⁸. Recenti sviluppi in ambito tecnologico hanno inoltre permesso di creare un *software* di riconoscimento automatico del parlato che trascrive direttamente il TP senza l’ausilio del rispeaker¹⁰⁹. Infine, ci si concentra anche sullo sviluppo dell’inserimento automatico della punteggiatura. Tuttavia, oggi sembra ancora impossibile pensare effettivamente a un *respeaking* senza operatore in quanto la trascrizione di dialoghi che si sovrappongono e l’uso corretto della punteggiatura costituiscono degli ostacoli concreti che la macchina, da sola, non può superare.

¹⁰⁵ Cfr. Eugeni 2008a:358.

¹⁰⁶ Cfr. Higgs 2006.

¹⁰⁷ Cfr. Romero-Fresco 2011:71.

¹⁰⁸ Cfr. Muzii 2006.

¹⁰⁹ Cfr. Tescari 2011.

2.4 Il *respeaking* in Europa

Il *respeaking* è stato introdotto per la prima volta in ambito televisivo nel 2001 in Gran Bretagna (*BBC*) e in Belgio (*VRT*). La *BBC* ha rispeakerato nell'aprile 2001 la *World Snooker Championship*, seguita da altri importanti eventi sportivi, sessioni parlamentari e, infine, telegiornali nazionali e regionali mandati in onda da *BBC News 24*. Oggi la *BBC*, tramite *RBM*, sottotitola il 100% dei suoi programmi, più di 50.000 ore all'anno, ma affianca ancora la stenotipia al *respeaking*. In Germania, il *respeaking* arriva nel 2006 grazie all'azienda *TITELBILD* che fornisce sottotitoli in tempo reale per l'*ARD*. Anche in Danimarca il *respeaking* è utilizzato dal 2006. In Spagna, è stata composta una squadra per il rispeakeraggio solo nel 2008 da *Mundovisión/RBM* e oggi un *team* di 21 rispeaker sottotitola in tempo reale 150 ore mensili di programmazione. Anche in Svizzera il *respeaking* è stato introdotto nel 2008 e oggi *SWISS TXT* fornisce 185 ore mensili di sottotitoli realizzati tramite rispeakeraggio¹¹⁰.

2.4.1 Il *respeaking* in Francia e la qualità della sottotitolazione

In Francia, i primi canali a utilizzare il *respeaking* sono stati *M6* e *TF1*, rispettivamente nel marzo e nell'aprile 2007. Ogni squadra è composta da tre persone: il rispeaker (*perroquet*) ascolta il TP e lo rispeaker; il suggeritore (*souffleur*) confronta il TM1 e il TM2 e se una frase ha bisogno di modifiche avvisa l'*editor*; l'*editor* (*correcteur*) verifica ed eventualmente correggere i sottotitoli prima della loro messa in onda. Un approccio del genere, utilizzato da *ST'501*¹¹¹, è ovviamente più costoso, ma anche più preciso (97% di accuratezza). Il solo problema è che il ritardo nella trasmissione dei sottotitoli aumenta anche fino a 16 secondi. *M6*, tramite *Red Bee Media France*, adotta il metodo inglese, con un ritardo di massimo 5 secondi, ma con molti errori¹¹². *RBM France* utilizza un *software dependent* e offre un periodo di formazione che va dai due ai tre mesi prima di mandare in onda i propri rispeaker. I rispeaker lavorano sulla sottotitolazione sia di programmi preregistrati, sia di programmi in diretta. Durante la diretta, i rispeaker si alternano ogni 15 minuti per non affaticarsi eccessivamente. Una volta terminati i sottotitoli in diretta, questi vengono corretti in modo tale da aggiornare il *software* di riconoscimento del parlato. Il rispeaker lavora da solo,

¹¹⁰ Cfr. Eugeni 2008b:46, Red Bee Media 2011 e Romero-Fresco 2011:6-9, 27-35.

¹¹¹ *ST'501* è un'azienda francese che fornisce servizi di sottotitolazione di tutti i tipi e che annovera tra i suoi clienti *TF1* e *France Télévisions*, *Arte France* e *TMC*. Cfr. <http://www.st501.fr/index.php/qui-sommes-nous/st-501> (ultimo accesso 13 ottobre 2011).

¹¹² Cfr. Romero-Fresco 2011:35-36.

provvedendo alla buona realizzazione dei sottotitoli, verificando ed eventualmente correggendo gli errori in tempo reale:

Pendant que j'effectue le sous-titrage en direct par reconnaissance vocale, je vérifie sur l'écran en face de moi que les sous-titres défilent normalement et qu'ils ne comportent pas de fautes. Nous regardons toujours l'écran lorsque nous sommes à l'antenne, ce qui nous place dans la même situation que le téléspectateur et nous effectuons les changements nécessaires. Par exemple, nous devons changer la couleur du sous-titre à chaque fois qu'un nouvel intervenant prend la parole, nous devons lever ou abaisser les sous-titres sur l'écran pour qu'ils ne cachent pas les bandeaux à l'écran et nous devons inclure des indications telles que « APPLAUDISSEMENTS » ou « MUSIQUE » pour retranscrire la bande-son du programme de la manière la plus complète qui soit.¹¹³

Se dal 2010 *France Télévisions*, *TF1*, *Canal+ Premium* e *M6* hanno raggiunto l'obiettivo di sottotitolare il 100% della loro programmazione, ciò significa che tali emittenti sottotitolano anche tutti i programmi in diretta. Per di più, il fatto che le aziende di sottotitolazione che offrono i loro servizi ai canali francesi (*RBM France* per *M6* e *Canal+*, *ST'501* per *TF1* e *France Télévisions*) cerchino, nei loro siti Internet¹¹⁴, solo *perroquets* e *correcteurs* per arricchire i loro *team* significa che il *respeaking* ha ormai preso piede in Francia ed è la tecnica più utilizzata nella sottotitolazione in tempo reale.

