UNIVERSITÀ DEGLI STUDI DI NAPOLI *FEDERICO II*

DOTTORATO IN LINGUA INGLESE PER SCOPI SPECIALI

XX CICLO

TESI DI DOTTORATO

LA SOTTOTITOLAZIONE IN DIRETTA TV
ANALISI STRATEGICA DEL RISPEAKERAGGIO VERBATIM DI BBC NEWS

CANDIDATO

Carlo Eugeni

COORDINATRICE

Chiar.ma Prof.ssa Gabriella di Martino

RELATRICE

Chiar.ma Prof.ssa Rosa Maria Bollettieri

TUTOR

Prof. Christopher Rundle

Anno Accademico 2007-2008

"Thou shalt say all that I command thee
and Aaron thy brother shall speak"
(Exodus 7:2)

Introduzione

Il termine rispeakeraggio è stato proposto per la prima volta in Eugeni (2006a) come traduzione italiana del più fortunato lemma inglese *respeaking*, che letteralmente significa 'riparlare' e che indica proprio la tecnica che è oggetto di questa tesi di dottorato. Da allora, grazie al diffondersi di questa tecnica e al successo ottenuto dalla 'Prima giornata di studi internazionale sulla sottotitolazione in tempo reale'[1], tenutasi nel 2006 nella sede di Forlì dell'università degli studi di Bologna, l'uso di questo termine e della sua versione originale si è diffuso in maniera esponenziale[2]. Alla base della decisione di optare per un lemma italiano è stata la volontà di

> tentare di approdare a una soluzione che evitasse l'ennesimo prestito integrale adattandosi il più possibile alle regole morfo-sintattiche della nostra grammatica e cercando di non chiamare in causa lessemi ambigui già in uso per identificare attività affini o più generiche, come 'ripetizione' o 'riformulazione'. Ecco, quindi, che partendo dall'ormai parzialmente acclimatato *speaker*, che identifica una persona che parla in un contesto ben definito, declinato nella forma *speakeraggio*, già in uso, si giunge, tramite il suffisso *ri-*, alla forma proposta. (Eugeni 2006a)

Quanto a una prima definizione, Eugeni definisce il rispeakeraggio come

> una riformulazione, una traduzione o una trascrizione di un testo [...] prodotta dal rispeaker ed elaborata dal computer in contemporanea con la produzione del testo di partenza [...]. [Un] *software* di riconoscimento del parlato procede alla trasformazione dell'*input* orale in testo scritto. (*ibidem*)

Applicato alla produzione di sottotitoli televisivi per non-udenti, il rispeakeraggio diventa un utile strumento per la sottotitolazione in tempo reale, sia intra-linguistica[3] (*verbatim* e *non verbatim*), sia traducente da una lingua all'altra[4]. Emerge quindi immediatamente che si tratta di una 'tecnica' che rientra nel dominio della traduzione, più in particolare della traduzione audiovisiva. Descrittivamente, la traduzione, e con essa anche la traduzione audiovisiva, è una disciplina che si ripartisce essenzialmente in due branche: pura e applicata (cfr. Holmes 1987, Toury 1995, Gottlieb 2005). Nella traduzione pura rientrano

1 Cfr. www.intralinea.it
2 Il termine ha superato le 250 pagine sul motore di ricerca libera Google, sulle sole pagine in italiano. Precedentemente alla conferenza e alla pubblicazione dell'articolo, le pagine contenenti il solo lemma inglese erano 3.
3 Cfr. Eugeni e Mack 2006.
4 Cfr. de Korte 2006.

gli approcci descrittivi e teorici, mentre nell'applicata, rientrano vari approcci, tra cui l'applicazione in ambito professionale e, in ultima istanza, didattico.

Alla luce di queste prime generali indicazioni e vista la natura prettamente pionieristica che caratterizza i contributi in materia di rispeakeraggio[5], sembra naturale destinare lo scopo generale della presente tesi allo studio approfondito della tecnica in questione dal punto di vista descrittivo (nel senso letterario del termine), teorico, applicato (in ambito professionale e accademico) e strategico (che costitisce una sorta di ponte tra le ultime due aree). Per questo motivo, nella prima parte (capitoli 1, 2 e 3), si delineeranno le caratteristiche precipue del rispeakeraggio e, sulla base degli elementi emersi, si svilupperanno dei quadri e dei modelli teorici di riferimento (raggiungendo così i due maggiori obiettivi delle scienze empiriche[6]) per lo studio di alcuni aspetti del rispeakeraggio come processo e come prodotto; nella seconda parte (capitoli 4, 5 e 6), si applicheranno i risultati del terzo capitolo all'analisi delle macro-strategie traduttive adottate dai rispeaker della BBC (considerati all'avanguardia nel settore) nel sottotitolare *verbatim* e in tempo reale il programma britannico dinformazione *BBC News*. Da quest'analisi scaturiranno le migliori pratiche di rispeakeraggio *verbatim* in contesto anglofono, che saranno prima adattate e applicate al contesto professionale italofono e, infine, utilizzate per abbozzare una didattica del rispeakeraggio in ambito universitario.

Entrando maggiormente nel dettaglio, nel primo capitolo, si abbozzerà una panoramica del rispeakeraggio inteso sia come processo traduttivo, sia come risultato finale e saranno considerate le rispettive funzioni, del prodotto e del processo traduttivi. Da questa iniziale introduzione descrittiva emergeranno le caratteristiche fondanti del rispeakeraggio per non-udenti inteso sia come processo, sia come prodotto. Una prima ipotesi è che questi due aspetti caratterizzanti ogni forma di traduzione abbiano notevoli affinità con il processo dell'interpretazione simultanea da una parte e dall'altra con il prodotto della sottotitolazione per non-udenti (di cui il rispeakeraggio è peraltro uno strumento). Grazie a una duplice analisi contrastiva, sarà possibile verificare la veridicità di questa ipotesi e posizionare esattamente il rispeakeraggio all'interno del più vasto panorama degli studi sulla traduzione. Così facendo, si potrà approntare la creazione di un quadro teorico di riferimento per lo studio del rispeakeraggio come processo e come prodotto in chiave strategica. Per

5 Cfr. Orero 2006.
6 Cfr. Hempel 1952.

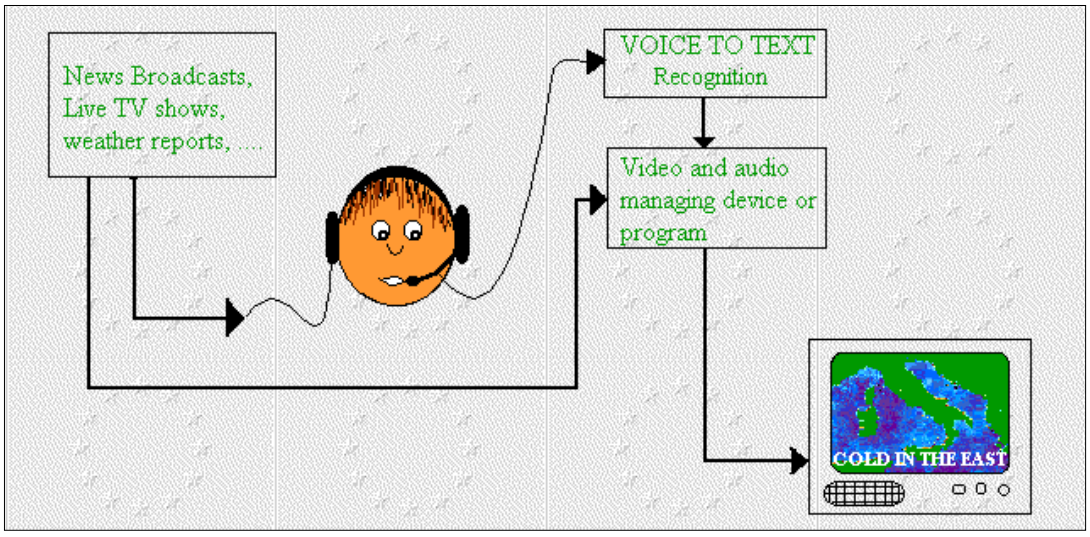
raggiungere questo obiettivo, si adatteranno gli insegnamenti degli studi sulle due discipline summenzionate alle caratteristiche del rispeakeraggio, tenendo in debita considerazione le somiglianze e le divergenze che saranno emerse dal secondo livello di analisi. Se si dovesse confermare la prima ipotesi, si farà ricorso ai contributi più prettamente psico-cognitivi e strategici dell'interpretazione simultanea (secondo capitolo) per delineare un ventaglio di possibili ragioni soggiacenti l'uso da parte di un rispeaker di una data operazione linguistica. Nel terzo capitolo, si cercherà invece di comprendere lo *skopos* del rispeaker e di definire una tassonomia di riferimento per l'individuazione e la catalogazione delle strategie traduttive emergenti dall'analisi comparativa tra un testo di partenza (TP) e il corrispondente testo di arrivo (TA).

Nel quarto capitolo, il modello strategico per lo studio di un testo rispeakerato (sviluppato nel capitolo 3) sarà adattato alle caratteristiche di genere e multimodali dei testi componenti il *corpus* in esame, composto di otto ore di *BBC News* sottotitolato intralinguisticamente e in tempo reale dai rispeaker dell'emittente britannica. Grazie a questa analisi, che costituirà la parte centrale della tesi, sarà possibile derivare le migliori prassi (tecnico-formali e linguistiche) in materia di rispeakeraggio intralinguistico *verbatim*. Oltre a comprendere meglio la natura del rispeakeraggio *verbatim* (maggiormente in uso rispetto al rispeakeraggio *non verbatim* per motivi che saranno analizzati successivamente), questi risultati permetteranno di redigere linee guida che saranno testate in un contesto italofono all'interno di un progetto internazionale volto all'accessibilità del TG in lingua italiana ai sordi segnanti (capitolo 5). L'ultimo capitolo sarà dedicato alla creazione di una didattica del rispeakeraggio come disciplina universitaria. Grazie alle prime esperienze in materia, alla conoscenza delle competenze necessarie al rispeakeraggio professionale descritte nei capitoli introduttivi e nel capitolo 5, alle linee guida derivate nel capitolo 4 e al contributo di Safar (1992 e 2006) in materia di insegnamento delle discipline audiovisive in contesto universitario, si proporrà un inquadramento didattico del rispeakeraggio in quanto professione da iniziare ad apprendere (in teoria e in pratica) in un corso universitario strutturato *ad hoc*.