Per quanto riguarda le convenzioni per la sottotitolazione in tempo reale, in generale, si richiede una sottotitolazione fedele e accurata con introduzione della punteggiatura e dei colori dei locutori: bianco per i locutori sullo schermo, giallo per i locutori fuori campo. Per ogni cambio locutore deve essere inserito un trattino seguito da uno spazio. Inoltre:

Vous ne pourrez certainement pas tout répéter mot pour mot. Vous serez souvent obligé(e) de résumer ce que vous entendez afin de rester au même niveau que les présentateurs. Quand vous résumez, assurez-vous de bien retransmettre le sens des propos répétés. De plus, la capacité du logiciel à reconnaître ce que vous dites dépend de votre élocution. Vous devez bien articuler et parler de manière posée.¹¹⁵

Per quanto riguarda la verifica della qualità della sottotitolazione, il *CSA* utilizza i parametri della fluidità, della leggibilità (il sottotitolo deve essere scritto in rilievo e ombreggiato per migliorare il contrasto oppure deve essere su sfondo nero) e possibilmente della sincronizzazione dei sottotitoli con il video¹¹⁶. La sottotitolazione deve essere la più

¹¹³ Cfr. *RBM France* http://www.redbeemedia.fr/docs/journee_type.pdf (ultimo accesso 13 ottobre 2011).

¹¹⁴ Cfr. *RBM France* <http://www.redbeemedia.fr/paris-jobs.html> (ultimo accesso 13 ottobre 2011) e *ST'501* <http://www.st501.fr/index.php/espace-emploi> (ultimo accesso 13 ottobre 2011).

¹¹⁵ Cfr. *RBM France* http://www.redbeemedia.fr/docs/descriptif_du_poste_Paris.pdf (ultimo accesso 13 ottobre 2011).

¹¹⁶ Cfr. *CSA* 2010:12.

accurata possibile, ma questo porta, come già detto in precedenza, all'evidente e inevitabile ritardo nella trasmissione dei sottotitoli, senza per altro garantire l'impeccabilità degli stessi.

2.4.2 *Il respeaking in Italia e il modello RAI*

In Italia, l'introduzione ufficiale del *respeaking* risale al 2008. La Colby¹¹⁷ ha introdotto per prima il monitor *touch screen* per la produzione di sottotitoli in tempo reale tramite *respeaking*. In questo modo il rispeaker può inserire la punteggiatura e le parole chiave selezionandole direttamente dal monitor. Il sistema può essere personalizzato per ogni programma. Presso Colby, i rispeaker per programma sono minimo due e si alternano su indicazione di un coordinatore, mentre un tecnico controlla il corretto funzionamento del sistema. Il coordinatore gestisce, inoltre, l'inserimento dei pulsanti, selezionabili attraverso il *touch screen* e invia manualmente in onda i servizi o testi precedentemente rispeakerati o importati nel sistema¹¹⁸. Per il momento, la Colby sottotitola per RAI e Mediaset.

Basandosi sull'esempio della RAI, si illustra di seguito la metodologia di lavoro adottata per la sottotitolazione in diretta tramite *respeaking*. L'azienda responsabile per la sottotitolazione riceve settimanalmente, dalla Direzione Televideo, l'elenco delle trasmissioni televisive da sottotitolare nella settimana successiva con gli orari della loro messa in onda. Prima dell'inizio della trasmissione da sottotitolare, la Direzione Televideo provvede a inviare al coordinatore del servizio gli eventuali testi o contenuti A/V in formato elettronico in modo tale che i rispeaker (selezionati dall'azienda) possano analizzare la scaletta della trasmissione e gli ulteriori materiali forniti. In questo modo, si ricavano i termini specifici utilizzati durante la trasmissione, si aggiorna il *software* e la trascrizione in diretta risulta più semplice e veloce. Ovviamente i rispeaker effettuano anche autonomamente la ricerca terminologica e preparano preventivamente il materiale e gli elementi utili al fine di facilitare il lavoro in diretta. Dopo aver effettuato l'interconnessione al TCP/IP e le prove tecniche, realizzate prima dell'inizio di ogni trasmissione, i rispeaker acquisiscono il TP tramite televisioni sincronizzate sui canali RAI per mezzo del digitale terrestre e inviano i sottotitoli, già formattati secondo lo standard *teletext*, tramite il protocollo *STLive*¹¹⁹. Alla fine della sottotitolazione di una trasmissione, la sessione di sottotitolazione viene chiusa mantenendo solo la connessione TCP/IP. Infine, i

¹¹⁷ La Colby è un'azienda che fornisce sottotitoli preregistrati per non udenti e multilingue, servizi di doppiaggio e *voice-over* e sottotitoli in tempo reale. Fondata nel 1988, nasce come società specializzata nella produzione di apparati per la generazione *teletext*. Cfr. <http://www.colby.it/pagine.php?p=content&txt=company&lang=it> (ultimo accesso 13 ottobre 2011).

¹¹⁸ Cfr. Romero-Fresco 2011:36-37.

¹¹⁹ Si tratta di un sistema informatico che gestisce sottotitolazioni contemporanee su più canali, su più pagine e in multilingua. Cfr. RAI 2011d.

sottotitoli realizzati in diretta sono corretti, sincronizzati e inviati, su richiesta, in formato STL alla Direzione Rainet per la pubblicazione sul portale Web della RAI. Per le operazioni tecniche e le eventuali problematiche tecniche, operative ed editoriali un *Service Manager* è messo a disposizione dall'azienda. L'effettiva erogazione del servizio e la qualità dello stesso sono verificate quotidianamente dalla Direzione Televideo su un campione del tutto casuale dei *file* prodotti dall'azienda e/o su segmenti registrati di sottotitolazione in diretta¹²⁰. Oggi, la RAI sottotitola alcuni dei suoi programmi in diretta tramite *respeaking*, arrivando a circa 4.000 ore all'anno¹²¹.

Per quanto riguarda le convenzioni da rispettare durante la sottotitolazione, queste sono precisate dalla RAI. I sottotitoli devono essere giustificati a sinistra e posizionati in basso, salvo alcune eccezioni. Parentesi e virgolette vanno aperte e chiuse in ogni sottotitolo, anche se la frase continua nel successivo. I punti interrogativi ed esclamativi devono essere preceduti da uno spazio. Se due o più persone parlano contemporaneamente e il dialogo è riportato nello stesso sottotitolo, i due turni devono essere contraddistinti da trattini seguiti da uno spazio. Se invece la frase è pronunciata da più persone contemporaneamente, deve essere indicato con l'etichetta '(insieme)'. Per quanto riguarda i colori, per il personaggio principale deve essere usato il bianco, per il secondo personaggio principale deve essere utilizzato il ciano, per gli altri personaggi principali devono essere utilizzati il verde e il magenta, per tutti gli altri personaggi il giallo. I rumori di fondo devono essere trascritti in sottotitoli centrati e di colore blu su fondo giallo. Se il rumore di fondo è accompagnato da un dialogo, anche il dialogo deve essere centrato. Le voci fuori campo devono essere precedute da un apposito simbolo che indica la presumibile provenienza della voce. Il pensiero deve essere scritto tra parentesi, mentre le frasi riportate o lette devono essere indicate tra virgolette. Le frasi cantate devono essere anticipate da un cancelletto (#) seguito da uno spazio che deve essere introdotto all'inizio di ogni sottotitolo. Il sottotitolo che conclude la canzone deve avere il cancelletto in testa e in coda¹²². Per ciò che concerne la qualità del servizio, si propone la seguente tabella riassuntiva i criteri di valutazione RAI:

¹²⁰ Cfr. RAI 2011a:15-18,28.

¹²¹ Cfr. Gruppo Rai 2011.

¹²² Cfr. RAI 2011c.

QUALITÀ DEL SERVIZIO DI SOTTOTITOLAZIONE		
Oggetto del monitoraggio	Indicatore	Valore target
Correttezza ortografica del sottotitolo	N° di parole con errori ortografici a programma	2 parole
Correttezza grammaticale e sintattica del sottotitolo	N° di frasi con errori di grammatica e/o sintassi a programma	1 frase
Adattamento del parlato in dialetto preceduto da indicazione tra parentesi – es. (in dialetto siciliano)...	N° di frasi in dialetto non trascritte a programma	1 frase
Aderenza del sottotitolo al significato del parlato	N° di frasi con significato differente dal parlato a programma	2 frasi
Correttezza dei colori convenzionali*	N° di frasi con colore convenzionale errato a programma	2 frasi
Correttezza della punteggiatura e della formattazione (es. punti, virgole, ecc.; maiuscole / minuscole, doppi spazi, lettere accentate, ecc.)	N° di frasi con errori di punteggiatura / formattazione a programma	1 frase
Indicazione tra parentesi di musica e rumori di sottofondo con valore diegetico (rilevante per la comprensione)	N° di rumori di fondo / musica rilevanti non indicati a programma	2 rumori di fondo / musica
Completezza dell'utilizzo della simbologia convenzionale (es. # per il cantato, parentesi per la voce-pensiero, ecc.)*	N° di simboli convenzionali non trascritti a programma	2 simboli convenzionali
Correttezza dell'utilizzo della simbologia convenzionale (es. # per il cantato, parentesi per la voce-pensiero, ecc.)*	N° di simboli convenzionali non corretti a programma	2 simboli convenzionali
Sincronia del sottotitolo con il parlato	Media dei secondi di latenza a programma	6 secondi
Adeguatezza della permanenza sullo schermo del sottotitolo	Media secondi di latenza	Tra 2 secondi e 6 secondi

Figura 8 – Tabella sulla qualità del servizio di sottotitolazione riguardo la produzione di sottotitoli tramite riconoscimento vocale. © RAI¹²³.

2.5 Il *respeaking* interlinguistico

Secondo la mappa proposta al paragrafo 2.1, il *respeaking*, a livello teorico, può essere indagato sotto diversi punti di vista, uno dei quali è quello della lingua. Il *respeaking*, infatti, può essere intralinguistico, cioè il *respeaker* produce i sottotitoli nella stessa lingua del prodotto televisivo in onda, oppure può essere interlinguistico, cioè il *respeaker* produce i sottotitoli in una lingua diversa rispetto a quella del prodotto televisivo in onda. Il primo fenomeno si rivolge in modo particolare agli audiolesi, il secondo è indirizzato perlopiù a un pubblico (di audiolesi e normoudenti) di lingua e cultura diverse da quelle dell'emittente.