Capitolo 1- Introduzione al rispeakeraggio televisivo

1.1 Introduzione

Nel rispeakeraggio televisivo, è possibile individuare immediatamente il prodotto (i sottotitoli di un programma in tempo reale) e la sua funzione (l'accessibilità dei programmi televisivi in tempo reale ad audiolesi, stranieri e a tutti coloro che si trovano nell'incapacità di poterne fruire agevolmente). Quanto al processo e alla sua funzione, però, l'analisi è resa particolarmente complicata da diversi elementi non presenti negli altri processi traduttivi affini. In particolare, come si è già brevemente visto, nel produrre i sottotitoli, il rispeaker detta il 'testo di mezzo' (TM)[7] a un computer che, grazie a un *software* di riconoscimento del parlato, trascrive l'*input* orale in testo scritto per poi proiettarlo in onda sotto forma di sottotitoli (figura 1). Nel passaggio dal testo di partenza (TP) al testo di arrivo (TA), al processo di elaborazione compiuto dal sottotitolatore si vanno ad aggiungere, quindi, l'interazione uomo-macchina tra il sottotitolatore e il *software* di riconoscimento del parlato e la trasmissione del TA dal *software* di riconoscimento a quello di sottotitolazione. Quanto alla funzione del rispeakeraggio come processo, il rispeaker tenderà alla produzione di un testo tecnicamente accettabile sia per il *software* di riconoscimento del parlato sia per quello di proiezione dei sottotitoli, onde evitare errori di *layout* del sottotitolo (rispettivamente, errori di trascrizione e mancato rispetto delle regole di messa in onda dei sottotitoli).



7 Il TM è il testo che il rispeaker detta al software di riconoscimento del parlato e che non coincide mai perfettamente con i sottotitoli che saranno poi letti dai telespettatori per cause imputabili al software o al rispeaker stesso. Si tratta quindi della versione ideale del TA. A parte i casi in cui si prevede la presenza di un professionista che corregge il TM prima della sua proiezione sullo schermo, i sottotitoli vanno in onda così come sono stati riconosciuti dal software, con tutti gli errori del caso. Ecco quindi che la distinzione tra TM e testo di arrivo ha una sua motivazione scientifica.

Figura 1: dopo aver ascoltato il TP, il rispeaker detta il TM al *software* di riconoscimento vocale che lo trascrive e lo trasmette al *software* di sottotitolazione, che lo mette in onda. © *JRC Voice project*.

Nel tentativo di approfondire ulteriormente la distinzione tra processo e prodotto e tra la funzione del processo e quella del prodotto, è forse utile dedicare due sottoparagrafi a parte a questi due binomi del rispeakeraggio.

1.1.1 Processo e prodotto

Gottlieb (2005) suddivide le diverse forme di traduzione facendo riferimento ai seguenti parametri:

- identità, o mancanza di identità, semiotica tra il TP e il TA (tipi di traduzione intrasemiotica e intersemiotica);
- eventuali cambiamenti nella composizione semiotica dell'atto traduttivo. Si avranno quindi tipi di traduzione isosemiotica (che utilizza gli stessi canali del TP), diasemiotica (che utilizza canali diversi), supersemiotica (che utilizza più canali), iposemiotica (che utilizza meno canali);
- grado di prescrizione delle norme (traduzione per convenzione vs. traduzione per ispirazione);
- grado di presenza di materiale verbale nel TA rispetto al TP. Si avranno quindi tipi di traduzione che restano verbali (traduzione di testi prettamente scritti), altri in cui si introducono elementi non-verbali (nella letteratura per l'infanzia si aggiungono illustrazioni a romanzi originariamente non intesi per l'infanzia), altri ancora in cui si introducono elementi verbali (è il caso della sottotitolazione per non-udenti) e infine traduzioni che restano non verbali (per un affresco ispirato a una statua raffigurante un episodio biblico).

Applicando questi criteri, risulterà che il rispeakeraggio, in quanto processo traduttivo, è una forma di traduzione intrasemiotica, per lo più intra-linguistica (anche se si conoscono esperienze di rispeakeraggio inter-linguistico, cfr. Marsh 2006 e de Korte 2006), isosemiotica, per convenzione e in cui vengono verbalizzati elementi para- ed extra-linguistici (come importanti rumori di sottofondo e l'intonazione di una data battuta). In

specifico, il processo del rispeakeraggio prevede un'identità semiotica tra il TP e il TM[8] (sono entrambi un'espressione linguistica); il TM viene prodotto all'interno dello stesso sistema linguistico; utilizzando gli stessi canali di produzione (la voce); nel rispetto di regole di produzione ben definite[9]; e con la verbalizzazione di elementi para- ed extra-linguistici tramite l'uso della voce. Un ultimo fondamentale aspetto, forse l'elemento più importante che lo contraddistingue dalla sottotitolazione per non-udenti in pre-registrato, riguarda la simultaneità del rispeakeraggio come processo traduttivo, visto che il TM è prodotto contemporaneamente al TP, senza essere sincronizzato. Differisce dalla sottotitolazione per sordi di programmi pre-registrati, in quanto quest'ultimo prevede una trasformazione *a posteriori* del TP.

Quanto al rispeakeraggio come prodotto finito, ossia il testo audiovisivo con i sottotitoli, può essere definito come traduzione intrasemiotica, intra-linguistica, supersemiotica, per convenzione e in cui si introducono elementi verbali e non-verbali. Se gli aspetti intrasemiotico, intra-linguistico e 'convenzionale' della traduzione rimangono immutati rispetto al rispeakeraggio inteso come processo, il TM è costituito da caratteristiche diverse rispetto al processo traduttivo dal quale deriva. I sottotitoli (testo (tra)scritto sovrapposto alle immagini) vanno infatti ad aggiungersi all'interazione delle componenti audio e video, verbali e non-verbali dell'originale. Inoltre, il TM presenta elementi verbali e non-verbali (punteggiatura, uso dei colori, didascalie esplicative, ecc.) necessari alla traduzione di tratti para- ed extra-linguistici (prosodia, tono e timbro di voce, effetti speciali, cambio di oratore, ecc.). Infine, la comparsa dei sottotitoli sullo schermo avviene in maniera non sincronica rispetto alla produzione del testo originale, ma con qualche secondo di ritardo. Come nel caso precedente, questo è un altro aspetto che diversifica il rispeakeraggio dalla sottotitolazione in pre-registrato.

1.1.2 Funzione del processo e funzione del prodotto

Come abbiamo già brevemente visto, la funzione del processo sta nell'interazione tra il rispeaker e la macchina e in particolare nel rispetto delle esigenze tecniche del *software* di riconoscimento del parlato e del *software* di sottotitolazione in uso. Un mancato rispetto di

8 Alla luce di quanto precedentemente affermato, in questa analisi del rispeakeraggio, quello che Gottlieb definisce il TA è da considerarsi rappresentato dal testo prodotto dal rispeaker, quindi il TM. Non sembra appropriato, in questa sede, parlare del TA, in quanto 'semplice' frutto della tecnologia.

9 In realtà esistono regole formali a cui aderire, ma non riguardanti il contenuto. Sarebbe quindi forse più corretto parlare di traduzione a metà tra la convenzione e l'ispirazione. Cfr. Eugeni 2006a e Eugeni 2007.

questi vincoli tecnici comporta una visualizzazione del TA diversa dagli intenti del sottotitolatore. In particolare, il sottotitolo potrebbe comparire su tre righe invece che su due, mal frammentato, con parole diverse rispetto a quelle del TM. Inoltre, la funzione del processo sarà anche quella di produrre sottotitoli che rispondano appieno alle esigenze e alle aspettative del pubblico a cui il TA è destinato. Per ragioni di ordine concettuale, però, questa funzione sarà considerata esclusivo appannaggio del prodotto.

Quanto alla funzione del prodotto, come si può leggere anche nella nota pubblicata nel sito della BBC[10] in materia di rispeakeraggio, l'obiettivo principale della produzione dei sottotitoli è l'accessibilità e l'inclusione di persone con problemi di udito: "BBC subtitles provide a transcript of the TV soundtrack, helping deaf and hard-of-hearing viewers to follow programmes".

Tuttavia, dal Libro Bianco del *Research and Development Department* della BBC, si evince che le funzioni del rispeakeraggio, sia come processo, sia come prodotto, vanno oltre quelle appena menzionate. In particolare, emerge che il ricorso al rispeakeraggio è stato effettuato per raggiungere l'obiettivo di sottotitolare la totalità dei programmi entro il 2008[11], ottemperando così alla legislazione in vigore nel paese, "whilst minimising the additional costs involved" (Marks 2003: 5). È infatti da sottolineare che la maggiore flessibilità del rispeakeraggio nei confronti dell'altro sistema utilizzato per produrre sottotitoli in diretta, la stenotipia, comporta anche un abbattimento dei costi in materia di reclutamento, formazione e remunerazione del personale. In sintesi, riprendendo le parole dell'attuale responsabile della formazione del *respeaking department* di RedBee Media (che produce sottotitoli per la BBC), le tre ragioni principali per cui la BBC ha iniziato a fare ricorso al rispeakeraggio non sono soltanto di natura sociale o tecnologica, ma anche e soprattutto di natura legislativa ed economica:

> Respeaking came into being for three main reasons. Firstly, there was a growing demand from deaf and hard of hearing audiences for a greater proportion of television broadcasts to be subtitled.
> Secondly, and perhaps consequently, the Broadcasting Act of 1990 stipulated that, from 1998, 50% of all television channels' output should be subtitled. That target rose to 90% by 2010, but the BBC's own target is to subtitle 100% of output by 2008.

10 Allo stato attuale della diffusione del rispeakeraggio nel mondo, la BBC è l'emittente televisiva che ne fa maggior uso. Cfr. Higgs 2006.

11 L'obiettivo è stato raggiunto nel mese di maggio 2008, cfr. http://www.bbc.co.uk/pressoffice/pressreleases/stories/2008/05_may/07/subtitling.shtml

> Thirdly, stenography is a highly specialised skill that takes years to master; therefore, stenographers are not only thin on the ground but also able to demand high salaries. To meet its subtitling targets, the BBC had to find an alternative method of subtitling live programmes that was both practical and cost-effective. (Marsh 2004: 22)

Oltre all'accessibilità, si aggiunge quindi come finalità del rispeakeraggio anche la volontà di rispettare le leggi nazionali e il risparmio in termini economici rispetto all'uso della stenotipia, un aspetto questo che, pur distinto, sembra essere in realtà una condizione indispensabile del concetto di accessibilità.

Riassumendo l'aspetto descrittivo del rispeakeraggio, esso consta di quattro aspetti fondamentali:

- *il processo*: TP → rispeaker → TM → macchina → macchina → TA;
- *funzione del processo*: il TM deve essere tecnicamente accettabile. In particolare, l'interazione uomo-macchina deve garantire una rapidità di produzione tale da permettere la sottotitolazione di programmi in tempo reale, una conseguente maggiore quantità di sottotitoli prodotti per l'emittente e una maggiore economicità rispetto alla stenotipia;
- *il prodotto*: i sottotitoli (spesso per non-udenti) così come compaiono sullo schermo;
- *funzione del prodotto*: garantire l'accessibilità di un programma in diretta al pubblico di destinazione.