2.5.1 Il *respeaking* interlinguistico in Europa

Tutto ciò che è stato detto finora vale sia per il *respeaking* intralinguistico, sia per quello interlinguistico. Tuttavia, è necessario fare una precisazione: mentre il *respeaking* intralinguistico sta pian piano prendendo piede in molti Paesi sostituendo o affiancando le

¹²³ Cfr. RAI 2011b:12-13.

altre tecniche per la sottotitolazione in tempo reale, il rispeakeraggio interlinguistico è quasi inesistente:

[...] despite the fact that respeaking can indeed entail translation, understood as an interlingual phenomenon, this only seems to occur in exceptional cases such as that of BBC Wales, where live programmes are respoken from Welsh into English. More often than not, though, respeaking is an intralingual phenomenon aiming at producing live subtitles for the Deaf and Hard-of-Hearing audience within the same country or language community.¹²⁴

La *BBC Welsh* offre quindi un servizio di sottotitolazione in tempo reale tramite rispeakeraggio, dal gallese all'inglese ma non viceversa, ed è l'unica emittente televisiva al mondo a utilizzare questa tecnica. In Danimarca, sono stati fatti alcuni esperimenti con i discorsi di Obama e altri eventi internazionali importanti:

On one occasion we have done interlingual respeak, namely the inauguration speech made by president Obama and when Obama spoke in Prague the same respeaker (who is a trained interpreter) did relay interpretation of the interpreter via respeak.¹²⁵

Anche nelle Fiandre, un primo test interlinguistico è stato effettuato con successo in occasione del discorso inaugurale del Presidente Obama. Questo esperimento è stato seguito dal rispeakeraggio di altri programmi, come *So You Think You Can Dance*, *The X Factor* (inglese-olandese) e *Henin* (francese-olandese)¹²⁶.

Negli Stati Uniti si utilizza un *software* di traduzione automatica dall'inglese ad altre lingue a partire dai sottotitoli intralinguistici prodotti¹²⁷, mentre Paesi come l'Italia preferiscono sottotitolare intralinguisticamente a partire dal TP commentato. Altri Paesi, come l'Olanda e il Belgio, praticano la sottotitolazione interlinguistica in tempo reale da parecchi anni, ma non tramite *respeaking*. Per esempio, la *NOB Cross Media Facilities*, un'affermata azienda di sottotitolazione olandese, offre un servizio di *live interlingual subtitling* fin dagli anni novanta. Tale servizio si è rivelato molto utile, registrando un incremento di *audience* grazie alla sottotitolazione di eventi internazionali importanti¹²⁸:

In the past four years, all kinds of events have been broadcast live, ranging from press conferences and speeches before and during the Iraq war, the American presidential election debates, the memorial service for the 2004 tsunami victims, the commemoration

¹²⁴ Cfr. Arumí Ribas & Romero-Fresco 2008.

¹²⁵ Cfr. Nothlev 2008.

¹²⁶ Cfr. Romero-Fresco 2011:30.

¹²⁷ Cfr. Díaz Cintas in Lavaur & Şerban 2008:36.

¹²⁸ Secondo de Korte (2006), l'emittente televisiva *SBS*, che negli scorsi anni ha trasmesso degli eventi speciali in diretta con sottotitoli interlinguistici, nonostante qualche problema tecnico, ha registrato un picco di ascolti superiore a quello delle reti che hanno trasmesso lo stesso evento senza sottotitoli.

of the liberation of Auschwitz (quite a challenge, as the languages used were English, Polish, and Hebrew), several royal weddings, to the 2006 Champions' League draw. There is a market for live subtitling, although it is not requested on a daily basis, but mainly for major international events, where people feel they want to be a witness to history.¹²⁹

La *NOB Cross Media Facilities* utilizza due sottotitolatori professionisti per *équipe*, uno come interprete, l'altro come dattilografo, che non trascrivono l'intero testo tradotto, ma lo sintetizzano quanto più possibile. Trattandosi di un lavoro molto stancante, due *équipe* si alternano in turni di 15-20 minuti per sottotitolare lo stesso programma¹³⁰. Un ritardo di 10 secondi nella trasmissione del segnale permette di sincronizzare i sottotitoli e correggere gli errori¹³¹. Il metodo adottato dalla *NOB* da una parte aggira l'ostacolo dell'interazione uomo-macchina, dall'altra è più lento e costoso rispetto al rispeakeraggio.

Un caso anomalo lo ritroviamo in Svizzera, dove i rispeaker si trovano spesso a dover tradurre il tedesco svizzero in tedesco standard, da dettare al *software* di riconoscimento del parlato. La Svizzera sarebbe un ottimo palcoscenico per la pratica del *respeaking* interlinguistico perché la tv pubblica svizzera è vista da francesi, italiani, tedeschi e romanci. La Svizzera grazie alla *SRG SSR*, cioè la Società Svizzera di Radiotelevisione, offre diversi canali televisivi in ognuna di queste lingue con sottotitolazione in tempo reale intralinguistica¹³².

Come si può notare, la sottotitolazione interlinguistica in tempo reale non è un'utopia: esistono già dei casi affermati, è utile per tenere il pubblico aggiornato sugli eventi internazionali in diretta e consente alle reti televisive di aumentare l'*audience*. Già da diversi anni, la TV satellitare dà la possibilità ai suoi clienti di accedere a dei canali stranieri senza muoversi da casa. Alcune aziende, come Sky, producono sottotitoli intralinguistici e interlinguistici di numerosi prodotti audiovisivi preregistrati, ma non di quelli in diretta e in semidiretta. Perché, allora, non si fornisce un servizio di sottotitolazione di tutti i programmi trasmessi via satellite, compresi quelli in tempo reale? E, nel caso si decida di farlo, perché non utilizzare il *respeaking*? Le risposte possono essere diverse. Innanzitutto, perché manca un interesse da parte dei *broadcaster* di investire in questo senso e la volontà di offrire un servizio in più al telespettatore. In secondo luogo, il rispeakeraggio intralinguistico ha impiegato diversi anni per farsi strada in ambito televisivo e tutt'oggi si trova a concorrere con altre tecniche di sottotitolazione in tempo reale. Convincere le reti televisive a utilizzare anche il rispeakeraggio interlinguistico sarà un'impresa ancora più ardua. Terzo, non tutti i

¹²⁹ Cfr. de Korte 2006.

¹³⁰ Cfr. de Korte 2006.

¹³¹ Cfr. Romero-Fresco 2011:13.

¹³² Cfr. Romero-Fresco 2011: 30-31.

Paesi hanno sviluppato dei *software* di riconoscimento del parlato per la propria lingua. Infine, il *respeaking* interlinguistico aggiunge un'ulteriore difficoltà a quelle elencate fino a ora: il *respeaker* deve tradurre in simultanea un testo audiovisivo. L'unica differenza che separa il *respeaking* interlinguistico da quello intralinguistico, cioè il passaggio da una lingua all'altra, avvicina definitivamente questa tecnica a quella dell'interpretazione simultanea.

2.5.2 *L'interpretazione simultanea e il respeaking interlinguistico: similitudini e differenze*

L'interpretazione simultanea è una tecnica che consiste nell'ascoltare un TP tramite cuffie collegate al canale di comunicazione del locutore e nel tradurlo contemporaneamente in un'altra lingua, parlando in un microfono, per le persone del pubblico che non capiscono la LP. Queste ultime ricevono il testo tradotto in simultanea tramite cuffie collegate al canale di comunicazione dell'interprete della lingua prescelta. Interpretazione simultanea e *respeaking* interlinguistico presentano molte caratteristiche comuni, ma anche molte differenze¹³³. Per quanto riguarda le similitudini, il simultaneista e il *respeaker* interlinguistico:

- devono conoscere bene entrambe le lingue di lavoro e normalmente traducono in passivo, cioè dalla lingua straniera (lingua B) alla lingua materna (lingua A);
- devono avere un minimo di conoscenze sull'argomento da trattare. Solitamente, sia l'interprete, sia il *respeaker* hanno conoscenze minori rispetto al pubblico per cui traducono riguardo all'argomento trattato. Perciò compiono entrambi uno sforzo psico-cognitivo in più. Gli argomenti di lavoro sono eterogenei, spaziano dal campo cinematografico, a quello medico, a quello politico, ecc.;
- si preparano in anticipo, a seconda dell'argomento, facendo ricerche terminologiche e redigendo glossari di possibili termini tecnici e nomi propri che verranno probabilmente utilizzati nel TP;
- hanno a che fare con TP aventi le stesse caratteristiche, cioè discorsi pronunciati da un unico oratore o dialoghi tra più persone. Nel primo caso si tratta spesso di un testo scritto in precedenza per essere letto, raramente è un discorso a braccio; nel secondo caso la comunicazione è imprevedibile, ma, nel *respeakeraggio*, può essere anche preparata in anticipo, per cui i personaggi sullo schermo seguono un copione;
- normalmente ascoltano solo una volta il TP e non hanno modo di chiedere chiarimenti a riguardo o di far ripetere una frase, non possono decidere i turni di parola e la

¹³³ Cfr. Arumí Ribas & Romero-Fresco 2008, Baaring 2006, Eugeni 2008a:362-380 e 2008c:33-72, Lambourne 2006, van der Veer 2006.

velocità di eloquio del locutore. In casi estremi, l'interprete può chiedere di rallentare o ripetere tramite appositi strumenti (tasto della *console*, comunicazione orale, ecc.);

- in alcuni casi, possono ricevere il testo in anticipo. Nello specifico, il rispeaker può ricevere in anticipo la trascrizione del TP e preparare i sottotitoli da sincronizzare in seguito oppure può semplicemente visionare in anteprima il programma e realizzare più facilmente i sottotitoli in tempo reale;
- sono i soli responsabili della produzione del TA (nel caso dell'interprete) e del TM1 (nel caso del rispeaker). Devono produrre un testo orale accettabile in termini di qualità, semanticamente e sintatticamente corretto, in tempi molto stretti, partendo subito dopo l'inizio dell'enunciazione del TP. Inoltre, dal momento che si rivolgono a un pubblico, seguono un 'codice etico' per la formulazione del testo. L'*output* è asincrono rispetto alla produzione del TP;
- ascoltano il TP, lo elaborano mentalmente e allo stesso tempo lo riproducono oralmente nella LA, con uno scarto minimo tra l'enunciazione e la riproduzione. Ciò significa che entrambi devono saper dividere la concentrazione tra tre diversi compiti (*Efforts*): prestare attenzione alla forma e al contenuto del TP (*Listening and Analysis Effort*), concentrarsi sulla forma e sul contenuto del messaggio da trasmettere in un'altra lingua (*Production Effort*), usare la memoria a breve termine per 'immagazzinare' le idee principali percepite e ancora da riformulare (*Short-term Memory Effort*). Nel caso in cui il simultaneista o il rispeaker dia troppo peso a uno o due degli sforzi psico-cognitivi appena menzionati, sovraccarica inevitabilmente la mente rischiando di commettere errori e/o ritardare la produzione del testo e/o perdere qualche passaggio del TP e/o dimenticare quello che deve dire¹³⁴. Interprete e rispeaker devono possedere quindi le stesse competenze psico-cognitive e attivano lo stesso meccanismo psicolinguistico per poter svolgere il processo. Nel caso del rispeaking, si aggiunge un ultimo compito: *Elocution Effort*¹³⁵, cioè lo sforzo di parlare in maniera chiara e inequivocabile per permettere alla macchina di elaborare meglio l'*input* ed evitare così eventuali errori di trascrizione;
- segmentano il TP in unità concettuali da rendere nella LA. La segmentazione è utile ai fini di equilibrare più facilmente i diversi sforzi psico-cognitivi. Il tempo di elaborazione del messaggio dipende anche dalla lunghezza delle unità concettuali, oltre che dal livello di conoscenza delle due lingue, dalla velocità e complessità dell'eloquio del locutore, ecc. Lederer (1981) afferma che, indipendentemente dalle

¹³⁴ Cfr. Gile 1995:159-190.