Ora che si sono chiarite le diverse componenti del rispeakeraggio, saranno analizzati gli aspetti più prettamente tecnici del riconoscimento del parlato e conseguentemente del rispeakeraggio televisivo.

1.2 Il riconoscimento del parlato[12]

Come si è visto, i *software* di riconoscimento del parlato sono una componente essenziale del rispeakeraggio e ne influenzano in maniera sostanziale sia il processo, sia il prodotto, tanto da renderlo una forma di sottotitolazione differente da tutte le altre. Dopo aver presentato i *software* di riconoscimento del parlato dal punto di vista tecnologico e

12 Sebbene i termini riconoscimento del parlato e riconoscimento vocale siano spesso utilizzati come sinonimi, in realtà è bene scindere queste due modalità di trattamento del linguaggio umano. La prima tecnologia provvede al riconoscimento di un testo prodotto oralmente e a trattarne il contenuto a seconda dell'uso che se ne vuole fare: in questo caso a trasporre le parole enunciate da un oratore in testo scritto. La seconda tecnologia invece riconosce le caratteristiche fisiche di una voce, identificandone l'oratore.

tecnico, sarà confrontato il rispeakeraggio prima con le altre forme di produzione rapida di testo e quindi con le possibili applicazioni del riconoscimento del parlato.

1.2.1 Aspetti tecnologici

Il riconoscimento del parlato è il processo attraverso cui un computer ascolta un testo prodotto oralmente, lo riconosce e trasforma le sue varie componenti in codici binari. A seconda dell'uso che se ne intende fare, l'*input* può essere trasformato in immagini, operazioni o, nel caso qui discusso, in parole. Più in dettaglio, l'attività della maggior parte dei riconoscitori del parlato si può suddividere in sette tappe fondamentali:

1. registrazione del suono;
2. riconoscimento dei singoli enunciati da elaborare. Per distinguerli, il *software* deve stabilire il punto di inizio e di fine di ogni singola parola. Il punto di inizio può essere determinato confrontando livelli audio dell'ambiente circostante con il campione appena registrato. Il punto terminale dell'enunciato è più difficile da determinare in quanto l'utente tende a inserire nel parlato elementi non lessicali, come respiri, rumore di denti, echi, ecc. Esistono essenzialmente due modalità di riconoscimento: i sistemi basati sul riconoscimento di *pattern* confrontano il parlato con dei *pattern* noti o appresi (per lo più morfemi) determinando così delle corrispondenze; i sistemi basati sulla fonetica acustica sfruttano invece conoscenze sul corpo umano (emissione della voce e discriminazione dei foni) per confrontare *feature* del parlato tra loro (proprietà fonetiche come il suono delle vocali o delle sillabe). La maggior parte dei sistemi moderni utilizza l'approccio di riconoscimento di *pattern* perché questo si adatta molto bene alle tecniche computazionali esistenti e tende a presentare migliori valori di accuratezza. Esiste infine un tipo di riconoscimento bimodale, che utilizza sia le informazioni acustiche sia quelle visive, integrandole opportunamente, per migliorare la precisione del riconoscimento soprattutto in ambienti rumorosi[13];

13 Questa strategia di integrazione di informazioni acustiche e visive nel riconoscimento del parlato è, d'altra parte, tipica degli esseri umani, come dimostrato sperimentalmente dal celebre 'effetto McGurk' (cfr. McGurk e MacDonald 1976). Da un punto di vista più tecnico, la giustificazione a priori viene anche dalla considerazione che il canale visivo può essere considerato ortogonale a quello acustico, capace quindi di fornire informazioni di natura diversa, possibilmente integranti quelle fonetiche (cfr. Cosi e Magno Caldognetto 1996). A rendere plausibile il ricorso alle informazioni visive come modalità di riconoscimento la constatazione che il rumore acustico non influenza i dati visivi, nemmeno quando la ricezione acustica non è ottimale. Da un punto di vista pratico, si deve infine sottolineare che il

3. pre-filtraggio (pre-amplificazione, normalizzazione, spostamento di banda, ecc.). I metodi di pre-filtraggio più comuni sono il metodo 'Banco di Filtri', che usa una serie di filtri audio per preparare il campione audio, e la 'Codifica Lineare Predittiva' che usa una funzione di predizione per calcolare gli scostamenti dalla pronuncia standard di una parola. Sono anche utilizzate diverse forme di analisi spettrale;
4. suddivisione dei dati in un formato 'pulito', utilizzabile nella fase successiva di elaborazione dell'*input*;
5. eventuale ulteriore filtraggio di ciascun dato (*frame* o banda di frequenze) durante il quale si effettuano gli ultimi aggiustamenti del campione prima delle fasi di confronto e *matching*. Spesso si fanno operazioni di allineamento temporale e normalizzazione;
6. confronto con le possibili combinazioni tra fonemi e grafemi e *matching* dell'*input* con l'enunciato corrispondente. La maggior parte delle tecniche utilizzate per attuare questa operazione è basata sul confronto del *frame* corrente con dei campioni noti. Ci sono, poi, metodo basati sui cosiddetti HMM (*Hidden Markov Models*), analisi della frequenza, analisi differenziale, tecniche di algebra lineare, distorsione spettrale, distorsione temporale, ecc. Tutti questi metodi sono usati per generare un valore di probabilità e accuratezza del *match*.
7. trascrizione.

Sebbene ogni passo sia distinto dagli altri, le sette operazioni qui descritte costituenti la fase di elaborazione è assai rapida e nei casi migliori è inferiore al secondo. Per funzionare al meglio, i *software* di riconoscimento del parlato hanno bisogno di un *input* accurato che si adegui ai vincoli di riconoscimento imposti dal singolo *software*. A seconda della natura dell'enunciato richiesta si distinguono:

- enunciati isolati (o parlato discreto): il *software* richiede che ciascun elemento da riconoscere (parola, sintagma o breve frase) presenti un periodo di pausa, cioè assenza di segnale audio, su entrambi i lati della finestra di campionamento, cioè sia prima, sia dopo l'enunciazione. All'utente è pertanto richiesto di fare una pausa tra

costo dei sistemi di acquisizione video, grazie al progresso tecnologico, tende a scendere e a rendere quindi concretamente possibile l'impiego di queste apparecchiature anche come apparati complementari.

un enunciato e l'altro, in attesa che il sistema elabori l'enunciato appena incamerato;

- enunciati connessi: un'evoluzione del precedente, che permette la sovrapposizione della fase di dettatura di un enunciato con l'elaborazione dell'enunciato precedente da parte del *software*. L'utente non deve quindi aspettare che il *software* elabori l'enunciato precedente, ma deve comunque scandire bene oltre che le parole anche le pause tra l'una e l'altra;
- parlato continuo: il *software* utilizza tecniche speciali per determinare i confini di un enunciato. Sistemi di riconoscimento basati su questa tecnica permettono all'utente di parlare in maniera quasi del tutto naturale. Sono i sistemi ancora più utilizzati per dettare un testo a un computer;
- parlato spontaneo: il *software* è in grado di elaborare un testo che sembra naturale e non preparato. Il sistema di riconoscimento del parlato basato su questa tecnica è in grado di riconoscere il parlato spontaneo. Il *software* quindi, oltre a determinare i confini di un enunciato, distingue anche le parole dagli elementi non lessicali che possono essere tipici della produzione di un testo orale non preparato e che normalmente inficiano la correttezza del riconoscimento. Si tratta di *software* in continua evoluzione e ancora non accurati al 100% per lingue diverse dall'inglese[14].

Tutti i tipi di riconoscimento del parlato appena descritti possono essere dipendenti o indipendenti da chi parla (rispettivamente *speaker dependent* e *speaker independent*). I sistemi dipendenti sono progettati per soddisfare le esigenze di uno specifico utente. Generalmente, presentano un'elevata accuratezza quando utilizzati dallo stesso utente, ma hanno prestazioni inferiori se usati nello stesso contesto da utenti differenti senza cambiare il profilo vocale. Assumono, infine, che l'utente non modifichi significativamente timbro e ritmo d'eloquio. Al contrario, i sistemi indipendenti sono progettati per essere usati da utenti diversi. Questi sistemi, definiti anche adattivi, di solito funzionano in una prima fase come sistemi indipendenti e poi, utilizzando tecniche di addestramento, si adattano al singolo utente per migliorare la qualità del riconoscimento. Mentre i *software* dipendenti consentono

14 Per l'italiano, il francese e lo spagnolo, il tasso di accuratezza dei software di riconoscimento del parlato è molto vicino al 100%, se usati nel rispetto dei vincoli tecnologici qui descritti.

un buon livello di accuratezza (oltre il 98% per l'inglese e livelli simili per le lingue con una corrispondenza quasi completa tra fonema e grafema), la sfida attuale dell'ingegneristica è l'adattamento al parlato spontaneo di ogni tipo di parlante.

1.2.2 Aspetti tecnici

Gli strumenti che vengono utilizzati per la produzione di sottotitoli si basano essenzialmente su uno dei tre programmi commerciali di riconoscimento del parlato più diffusi:

- *Via Voice* prodotto da IBM;
- *Speech Magic* prodotto da Philips;
- *Dragon NaturallySpeaking* prodotto da Nuance.

Dal punto di vista operativo, un software di riconoscimento del parlato ha bisogno soltanto di un semplice microfono esterno per la canalizzazione della fonte acustica. Per poter funzionare in maniera otimale, i *software* di riconoscimento del parlato spontaneo *speaker dependent* chiedono ad ogni utente di creare il proprio profilo vocale. In altre parole, l'utente legge alcuni brani già presenti nella memoria del *software*. Nell'operare il *matching* (l'abbinare il suono alle parole corrispondenti), il programma adatta gli algoritmi utilizzati al modo di parlare dell'utente, registrando informazioni fisiche riguardo la sua voce e il suo modo di pronunciare le parole (timbro, prosodia, volume, tono, ritmo, ecc.). Se l'utente è in grado di ripetere senza variazioni significative un certo enunciato, il sistema di riconoscimento del parlato dovrebbe essere in grado di adattare il modello costruito per una lingua all'utente in questione effettuando così il riconoscimento con successo.