¹³⁵ Cfr. Eugeni 2008c:65.

lingue utilizzate, l'interprete (e in questo caso anche il rispeaker) può produrre facilmente il TA una volta recepite le unità di senso del discorso. Questo perché la traduzione diventa automatica nella mente dell'interprete ('Teoria del senso'¹³⁶), permettendo un minore sforzo nella produzione del messaggio. È da sottolineare, tuttavia, che l'interprete segmenta il testo per trovare un equivalente nella LA, il rispeaker lo segmenta per riformulare un testo nella LA da presentare sotto forma di sottotitoli;

- tutto ciò che non traducono può essere capito dal contesto comunicativo e dalla comunicazione 'fisica' del locutore (informazioni para- ed extralinguistiche). Inoltre, la compresenza del TP contribuisce a un senso di autenticità del discorso (questo punto non vale per il rispeakeraggio destinato agli audiolesi anche se la componente video contribuisce all'informatività del TP);
- modulano il tono e il volume della voce;
- monitorano continuamente l'*output* per correggere, se possibile, eventuali errori, ma anche per essere coerenti e lineari nella formulazione del proprio testo;
- lavorano costantemente sotto un forte carico di stress psicologico a causa del lavoro in tempo reale, del breve tempo a disposizione per la traduzione del messaggio e dell'impossibilità o quasi di correggere eventuali errori;
- se presenti fisicamente nell'ambiente lavorativo, lavorano all'interno di cabine insonorizzate con un paio di cuffie e un microfono, generalmente in coppia e per un tempo massimo di 30-40 minuti per turno; possono lavorare da casa o comunque a distanza. Nelle tele- e videoconferenze, il segnale impiega più tempo a essere trasmesso a causa dell'uso di computer e reti di comunicazione a distanza. Il ritardo nella trasmissione del messaggio sarà quindi maggiore.

Per quanto riguarda le differenze, il rispeaker interlinguistico:

- usa le liste redatte in anticipo per addestrare il *software* di riconoscimento del parlato e aggiorna costantemente il proprio profilo vocale in modalità *offline*;
- traduce un TP audiovisivo e, nel caso preso in considerazione dalla presente tesi, lavora in ambito televisivo trattando generi diversi, come notiziari, sport, sedute parlamentari, concerti, ecc. L'interprete, al contrario, lavora perlopiù in conferenze, *summit*, *meeting* e corti di tribunali. La situazione è sempre altamente convenzionale, di conseguenza, anche l'*audience* è diverso. Generalmente, l'interprete ha un pubblico

¹³⁶ Lederer, *op. cit.* in Eugeni 2008c:65-67.

omogeneo ed esperto in materia, il rispeaker ha un pubblico eterogeneo e imprevedibile;

- produce un testo orale ‘di mezzo’ il cui scopo è quello di essere elaborato e trascritto dalla macchina sotto forma di sottotitoli e quindi letto (non sentito) dal pubblico di destinazione;
- sintetizza ed edita il testo per presentarlo sotto forma di sottotitoli accessibili al pubblico nel rispetto del TP e dell’intenzione comunicativa. Per rendere i sottotitoli immediatamente comprensibili, il sottotitolatore tende a semplificare la forma preservando il contenuto. L’interprete segue nei limiti di tempo e di velocità di eloquio e in base alle peculiarità lessico-grammaticali delle due lingue coinvolte, la forma e il contenuto dettati dal locutore senza semplificare il discorso, cercando di mantenere lo stesso registro e le varietà linguistiche del locutore. La produzione di sottotitoli, inoltre, impone alcune costrizioni riguardo tempo e spazio di proiezione (l’interprete ha solo limiti di tempo);
- inserisce la punteggiatura (manualmente o oralmente) ed eventuali informazioni para- ed extralinguistiche (tramite diversi colori, *font*, punteggiatura, *tag*), elimina tratti tipici della lingua orale ed elementi ridondanti, rende il testo il più comprensibile, lineare e coerente possibile. L’impaginazione del sottotitolo è fondamentale per la leggibilità e l’accessibilità dello stesso, ma ‘ruba’ parte dell’attenzione che dovrebbe essere dedicata alle altre operazioni. È perciò importante che queste attività diventino automatiche o siano delegate a un *editor* in modo da non sovraccaricare lo sforzo cognitivo e la memoria a breve termine¹³⁷;
- non si rivolge direttamente a un essere umano, ma lavora ‘in collaborazione’ con un computer, per cui deve avere competenze fonetiche precise e rispettare dei limiti per farsi facilmente riconoscere dal *software*. L’interprete, invece, deve prestare attenzione al timbro di voce, alla dizione e all’articolazione per rendere il TA piacevole all’ascolto, ma la sostanza del TA da lui prodotto non è compromessa dalla scarsa articolazione quanto dalla presenza di un’inflessione regionale. Il contrario succede nel caso del rispeaker;
- deve anticipare gli eventuali errori della macchina, scegliere la terminologia e produrre il TM1 di conseguenza. Il TA non corrisponde sempre con esattezza al TM1. Se il rispeaker deve correggere un errore grave già proiettato sullo schermo, lo corregge e lo invia nuovamente corretto;

¹³⁷ Cfr. Eugeni 2008c:39.