Una volta creato il profilo vocale, per poter funzionare, il programma richiede all'utente di operare il test audio della scheda sonora del *computer* e dell'ambiente circostante, in modo da permettere al *software* di distinguere il brusio di sottofondo (riverbero, rumori esterni continui, ecc.) dai suoni fonetici emessi dalla voce dell'utente. Terminata questa rapida operazione, il riconoscitore del parlato è pronto per la dettatura. Nel caso di un testo scritto, il *software* richiede all'utente di dettare anche la punteggiatura e tutti i comandi tipografici necessari a dare la corretta forma al testo. Oltre a questo *matching*, tra comandi vocali e operazioni, i *software* di riconoscimento del parlato offrono una vasta gamma di strumenti per migliorare l'accuratezza del riconoscimento. Essi sono:

- il vocabolario di base: si tratta di liste di parole o enunciati che il sistema riconosce senza che l'utente debba correggerle,[15] dettarne l'ortografia o inserirle nel vocabolario seduta stante o in un secondo momento. Generalmente, vocabolari di dimensioni minori permettono un riconoscimento migliore da parte del computer (perché meno sono le parole concorrenti a un'unica pronuncia), mentre vocabolari più estesi creano maggiori difficoltà di riconoscimento. A differenza dei normali dizionari, ciascun elemento presente nel dizionario di un sistema non deve necessariamente essere una singola parola, ma può anche essere una o più frasi. Tali dizionari, infine, sono aperti, consentono cioè di introdurre un numero illimitato di elementi, a detrimento però della rapidità e dell'esattezza del riconoscimento. Ogni elemento presente nel vocabolario di base ha infatti un indice di frequenza che aumenta a seconda dell'uso che si fa dell'elemento stesso. A *input* simile, quindi, la parola con l'indice di frequenza più alto sarà trascritta. Come appena accennato, meno saranno i concorrenti, più rapida e corretta sarà quest'operazione;
- i vocabolari specialistici: creati per evitare di appesantire troppo il vocabolario di base, i vocabolari specialistici sono composti da parole caratterizzate da un indice di frequenza superiore a qualsiasi altra parola del vocabolario di base. Nel caso di contesti particolari (come per esempio la sottotitolazione di telecronache dei campionati mondiali di calcio, la resocontazione dell'audizione della commissione Bilancio della Camera dei Deputati, la trascrizione della telefonata con un tecnico informatico, ecc.), il vocabolario specialistico, opportunamente creato e attivato, permette di far riconoscere al *software* termini che in condizioni normali sarebbero stati di difficile riconoscimento. Questi vocabolari sono utili nel caso di nomi propri, tecnicismi o formule specifiche;
- analisi documenti: è una funzione che analizza determinati documenti scritti alla ricerca di termini dalla grafia ignota, sia termini sconosciuti, sia termini noti ma con una grafia differente da quella contenuta nel dizionario di base. In quest'ultimo caso, il sistema chiederà all'utente di disambiguare il termine dal relativo omofono.

15 Qualora il software trascriva una parola in luogo di quella desiderata, per la correzione, l'utente è chiamato a dettare le due pronunce in modo che in un secondo momento il riconoscitore non le confonda ancora. Nel caso invece di parole sconosciute, il programma comunque scrive qualcosa di foneticamente analogo all'input ricevuto. Dopo aver corretto l'errore di trascrizione, l'utente dovrà introdurre la parola nuova e fornirne l'ortografia.

Attivata prima di iniziare la dettatura, la funzione 'analisi documenti' prepara il riconoscitore a parole che possono comparire nel contesto dato;

- *house-style*: si tratta di una funzione specifica dei riconoscitori per cui frequenti errori ortografici o di trascrizione sono corretti e scritti nella maniera opportuna. Sono molto impiegati per le sigle e nomi propri che non siano omofoni di altri termini dei dizionari. Anche in questo caso, si distingue tra *house-style* di base, che si applicano cioè alla dettatura in generale, e *house-style* specialistiche, che compaiono cioè solo se attivate;
- macro di dettatura: molto utilizzate nel caso di formule rituali (apertura dei lavori, titolo di una persona, espressioni ricorrenti, ecc.), nomi propri omofoni di altri termini presenti nei vocabolari (come per esempio Prodi, Tasso, Elefante, ecc.), termini che svolgono una particolare funzione nel testo (nome dell'oratore, rappresentazione della componente non verbale di un programma, ecc.) o altro ancora, sono una scorciatoia per ottenere il massimo risultato con il minimo sforzo. Come riassume bene Marsh (2005)

> (f)or example, in sport where you have lots of crowd noises and you have to label them like 'APPLAUSE' or 'CHEERING' or 'LAUGHTER' and you want a label to come out basically defining the noise, you need a label which respects certain criteria. It has to be centred and in white capitol letters, respecting the BBC style. If I couldn't use macros, to respect the BBC style, instead of simply saying 'applause-macro', I had to say 'new line, centred, white, upper case, applause'. So, simply saying 'applause macro' is much, much quicker than saying all that.

Durante la dettatura, l'utente deve pronunciare il testo e i comandi per la corretta formattazione dello stesso (grassetto, maiuscole, a capo, giustificato, ecc.). La correzione del testo in corso di formazione può essere effettuata sia in tempo reale, grazie a *software* specifici che consentono una rapida manipolazione del testo riconosciuto prima della 'pubblicazione' del testo, sia in un secondo momento. Da queste correzioni, il software imparerà nuovi termini che dovranno essere opportunamente inseriti nel sistema, onde evitare che, in un secondo momento, gli stessi termini siano nuovamente riconosciuti in maniera scorretta.

1.2.3 Difficoltà operative

Indipendentemente dalla professionalità del rispeaker e dalla sua dimestichezza con i *software* di riconoscimento del parlato, esistono delle difficoltà intrinseche nella professione del rispeaker a cui sono state trovate soluzioni, ma di cui è bene essere a conoscenza. È possibile infatti che alcune di queste difficoltà influenzino in maniera definitiva l'*output*. Queste difficoltà possono essere dovute ai limiti tecnologici del *software* in uso o a limiti umani. Nel primo caso, la prima difficoltà, o meglio il primo problema con cui il rispeaker deve convivere è il *décalage* con cui compare il testo scritto rispetto al momento della sua ideazione. Nel caso della dettatura di una lettera, il tempo che intercorre è di circa un secondo nel migliore dei casi. Qualora si volesse sottotitolare qualsiasi testo orale, i sottotitoli comparirebbero sullo schermo con un *gap* rispetto al momento della sua produzione spesso frustrante. Nelle conferenze, in cui l'eloquio non è alternato a una componente video significativa (diapositive, foto, grafici, ecc.), anche solo cinque secondi non sarebbero troppi, perché non vi è necessità di sincronizzare il testo alle immagini. Nel caso di programmi televisivi (TG, competizioni sportive, documentari, ecc.), in cui il significato del testo, nella sua totalità, è il prodotto di una forte interazione multimodale (cfr. Baldry e Thibault 2005), gli stessi cinque secondi causerebbero una diacronia tale tra i sottotitoli e le immagini da rendere difficile l'attribuzione di un dato enunciato a un dato oratore e/o a una data sequenza di immagini, rendendo così incomprensibile il senso generale del discorso. Un'altra difficoltà è dovuta all'incapacità dei *software* oggi sul mercato di adattarsi talmente tanto al parlato da garantire una trascrizione esatta di eventuali parole enunciate dal rispeaker non presenti nel suo vocabolario. Grazie a una politica di correzione *ad hoc* è possibile evitare tale problema a discapito però della velocità di trascrizione. Con un'interfaccia di correzione, il rispeaker stesso o un assistente può infatti rettificare un'eventuale errore del *software* di riconoscimento prima della sua messa in onda, ma la perdita di anche pochi secondi può addirittura raddoppiare i tempi di comparsa del sottotitolo sullo schermo. Un'altra difficoltà tecnica è dovuta alla qualità dell'*input*: anche se il rispeaker evitasse ogni azione di disturbo alla buona comprensione da parte del *software*, articolando bene le parole, annullando il suo accento il più possibile, evitando di produrre pause piene, è comunque possibile che un minimo cambiamento nel rumore ambiente sia interpretato come un prodotto linguistico da parte del *software* allungando così i tempi di elaborazione dell'*input*. Sorte simile è dovuta a fattori del tutto aleatori, peraltro non rari. Un

esempio abbastanza comune è il *software* che interpreta una parola come un suo omofono. Visto che tutti i *software* di riconoscimento del parlato operano una minima analisi sintattica prima di ogni trascrizione, una diversità morfo-sintattica tra gli omofoni causa un rallentamento di elaborazione dell'*input*. Un'ultima difficoltà tecnica è dovuta alla qualità della scheda audio. Secondo studi condotti dal dipartimento di ricerca della CNN (cfr. Mellor 1999), una buona scheda audio (molto più di un buon microfono) può aumentare il livello di accuratezza del 5%, sebbene un'accuratezza del 100% sia ancora oggi impossibile. Questo dato è sicuramente cambiato in questi ultimi anni, dato che anche pochi anni rappresentano un salto generazionale importante nel mondo dell'informatica. Sicuramente, l'accuratezza è oggi meno influenzata da alee sonore e il motore di riconoscimento riesce a fare maggiore astrazione dagli eventi non-lessicali, ma la qualità della scheda audio resta un elemento importante ai fini di un buon riconoscimento del parlato.

Per quanto riguarda la seconda tipologia di difficoltà, quelle derivanti da limiti umani, il primo ostacolo che il rispeaker deve affrontare è dato dalla velocità di produzione del TP. Si tratta di una variabile indipendente dal rispeaker e che determina la quantità di compressione che un operatore deve effettuare se vuole mantenere un certo standard di accuratezza. Sebbene l'*access service* della BBC parli di una soglia massima tollerabile di 300 parole al minuto (cfr. Marsh 2005), in realtà la velocità media di eloquio dei programmi in diretta e semidiretta per cui viene fatto uso del rispeakeraggio non supera mai le 200 parole al minuto. Al fattore temporale si aggiunge quello metalinguistico. Per rendere comprensibile un sottotitolo è infatti necessaria una minima impaginazione. Sebbene sia già stato fatto notare che alla BBC si fa uso della modalità di proiezione *scrolling*, per cui i sottotitoli compaiono parola per parola sulla schermo e non in blocco come nel caso dei sottotitoli in pre-registrato, la leggibilità del sottotitolo può comunque essere ottenuta tramite la dettatura della punteggiatura, delle maiuscole e delle didascalie esplicative. Il continuo alternarsi dell'uso della lingua per dettare il TA e per impaginarlo non solo aumenta il numero di parole al minuto che il rispeaker deve dettare al *software*, ma lo distrae dalle quattro fasi di ascolto e comprensione del TP e di elaborazione e dettatura del TA, sovraccaricando la sua *capacité de traitement* (cfr. Gile 1995). Con l'esperienza, questa operazione metalinguistica diventa tuttavia quasi automatica (cfr. Marsh 2005), favorendo così un maggiore equilibrio tra gli sforzi che il rispeaker deve compiere. Sempre in ambito metalinguistico, altre sono le operazioni quasi automatiche che devono essere attuate per una

corretta impaginazione dei sottotitoli, come il cambio di colore per notificare il cambio di oratore e il posizionamento del sottotitolo sullo schermo onde evitare di sovrapporlo alla bocca degli oratori o a eventuali didascalie. Queste operazioni vengono effettuate tramite un apposito *hardware* esterno delle dimensioni di una calcolatrice, che sfrutta il potenziale manuale del rispeaker a vantaggio di quello vocale, comportando così un sovraccarico minore rispetto all'eventuale impaginazione effettuata con la voce.

1.2.4 Tecniche di scrittura rapida

La produzione rapida di testo è possibile grazie a varie tecniche. In generale, essa permette all'operatore di ottenere una traccia scritta di un testo orale non letto o di un testo scritto non digitalizzabile. L'operazione dovrà avvenire nel minor tempo possibile, nel rispetto del messaggio originale e in ottemperanza alle convenzioni grammaticali e tipografiche del TA. Applicata alla sottotitolazione televisiva, la produzione rapida di testo deve anche tenere in debita considerazione la leggibilità dei sottotitoli e la loro interazione con le altre componenti del TP (componente audio verbale e non-verbale e componente video verbale e non-verbale). A tal proposito, il rispeakeraggio è soltanto una delle possibili modalità di produzione di sottotitoli in tempo reale. In particolare, come sottolinea Lambourne (2006), esistono anche altre forme concorrenti che sono state comunque sviluppate precedentemente l'introduzione del riconoscimento del parlato:

- *Dual QWERTY system*;
- *Stenotyping*;
- *Velotyping*.

Il *dual QWERTY system* è una forma di scrittura veloce ancora in uso e che consiste nella produzione di testo da parte di due dattilografi che, coordinati tra di loro in modo da concentrarsi ognuno su una parte di testo diversa e grazie all'uso di forme abbreviate, riescono a produrre mediamente 550 caratteri al minuto.

Figura 2: *Dual QWERTY system*. Due dattilografi trascrivono ognuno una parte di testo diversa.
© SysMedia LTD.

La stenotipia è la forma di scrittura veloce più utilizzata per produrre testi in maniera rapida e accurata. Si tratta di una tecnica di "scrittura abbreviata su una tastiera che permette di digitare sillabe anziché lettere isolate" (Trivulzio 2006), come avviene invece in dattilografia. Grazie alla possibilità di scrivere sillabe (talvolta anche bisillabi) invece che lettere, i tempi di trascrizione possono arrivare anche a 900 caratteri al minuto.

Figura 3: Stenotipia. Il sistema sillabico su cui si basa
permette di digitare sillabe anziché caratteri. © IVD Spinea.

La macchina *Velotype* assomiglia per molti versi alla stenotipia e permette di digitare il corrispettivo grafemico dei singoli fonemi anziché l'ortografia standard. Si tratta di un sistema molto utilizzato per quelle lingue con una rispondenza grafema-fonema lontana dal 100%. Applicato alla lingua inglese può raggiungere una velocità di 600 caratteri al minuto (cfr. Lambourne 2006).

Figura 4: *Velotyping*: Il sistema fonetico su cui si basa permette di digitare la forma grafemica di singoli fonemi anziché l'ortografia. © SysMedia LTD.

1.2.5 Applicazioni del riconoscimento del parlato

Come si è visto, la tecnologia di riconoscimento del parlato non è l'unica a garantire un'immissione rapida di dati di modo da poter produrre un sottotitolo in tempo reale. Allo stesso modo, la sottotitolazione in diretta non è l'unica applicazione possibile del riconoscimento del parlato. Anzi, la sottotitolazione è una delle ultime applicazioni di questa tecnologia che ha trovato una sua applicazione già a partire dagli anni Settanta, quando veniva utilizzata come ausilio dai professionisti in ambito medico, politico, giuridico, meccanico, ecc., vale a dire al servizio di persone con esigenze di produrre un testo scritto, ma che per motivi professionali (l'esigenza di avere il testo in tempi strettissimi, non poter utilizzare le mani, ecc.) non potevano, e tuttora non possono, utilizzare le altre tecniche esistenti. Nonostante si trattasse di testi abbastanza brevi, con un vocabolario limitato e dalla

forma molto ritualizzata (referto medico, sentenza giudiziale, diagnosi medica o meccanica, ecc.), il riconoscimento del parlato permetteva alle persone che ne facevano uso di risparmiare tempo e denaro. Oltre che per la scrittura di un testo, i *software* di riconoscimento del parlato sono stati utilizzati anche come centralini automatici dei più svariati servizi, come il servizio informazioni delle Ferrovie dello Stato. Sempre negli anni Settanta, si è iniziato a intuire che questa tecnologia poteva venire in aiuto ai disabili. Ecco quindi che si è arrivati all'ideazione del Dispositivo Telefonico per Sordi (DTS), un apparecchio che, come dimostra la figura 5, permette a una persona con problemi di udito di avere una conversazione telefonica con un udente o con un'altra persona sorda. In particolare, nel caso di un sordo che chiama un normoudente, dopo aver composto il numero di telefono desiderato, la persona non udente aspetta che l'udente risponda. Automaticamente, il *software* di riconoscimento del parlato trascrive la risposta dell'udente. Il sordo legge 'la voce dell'udente' e ribatte tramite l'uso della propria voce (qualora la sordità non abbia affetto il suo apparato fonatorio) o scrivendo un testo che viene poi messo in voce da un sintetizzatore vocale.

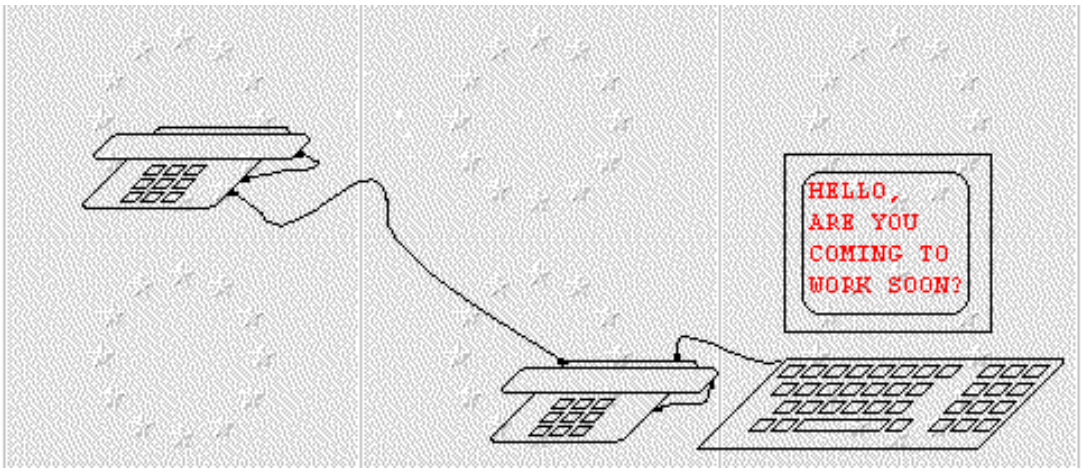


Figura 5. DTS: La voce di chi risponde viene riconosciuta dal *software* e trascritta sullo schermo della persona non udente. © *JRC VOICE project*.

Un'evoluzione di questo dispositivo è rappresentata dall'applicazione dello stesso sistema al video-telefono. In questo caso, il riconoscimento del parlato si interfaccia con uno speciale dispositivo, *Voicemeeting*, frutto del progetto VOICE del Centro Comune di Ricerca della Commissione Europea[16]. Si tratta di un'interfaccia che si applica al *software* di

16 Cfr. http://voice.jrc.it

riconoscimento del parlato Dragon NaturallySpeaking e che è in grado di mixare il testo e l'immagine proiettando così sullo schermo della persona sorda il volto dell'interlocutore e i relativi sottotitoli (figura 6).

Figura 6. Videochiamata: il volto della persona chiamata compare al di sopra dei sottotitoli. © *JRC VOICE project*.

Come risulta evidente dalla figura 6, si può parlare in questo caso di una forma di sottotitolazione in diretta. Sempre in ambito di sottotitolazione in tempo reale, sembra interessante sottolineare l'importanza di un'altra applicazione del riconoscimento del parlato dal forte impatto sociale, la sottotitolazione delle conferenze, delle lezioni universitarie e scolastiche, delle omelie e di ogni evento simile. In questi casi, l'uso del riconoscimento del parlato può avvenire in due modi: sottotitolazione automatica e rispeakeraggio. Nel primo caso, l'oratore si autosottotitola, nel senso che il *software* di riconoscimento del parlato viene utilizzato per generare sottotitoli direttamente dalla voce dell'oratore (figura 7). Si tratta di un'operazione macchinosa, che necessita, da parte dell'oratore, la consapevolezza della presenza di questa tecnologia e la conseguente attuazione di determinate precauzioni nel rispetto dei criteri di leggibilità (presenza della punteggiatura, assenza di elementi non-lessicali[17], rispetto delle regole lessico-grammaticali sottostanti la produzione di testi scritti, ecc.). Non c'è da dimenticare, inoltre, che i *software* di riconoscimento del parlato non sono

17 Cfr. Savino et al. (1999: 2).

speaker-independent al 100%, ma hanno bisogno di un profilo vocale per ogni oratore da cui poter trarre la giusta chiave per riconoscere il parlato di ognuno. Risultano quindi chiari la macchinosità e i limiti di una tale operazione.

Figura 7. Conferenza autosottotitolata. Il volto dell'oratore compare al di sopra dei sottotitoli.
© *JRC VOICE project*.

Nel secondo caso, invece, la presenza di un professionista garantisce, oltre a una maggiore correttezza formale del sottotitolo, anche una maggiore flessibilità nell'alternarsi degli oratori, nel senso che chiunque potrà essere sottotitolato senza dover prima addestrare il *software* (figura 8).

Figura 8. Rispeaker sottotitolano una conferenza. © AIR.

Oltre alla sottotitolazione in tempo reale, il riconoscimento del parlato è utilizzato anche per la sottotitolazione di programmi pre-registrati. In particolare, il *software* di riconoscimento del parlato si può interfacciare sia direttamente con la fonte da cui riceve in *input* il TP e conseguentemente con il *software* di sottotitolazione per la messa in onda del filmato sottotitolato (figura 9), sia con un operatore che introduce il TA tramite la sua voce.