- a seconda del *software* utilizzato, deve o meno regolare le pause per l'invio di sottotitoli in blocco perché il *software* riconosce la pausa naturale come la fine di un sottotitolo e lo invia in automatico se viene selezionata questa opzione (in questo caso, è necessaria la comprensione del TP da parte del rispeaker); deve controllare il ritmo di eloquio in caso di sottotitoli inviati parola per parola (in questo caso, il rispeaker non è obbligato a capire e riformulare il TP). L'interprete produce un TA continuo seguendo il ritmo di enunciazione del TP;
- non sempre vede il locutore sullo schermo e non ha alcun rapporto con esso (non può chiedere di rallentare, ripetere, ecc.);
- non ha alcun contatto con il pubblico di destinazione del TA. L'interprete si rivolge direttamente al proprio pubblico e, di solito, ha la possibilità di vederlo e ricevere eventuali *feedback* se c'è, per esempio, qualche problema tecnico;
- permette di superare barriere sensoriali e linguistiche. L'interprete si rivolge a un pubblico con limitata o assente accessibilità linguistica;
- mentre, solitamente, l'interpretazione avviene in un unico luogo in cui sono presenti locutore, interprete e pubblico (tranne nel caso di tele- e videoconferenze), il rispeakeraggio avviene al contempo in diversi luoghi: da dove provengono i dati video (campo di calcio, studio televisivo, ecc.); da dove provengono i dati audio (lo stesso luogo dei dati video o un altro, per esempio la cabina in cui si trova il telecronista); gli studi TV; gli uffici del *teletext* o la casa del rispeaker o qualsiasi altro luogo in cui questo e la sua cabina siano collocati; ogni schermo televisivo su cui il programma è trasmesso. L'interprete, se presente fisicamente, a volte non si sente a suo agio nel posto in cui lavora. Il rispeaker, al contrario, si adatta sempre molto facilmente all'ambiente che lo circonda e che conosce bene¹³⁸.

Eugeni propone un'analisi comparativa a partire dall'approccio socio-linguistico alla comunicazione elaborato da Dell Hymes¹³⁹, grazie al quale è possibile analizzare ogni tipo di discorso. Il modello di Hymes si compone di sedici aspetti raggruppati in otto categorie le cui iniziali generano il celebre acronimo *SPEAKING*: *Situation* (situazione), *Participants* (partecipanti), *Ends* (fini), *Act sequences* (forma e contenuto del messaggio), *Key* (il *set* di elementi che fissano “*le ton, l'esprit et la couleur de l'act de communication*”¹⁴⁰), *Instrumentalities* (strumenti), *Norms* (norme), *Genres* (generi). Gran parte delle categorie sono state trattate nelle pagine precedenti, ma ne rimane ancora una: i partecipanti.

¹³⁸ Cfr. Eugeni 2008a:364.

¹³⁹ Hymes 1974, *op. cit.* in Eugeni 2008a:362-380.

¹⁴⁰ Cfr. Eugeni 2008c:51.

Nell'interpretazione simultanea è difficile separare con precisione i ruoli (mittente, interprete, destinatario) in quanto il mittente può diventare destinatario del messaggio e viceversa, l'interprete è destinatario indiretto del messaggio e a sua volta mittente del messaggio tradotto, in più vi è una distinzione tra destinatari diretti del TP e destinatari diretti del TA. Allora, si può dire che il locutore è in quel preciso istante il mittente del messaggio (tranne se legge un testo scritto da altri, allora sarà un semplice intermediario), il quale si rivolge a dei destinatari diretti che capiscono la LP e all'interprete che è il destinatario indiretto. Questo ascolta il messaggio, lo rielabora nella LA e lo trasmette ai suoi destinatari diretti, i quali capiscono la LA e ricevono le informazioni para- ed extralinguistiche dal mittente 'reale' del messaggio. I ruoli possono scambiarsi (tranne quello dell'interprete) nel momento in cui non vi è più un monologo, ma è in corso un'interazione (per esempio la fase domande-risposte in un seminario). L'interprete, inoltre, non ha quasi mai le stesse competenze del pubblico di destinazione (sia della LP, sia della LA) per quanto riguarda l'argomento trattato. Un caso particolare è quello del *relais*: quando le lingue utilizzate sono più di due e un interprete non capisce la lingua parlata dal mittente, per poter fare il suo lavoro, l'interprete deve affidarsi a un suo collega (il *pivot*) che traduce dalla lingua del mittente a una lingua da lui/lei conosciuta. I partecipanti all'atto dell'interpretazione in questo caso sono: il mittente, il *pivot* (interpreta dalla lingua A alla lingua B), l'interprete dalla lingua B alla lingua C, i destinatari della lingua C. A questi possono eventualmente aggiungersi i destinatari della lingua A e quelli della lingua B.

Nel rispeakeraggio, i produttori di testo sono lo speaker e il rispeaker, i destinatari del messaggio sono il rispeaker (destinatario indiretto) e i telespettatori (destinatari diretti) che possono essere interessati o no al servizio di sottotitolazione. Lo speaker non è il solo mittente del messaggio. Gli autori, i tecnici del suono e del video e altre figure professionali che permettono il corretto svolgimento della trasmissione contribuiscono all'emissione del messaggio. Inoltre, un eventuale partecipante al rispeakeraggio è l'*editor*. Se si considera il rispeakeraggio come processo, il rispeaker si trova nella stessa situazione dell'interprete di simultanea. Ma se si considera il rispeakeraggio come prodotto, ci si accorge subito che vi sono delle differenze: il TP e il TA coesistono nello stesso prodotto audiovisivo e possono essere percepiti da chiunque sia capace di leggere e ascoltare. Nel caso dell'interpretazione simultanea, anche se un partecipante conosce sia la LP, sia la LA, questi riesce a seguire solo uno dei testi. È interessante notare che, mentre l'interprete mette in comunicazione diretta due parti, nel rispeakeraggio non vi è interazione tra mittente, rispeaker e *audience* da una parte e tra rispeaker e *audience* dall'altra. Il mittente (l'emittente televisiva in questo caso) fornisce un servizio e chi vuole ne può usufruire, ma non c'è un'interazione perché nessuna delle parti

si trova nello stesso luogo delle altre. Il rispeaker può ricevere dei *feedback* in seguito tramite e-mail o grazie ai dati sullo *share*, ma non ha alcun modo di entrare in contatto col pubblico durante il suo lavoro.

Il rispeakeraggio può essere utilizzato anche nelle conferenze, in seminari, ecc. In queste occasioni, rispeaker e interprete possono collaborare: l'interprete può approfittare dei sottotitoli intralinguistici prodotti dal rispeaker per capire meglio la LP, il rispeaker può approfittare della traduzione interlinguistica dell'interprete per generare i sottotitoli. È ovvio che, in quest'ultimo caso, il ritardo nella produzione dei sottotitoli sarà notevole perché si aspetta la traduzione dell'interprete e non si traduce direttamente dal discorso del locutore. L'utilizzo del *respeaking* interlinguistico ridurrebbe notevolmente i tempi di trasmissione dei sottotitoli.

2.6 Conclusioni

Nel presente capitolo si è voluta presentare la tecnica del *respeaking* partendo dalla mappa sui *Translation Studies* ideata da Holmes-Toury. Una volta spiegate le varie voci della mappa, si è passati all'analisi descrittiva vera e propria che ha interessato il prodotto, il processo e le loro funzioni, per poi arrivare alla ricerca applicata mediante la spiegazione del funzionamento degli strumenti, le proposte di formazione professionale e, infine, le critiche a questa tecnica. Descrivendo questi vari aspetti, il capitolo si pone come una breve teoria generale sul *respeaking*, utile per un approccio iniziale a questa tecnica. Per quanto riguarda le teorie parziali, che permettono al ricercatore di focalizzarsi su un aspetto specifico del *respeaking*, si è per ora affrontata la questione della lingua introducendo il rispeakeraggio interlinguistico, che è quello che più ci interessa e che ci permette di proseguire l'analisi e, soprattutto, la sperimentazione nei capitoli successivi. Sempre per quanto riguarda le teorie parziali, si è accennato anche all'uso e all'insegnamento del *respeaking* in varie aree geografiche con particolare attenzione alla situazione francese e italiana.

Ricapitolando, il fine ultimo del rispeakeraggio è quello di realizzare sottotitoli in tempo reale affinché il prodotto audiovisivo in onda sia facilmente accessibile al pubblico di destinazione. Perciò, tenere conto di una buona fruibilità del prodotto e del suo impatto nel contesto socioculturale ricevente è di fondamentale importanza: nel caso di sottotitolazione in tempo reale tramite rispeakeraggio intralinguistico, si migliora la vita di persone audiolese che hanno accesso a sempre più prodotti audiovisivi¹⁴¹; nel caso di sottotitolazione in tempo reale

¹⁴¹ “I sottotitoli sono uno dei mezzi di comunicazione più importanti che aiutano i sordi a sentire con gli occhi e che interrompono il silenzio che ci separa da loro. Lo sviluppo intellettuale e culturale dei sordi, come quello

tramite rispeaking interlinguistico, si fornisce un servizio in più ai telespettatori che, nell'era della TV digitale e dei canali satellitari, desiderano accedere anche alle reti televisive straniere. Per fare ciò, è necessario che il rispeaker disponga degli strumenti giusti e di una buona formazione professionale per acquisire, in particolar modo, le seguenti competenze:

- fonetiche, per una buona interazione uomo-macchina;
- metalinguistiche, per inframmezzare la produzione di un testo con la dettatura della punteggiatura;
- sintetiche, per la creazione di sottotitoli da un testo orale;
- psico-cognitive, per fare più cose allo stesso tempo;
- di genere, per evitare degli sforzi di memoria inutili ed essere più preciso nella resa dei termini tecnici¹⁴².

Infine, come afferma Mack (2006):

Colpisce in ogni caso la palese analogia con il compito del traduttore e interprete interlinguistico, che pure sa di non potere mai produrre una replica identica dell'originale ma, consapevole dell'inevitabile destino di dover interpretare soggettivamente quanto ascoltato prima di riformularlo in un'altra lingua, deve compiere continue scelte, in gran parte imposte da fattori esterni, potendo tutt'al più aspirare a 'dire quasi la stessa cosa'.

degli udenti, passa attraverso la comunicazione. I sottotitoli costituiscono quindi per i sordi un'ottima opportunità per stimolare il loro sviluppo culturale e lessicale e sono un sussidio all'apprendimento dell'italiano.” (Di Silvio 2006)

¹⁴² Cfr. Eugeni 2008b:44-45.