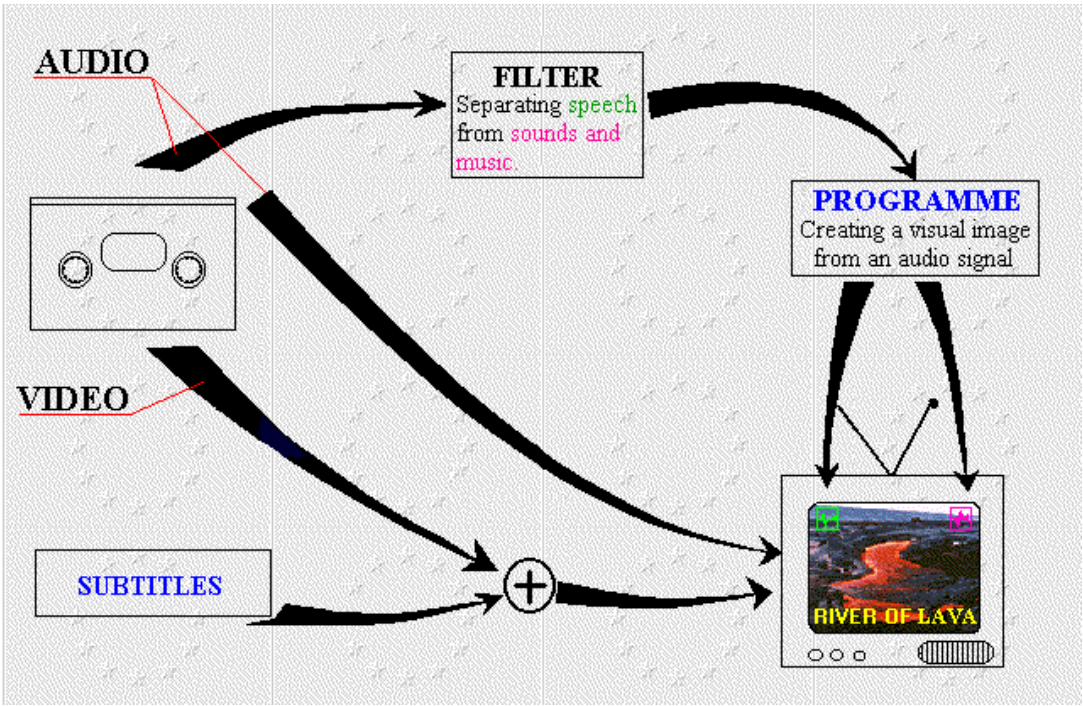


Figura 9. Procedura di sottotitolazione automatica di un testo audiovisivo. © *JRC VOICE project*.

Nel primo caso, l'aspetto maggiormente problematico sta nella necessità di correggere, in un secondo tempo, il testo così come è stato riconosciuto dal *software* ripulendolo dagli inevitabili errori di riconoscimento. Una possibile soluzione sta nell'ulteriore interfacciamento di questi tre *software* con un programma di testo. In questo caso, il *software* di riconoscimento del parlato non farebbe altro che abbinare il testo all'audio azzerando così i tempi di correzione. Si otterrebbe così un filmato perfettamente sincronizzato e assolutamente fedele al TP[18], anche se in violazione delle norme di leggibilità che vorrebbero che i sottotitoli, opportunamente rivisti nella forma, diano allo spettatore il tempo necessario alla lettura e che all'interno di ogni sottotitolo sia presente un'intera unità sintattica.

Nel secondo caso, la presenza di un operatore garantisce in qualche modo un controllo continuo sul TA in fase di produzione. Purtroppo, oltre ad aumentare i costi di produzione, la sua presenza pone anche dei dubbi deontologici sulla natura dei sottotitoli dibattuti sia in sedi professionali, sia nel settore della ricerca. Con l'uso di questa tecnologia, infatti, vengono sì abbassati i tempi (rispetto a un sottotitolatore che introduce testo con la classica tastiera) e i costi della manodopera (rispetto a un sottotitolatore che introduce testo con la stenotipia o con la macchina *velotype*), ma, come afferma Díaz-Cintas, "the quality of the final product would invariably decrease" (2007). Paradossalmente, il motivo di questa affermazione sta proprio nella velocità di immissione del testo scritto e nella conseguente difficoltà da parte dei sottotitolatori di rendersi conto di eventuali errori di trascrizione e soprattutto di impaginazione del sottotitolo che, come si vedrà più avanti, deve rispettare precise norme spazio-temporali per poter assolvere al suo compito. Una volta introdotto il testo nel *software* di sottotitolazione, infatti, come affermava già Donaldson, esperto di sottotitolazione per non-udenti e audioleso a sua volta, il sottotitolatore "should revise, afterwards, what he or she has produced, thus cancelling out the gain in terms of time"[19]. Ecco quindi che l'uso del rispeakeraggio in televisione, pur garantendo il rispetto degli impegni presi in termini di ore di programmazione sottotitolata, molti errori rimarranno, a discapito dell'utente finale.

18 Questa operazione da parte dello spettatore sordo si è dimostrata più lenta rispetto alla normale codifica di un testo orale tramite l'apparato acustico (cfr. Volterra 1986 e Karamitroglou 1998).
19 Intervento alla tavola rotonda sull'accessibilità all'interno della conferenza internazionale Languages and the Media 2004.

Sempre nell'ambito dell'accessibilità, i *software* di riconoscimento del parlato sono, da alcuni anni, applicati alla navigazione su internet da parte di utenti disabili motori. Nel caso di un disabile non in grado di utilizzare le mani per scrivere, il *software* di riconoscimento del parlato gli permette di navigare al pari di un normodotato, previo un certo periodo di applicazione e di addestramento del *software*.

Un'ultima interessante applicazione dei *software* di riconoscimento del parlato risale, per l'Italia, al 2001, anno in cui è iniziato ufficialmente l'uso del riconoscimento del parlato alla Camera dei Deputati Italiana per la produzione di resoconti in luogo della tradizionale stenografia. Grazie a questa tecnologia, la produzione di un resoconto integrale è resa assai più dinamica con un conseguente risparmio in termini di tempi, costi e manodopera (cfr. Arma 2007).

1.3 Il rispeakeraggio televisivo

Finora si è visto il rispeakeraggio in generale come tecnica di produzione di testo in tempi rapidi. Nel tentativo di affrontare più da vicino il rispeakeraggio come tecnica di sottotitolazione in tempo reale per la televisione, si cercherà, ora, di approfondirne tutti gli aspetti inerenti la fase di produzione. Il prodotto finito, che è l'oggetto di questa tesi, sarà invece affrontato in un secondo momento, grazie al contributo degli studi sulla traduzione audiovisiva e all'analisi delle migliori prassi, raccolte in un corpus di otto ore di registrazione di programmi rispeakerati.

Come si è visto, il rispeakeraggio come processo è un'operazione isosemiotica che comporta, per la produzione del TA (o meglio del TM), l'utilizzo dello stesso canale impiegato dal mittente del TP. In realtà, il rispeakeraggio non è una semplice ripetizione del TP e non è nemmeno una modalità applicabile a tutti gli ambiti. Anche all'interno della stessa applicazione, come si vedrà, molti sono i fattori che influenzano la fase di produzione del TA, nel caso in questione il sottotitolo. Inoltre, proprio per le difficoltà intrinseche, sia tecniche sia linguistiche, il rispeakeraggio è un'operazione molto più complessa che coinvolge numerose attività intellettuali e tecniche. Sempre in chiave descrittiva, si può quindi scorporare il processo in due sottocategorie: i fattori che influenzano la fase di produzione e le competenze che deve applicare il rispeaker in ogni situazione di lavoro in cui si trova a operare.

1.3.1 I fattori d'influenza

Per mostrare le varie sfaccettature del rispeakeraggio come tecnica di sottotitolazione televisiva, Lambourne (2007) elenca una serie di fattori che hanno un'influenza importante sul lavoro del sottotitolatore in tempo reale. Essi sono:

- metodo di trascrizione (rispeakeraggio, stenotipia, *dual-keyboard*, *velotyping*);
- modalità di produzione (differita, diretta, semidiretta);
- proiezione del programma (differita, diretta, semidiretta);
- politica di *editing* (sottotitoli integrali o *verbatim* e adattati o *non verbatim*);
- metodo di correzione dell'*output* (autocorrezione, da altro operatore, nessuno);
- metodo di visualizzazione (*pop-on* e *scrolling-rolling up*).

Ai fini prefissati, mentre il primo fattore è stato già ampiamente discusso, sembra interessante riprendere gli altri cinque fattori e approfondirli anche alla luce di quanto già emerso da un prima disamina del rispeakeraggio.

Modalità di produzione

La modalità di produzione ha un forte impatto sul lavoro del rispeaker. Lavorare in differita o in diretta ha delle conseguenze enormi in termini di concentrazione. Nel caso della differita, infatti, il rispeaker procede alla sottotitolazione di un programma tramite dettatura. Sebbene la maggior parte del lavoro avvenga in contemporanea con l'ascolto del TP, il rispeaker sa che il TM non sarà ancora proiettato sullo schermo. Questo implica che si può fermare in qualsiasi momento e riascoltare un pezzo che non ha ben compreso. Per quanto riguarda le parole non conosciute dal *software*, il sottotitolatore non si deve preoccupare di pensare quali possono creare problemi di riconoscimento e può sia correggere parole scritte male, sia prendersi il tempo di addestrare il *software* a una parola ricorrente. Infine, non si deve preoccupare del ritardo che accumula nei confronti del TP perché sa che *a posteriori* sarà effettuata la sincronizzazione automatica con il TA.

Un'esperienza profondamente diversa ma con qualche aspetto comune alla modalità precedente è vissuta nel caso di una produzione in semidiretta, cioè poco prima della reale messa in onda del programma che si sta sottotitolando. È il caso della sottotitolazione dei telegiornali britannici durante i quali i rispeaker possono avere accesso ai servizi che andranno in onda dopo qualche minuto, ma il cui ordine di proiezione non è ancora deciso.

Il rispeaker ascolta i servizi e li prova a sottotitolare alla ricerca di eventuali parole ignote che provvederà a inserire nel vocabolario del *software*. Quando il notiziario inizia, il rispeaker lo sottotitola come se fosse in diretta, ma con il vantaggio di conoscere anticipatamente il contenuto dei servizi e sopratutto di non doversi preoccupare di eventuali parole ignote che, anzi, saranno riconosciute correttamente.

Più prototipica è infine la situazione lavorativa del rispeaker che deve sottotitolare un programma in diretta, come avviene nel caso delle telecronache sportive. In questo caso, il rispeaker si trova a sottotitolare in tempo reale un programma che non ha mai visto prima. In queste situazioni, i cui aspetti psico-cognitivi saranno approfonditi ulteriormente nel corso della tesi, il rispeaker è sottoposto a uno *stress* continuo determinato dall'impossibilità di commettere errori, dalla necessità di evitare che il *software* commetta errori di riconoscimento, dall'obbligo di colmare il più possibile il divario in termini di tempo tra il testo originale e i sottotitoli. Nel paragrafo successivo, si analizzeranno le competenze necessarie a espletare questa modalità di sottotitolazione.

Proiezione del programma

Affine e strettamente collegato al fattore precedente è la modalità di proiezione del programma. Per quanto riguarda i programmi in differita e i programmi in diretta, le ripercussioni sul rispeaker saranno le stesse descritte sopra. Per quanto riguarda la semidiretta, invece, la situazione cambia sensibilmente rispetto a quanto sopra riportato perché il rispeaker lavora in una modalità più simile alla differita che alla diretta. In specifico, il rispeaker sottotitola il programma in differita, ma, visti i tempi ristretti, attribuirà meno importanza al *layout* del sottotitolo e alla sua sincronizzazione con il testo audiovisivo, che avverrà manualmente o con i *software* di sottotitolazione automatica brevemente descritti nel paragrafo precedente (figura 9). Un esempio di questa modalità di produzione è la sottotitolazione dei notiziari italiani o la sottotitolazione di eventi particolari. Ad esempio, nel caso del dibattito televisivo tra Romano Prodi e Silvio Berlusconi, candidati alle legislative del 2005 per la costituzione della nuova legislatura. Nel caso in questione, la società di resocontazione Cedat85 ha rispeakerato il testo in diretta senza metterlo in onda. Parallelamente alla sua produzione, un secondo operatore correggeva il TM e lo allineava con il TP. Con solo mezz'ora di ritardo rispetto alla diretta, il dibattito

perfettamente sottotitolato è stato proiettato in *streaming* sul sito del Centro d'Ascolto della Rai.

Politica di editing

La decisione di intervenire sul testo può dipendere da diversi fattori, ma ha un forte impatto sulla memorizzazione del TA e la produzione del TM da parte del rispeaker. In particolare, questa decisione dipende sia da fattori esogeni, come la politica di accessibilità adottata dall'emittente (accessibilità significa far comprendere i propri programmi agli utenti o garantire loro lo stesso testo dei normoudenti), le richieste fatte dalle associazioni in difesa dei non-udenti (molte associazioni richiedono la trascrizione esatta del TP) o le pressioni esercitate dalle categorie sindacali interessate[20]; sia da fattori endogeni come la velocità di trascrizione e di proiezione del TM, la velocità di eloquio degli oratori e il genere del programma da sottotitolare. Si tratta, quest'ultimo, di un aspetto che merita sicuramente un approfondimento, in quanto, a seconda dei programmi, molti fattori interverranno a favore di una forma di sottotitolazione o l'altra. In particolare, qualora il TP è particolarmente lento, il rispeaker potrà, e per non annoiare troppo lo spettatore sordo dovrà, riportare integralmente il TP. Nel caso invece di un programma in cui la velocità di eloquio è molto elevata, come per esempio le sedute parlamentari o le telecronache sportive, l'approccio è duplice: se il TP è troppo rapido, al di sopra delle possibilità di dettatura, trascrizione e proiezione del TA, allora sarà necessario apportare alcune riduzioni del TP. Ma quando le condizioni non ostacolano una trascrizione *verbatim*, la responsabile per la formazione dell'ufficio respeaking della società che produce i sottotitoli per la BBC, RedBee Media, sostiene che:

> The news and the parliamentary sessions, being particularly fast, you have to go along with them. While, if you subtitle sport, the idea is that you describe the action you can see on the screen so you do not need to speak all the time. We chose to edit much more with sport events than with the news or the parliament. (Marsh 2005)

Un'ultima considerazione va forse fatta a proposito della capacità di rielaborazione dell'operatore. Bisogna ricordare infatti che solo da pochi anni si è iniziato a utilizzare i

20 In RAI, il sindacato dei giornalisti richiede la trascrizione esatta del TP per i testi in differita e in semidiretta e, nel caso della diretta, la presenza di un giornalista che riformula il TP e detta alla stenotipista quello che dovrà essere il TA che lo trascrive (cfr. de Seriis 2006).

rispeaker per produrre sottotitoli televisivi e la domanda è ancora molto bassa. Conseguentemente, non esiste ancora una vera e propria didattica del rispeakeraggio televisivo e molti rispeaker si sono improvvisati tali senza una vera formazione *ad hoc*. Intuitivamente, ne consegue che un rispeaker non abituato a operare in situazioni di lateralizzazione (una parte del cervello opera l'ascolto e l'altra procede, oltre che alla dettatura del TM, anche alla sua rielaborazione[21]), tenderà più a una trascrizione *verbatim* del TP o all'eliminazione di alcune componenti piuttosto che a una sua riformulazione.

Indipendentemente dai fattori che influenzano l'intervento da parte del rispeaker sul TP, la 'manipolazione' che viene effettuata in fase di riformulazione avviene sostanzialmente su due livelli: quantitativo e qualitativo. Il primo concerne la riduzione del tasso di parole pronunciate al minuto e può essere effettuato sia rimuovendo le caratteristiche tipiche dell'oralità (ripetizioni, false partenze, evidenti ridondanze, ecc.), sia effettuando alcuni tagli strategici a livello sintattico (frasi incidentali, elementi simili in una lista, ridondanze, ecc.). In nessun caso si procede a una riformulazione volta alla semplificazione, o meglio all'accessibilità del TP per una determinata categoria di spettatori. Questa operazione è invece tipica del secondo livello. Si tratta di un'operazione molto comune in interpretazione di conferenza in quanto inevitabile nel passaggio tra lingue diverse, per definizione non isomorfe[22]. Tuttavia, sembra che la tendenza nella sottotitolazione intra-linguistica sia la resa *verbatim* del testo, sia per quanto riguarda il mercato dei DVD[23], sia i sottotitoli in tempo reale[24]. Intuitivamente, la ragione sta nella co-presenza, all'interno dello stesso testo, del TP e del TA e nella conseguente possibilità di confrontare le due versioni[25]. In situazioni del genere, è facile da parte di un non-udente cadere nell'equivoco e derivare da una superficiale analisi contrastiva eventuali intenti paternalistici da parte dell'emittente o del sottotitolatore nei confronti dei sordi.[26]

Metodo di correzione dell'output

21 Cfr. Gran 1999.
22 Cfr. Gile 1995 e Pöchhacker 2002.
23 Cfr. Neves 2004.
24 Cfr. Eugeni 2007.
25 Si ricorda che la maggioranza delle sordità non è di tipo genetico e che, solo in pochi casi, tutti i membri di una famiglia di sordi sono realmente sordi. È quindi abitudine ricorrente confrontare il TP e il TA quando si guardano programmi sottotitolati intralinguisticamente.
26 Cfr. Mereghetti 2006.

La correzione del TM prima che venga messo in onda è un'opzione non sempre possibile e, in definitiva, dipende anch'essa da numerosi fattori, primi fra tutti il divario medio tra l'emissione del TP e la comparsa dei sottotitoli sullo schermo e l'interfaccia utilizzata per la visualizzazione dei sottotitoli in fase di produzione. La questione del divario è la maggiore imputata per la presunta scarsa qualità dei sottotitoli in tempo reale. Avere dei sottotitoli che compaiono sullo schermo non in perfetta sincronia con la pronuncia del TP appare infatti come un difetto di questa modalità a causa del quale emittenti di prestigio non possono permettersi di adottarla. In realtà, si tratta di una necessità fisiologica. Produrre sottotitoli in tempo reale richiede un certo tempo di reazione, oltre che da parte del sottotitolatore (che in un futuro assai prossimo potrebbe non essere più necessario[27]), anche da quella del *software*. In particolare, oltre ad alcune ragioni già menzionate (comprensione e reazione del rispeaker, tempo di elaborazione del *software*, ecc.), questo ritardo è da imputarsi anche alla 'comprensibilità' del TP, alla qualità dell'*input* vocale del rispeaker e al tipo di rilascio delle parole da parte del *software*.

A seconda del ritardo, ma anche a seconda del tipo di errore riscontrato e dello status dei sottotitoli all'interno del programma sottotitolato (se i sottotitoli sono parte integrante della pellicola o se sono attivabili dal *teletext*[28]), la correzione può essere consigliabile o meno. Nel caso più comune di sottotitoli attivabili dal *teletext*, un errore che non inficia la comprensione del TA è solitamente tralasciato. Un errore che invece viene considerato importante è solitamente corretto se il ritardo lo consente[29]. Altrimenti, la scelta starà al sottotitolatore, che potrebbe continuare ad aumentare il divario o decidere di recuperarlo, riassumendo o eliminando quel che segue l'errore.

Il secondo fattore importante da cui dipende la correzione è l'interfaccia utilizzata per la visualizzazione dei sottotitoli in fase di produzione. Nel caso della classica interfaccia (figura 10), ci sono tre schermi: sul primo compare il TP, ossia il filmato originale così come viene trasmesso dall'emittente; nel secondo, il TM, che può essere trattenuto per essere corretto o essere direttamente spedito al terzo; il terzo schermo, infine, è quello in cui è visualizzato il TA, cioè il filmato originale e i sottotitoli dopo il via libero del correttore.

27 Cfr. Accademia Aliprandi et al. 2007.
28 Cfr. de Korte 2006.
29 Cfr. Marsh 2005.

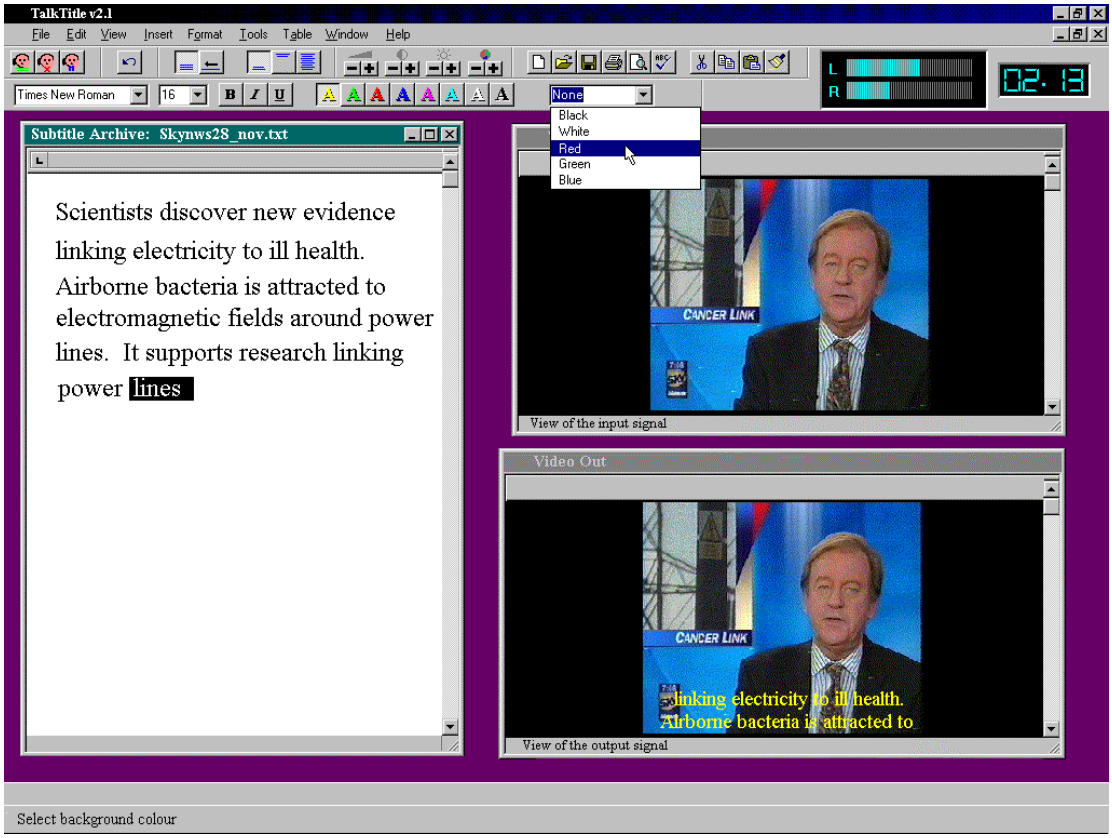


Figura 10. Interfaccia per la visualizzazione dei sottotitoli in fase di produzione. © *JRC VOICE project*.

A seconda della flessibilità del *software* in questione, le ripercussioni sulla correzione possono essere rilevanti. Nella migliore delle ipotesi, il correttore deve continuamente tenere d'occhio il secondo schermo con il TM, in continua formazione, e la sua rispondenza al TP. Nel momento in cui individua un errore che valuta come da correggere, il correttore seleziona la parola da correggere, la corregge prestando attenzione che si accordi con il resto del sottotitolo e da il via libera definitivo alla sua messa in onda. In questa operazione, ha abbassato la concentrazione sia sul TP, sia sul TM e ha attirato l'attenzione del collega (se fisicamente presente nella stessa cabina, come di sovente è il caso) su un errore. Tutto questo ha delle conseguenze sul rispeaker momentaneamente operativo: la momentanea assenza di supervisione del collega lo obbliga a prestare maggiore attenzione a come detta il TM, onde evitare errori che il collega non sarebbe in grado di notare; la consapevolezza di aver prodotto un errore mette il rispeaker in una situazione di stress; l'inevitabile aumento del ritardo della comparsa del sottotitolo, infine, obbliga il rispeaker ad attuare una strategia compensatoria che diminuisca il divario.

Quanto alla correzione in se, qualora possibile, essa può essere effettuata dal collega che non sta lavorando o da una persona ad essa deputata o dal rispeaker stesso. Le

ripercussioni sul rispeaker saranno di peso: poter contare sull'assistenza di un terzo permette al rispeaker di lavorare in tutta serenità, per turni consistenti e con un tasso di accuratezza accettabile; di contro, lavorare prestando attenzione alla qualità dell'*output* ed eventualmente correggere alcuni errori comporta un sovraccarico non indifferente delle operazioni intellettuali che il singolo rispeaker deve effettuare. Inoltre, è intuibile che il tasso di errori tralasciati sia superiore al caso precedente per tre motivi tra loro interconnessi:

- l'attenzione dedicata all'attività di correzione comporta un'ulteriore riduzione dell'attività di ascolto del TP e quindi l'accuratezza dell'*output* successivo alla correzione sarà intuitivamente inferiore, quanto meno in termini di contenuto;
- l'attenzione dedicata all'attività di produzione del TM è ridotta e quindi gli errori di riconoscimento saranno intuitivamente maggiori;
- l'attenzione dedicata all'attività di monitoraggio della trascrizione è ridotta e quindi eventuali errori che seguono l'errore in corso di correzione saranno intuitivamente più difficili da individuare.

Metodo di visualizzazione

Allo stato attuale dell'evoluzione della ricerca, le differenze tra i tre *software* di riconoscimento del parlato sopra menzionati sono essenzialmente di carattere tecnico e interessano l'ambito di visualizzazione dei sottotitoli sullo schermo. Più specificatamente, mentre nel caso di *Via Voice* e *Voice Suite* le singole parole, una volta che sono state elaborate dal *software*, compaiono secondo la modalità *scrolling-rolling-up* (ogni parola scorre da destra verso sinistra fino a riempire una riga, per poi salire alla riga superiore lasciando così spazio alla nuova riga in corso di formazione, cfr. figura 11); *Dragon NaturallySpeaking* proietta l'intero testo in modalità *pop-on* (ogni didascalia scompare dallo schermo sostituita dalla didascalia successiva, cfr. figura 12) solo quando riconosce nell'eloquio del sottotitolatore una pausa naturale.

Figura 11: sottotitoli *scrolling/rolling-up* fanno comparire una parola alla volta. © *RedBeeMedia*.

Figura 12: sottotitoli *pop-on* compaiono in blocco. © *JRC VOICE project*.

Questo ha delle ricadute sul lavoro del rispeaker e sull'utenza finale. Visto che *Via Voice* e *Voice Suite* riconoscono le singole parole, queste devono essere tutte accentuate omogeneamente come fossero una stringa di parole non correlate tra di loro, se si vuole ottenere un migliore riconoscimento da parte del *software*. Si tratta sicuramente di uno svantaggio per il rispeaker che non può usare la prosodia per dare coesione al testo che va pronunciando, ma deve fare affidamento esclusivamente sulla sua memoria. L'utente finale, dal canto suo, vede comparire le parole una per una avendo così l'impressione di assistere a un processo in corso, a discapito, però, della visione d'insieme. A livello grafico, quindi, non si avranno i consueti blocchi di sottotitoli, quanto un testo in continua evoluzione. Il vantaggio di questa tecnica di proiezione dei sottotitoli sta nel fatto che viene garantita una maggiore sincronia tra il testo sottotitolato e quello enunciato.

Dragon *NaturallySpeaking*, invece, proietta le stringhe di testo riunite in blocchi segmentati secondo le pause naturali prodotte dal rispeaker. Se da un lato questo ha il

vantaggio di garantire una migliore visione dei sottotitoli, la sfida per il rispeaker sta nel saper intervallare pause naturali e frasi di senso compiuto nel rispetto di quelli che sono forse i due principi base di ogni sottotitolatore di programmi pre-registrati: evitare di interrompere un sintagma a metà e garantire una certa permanenza della didascalia sullo schermo.

1.3.2 Le competenze del rispeaker

Dal paragrafo precedente sono emersi i vari fattori che influenzano il processo del rispeakeraggio tanto da determinarne un cambiamento di approccio. Eccezion fatta per gli aspetti più strettamente dipendenti dal genere da sottotitolare, però, si può dire che il rispeakeraggio ideale dovrebbe garantire una rapida trascrizione del TP, il più possibile completa, accurata e in sincronia e in armonia con il TP (Lambourne 2007). Le sfide poste da questo *optimum* sono numerose e gravano sui processi psico-cognitivi del professionista proteso al raggiungimento di un tale obiettivo.

In particolare, il rispeaker deve possedere altre competenze rispetto a quelle strettamente linguistiche necessarie al buon esito di un prodotto come la sottotitolazione pre-registrata. Come si è visto prima, la fase di produzione di sottotitoli tramite rispeakeraggio è duplice: l'uomo produce il TM e la macchina produce il TA nei tempi e nei modi dettati dal *software* di riconoscimento del parlato in uso. In quest'ultimo caso, la macchina può commettere errori nel riconoscere l'eloquio dell'operatore per motivi dipendenti da quest'ultimo o dalla macchina stessa con conseguenti errori nel TA e ritardi nella proiezione. Per evitare che questo si verifichi, il rispeaker dovrà possedere le seguenti caratteristiche che si riferiscono rispettivamente al processo traduttivo e alla forma del TM:

- Fonetiche: il rispeaker deve pronunciare le singole parole nella maniera più chiara possibile onde evitare 'malintesi' con la macchina; (Eugeni 2007)
- Psico-cognitive: il rispeaker deve simultaneamente ascoltare e comprendere il TP ed elaborare e produrre il testo di mezzo[30] nei limiti spazio-temporali dettati dalla tipologia di sottotitoli da produrre. (Eugeni 2007)

30 Cfr. Gran 1998.

Più in specifico, dal punto di vista fonetico, il rispeaker deve poter essere in grado di pronunciare ogni singola parola nella maniera più chiara possibile evitando quelle che Savino *et al.* (1999: 2) chiamano "eventi non-lessicali", cioè:

> - quelli che sono espressione di intenzionalità comunicativa (grounding, feedback, ecc). A questa categoria vengono solitamente attribuiti fenomeni quali gli allungamenti in finale di parola, le pause piene con vocalizzazione e con nasalizzazione, le nasalizzazioni e vocalizzazioni caratterizzate da particolari andamenti melodici;
> - [...] e quelli non esprimenti intenzionalità comunicative, a cui appartengono fenomeni come la tosse, lo starnuto, lo schiocco di lingua, il raschiamento, ecc. (un colpo di tosse o uno starnuto non implicano necessariamente che il parlante intenda comunicare che è raffreddato).

Benché i programmi di riconoscimento del parlato siano dotati di ausili linguistici che permettono di selezionare coppie minime in base al contesto, in alcuni casi l'omofonia può comportare un'erronea trascrizione. Sarà allora compito del rispeaker agevolare il *software*, laddove possibile, per esempio scandendo bene i confini tra le varie parole. Nel caso di 'and light', il rispeaker dovrà pronunciare separatamente le due parole di modo che il programma non le confonda con 'enlight'. Viceversa, dovrà pronunciare quest'ultimo senza pause all'interno della parola per evitare che sia riconosciuto come due parole distinte.

Dal punto di vista psico-cognitivo, il rispeaker deve avere, oltre che competenze linguistiche, anche un'ottima gestione del carico cognitivo, dovendo ascoltare il TP, ideare il TM e pronunciarlo allo stesso tempo, nel pieno rispetto dei vincoli tecnologico e linguistico imposti dal contesto comunicativo. Infine, come nel caso dell'interprete di simultanea con cui le analogie sembrano peraltro notevoli, il rispeaker, mentre lavora, deve non solo controllare il flusso della sua stessa voce, ma anche cercare di non demoralizzarsi a causa della presenza di eventuali errori presenti nei sottotitoli, risultanti non solo da imperfezioni nell'*input* vocale, ma anche dal non perfetto funzionamento del *software* stesso.

Da sottolineare è poi la necessità di alternare la produzione di testo con la produzione di metatesto. In particolare il rispeaker deve anche dettare la punteggiatura rendendo così chiaro il TA, che, sfruttando un canale che non ha a disposizione tutti gli strumenti per trasmettere appieno il senso del TP, deve scendere a compromessi con i sottotitoli (o meglio con le convenzioni della lingua scritta) per essere più facilmente ricevibile dagli spettatori.

Queste competenze sono di carattere assoluto, valgono cioè per ogni rispeaker in ogni occasione[31]. Per quanto riguarda le caratteristiche del contenuto del prodotto, il TA, invece, la scarsa letteratura in materia non fornisce strumenti utili alla redazione di linee guida per il buon esito della sottotitolazione in diretta. Sempre in ambito di ricerca, la già citata Marsh dice chiaramente che, nonostante la BBC occupi il primo posto al mondo in materia di sottotitolazione in diretta[32]; siano stati sperimentati diversi tipi di *font*; e *feedback* sia stato fornito dagli spettatori sordi nelle fasi sperimentali del rispeakeraggio, la BBC non ha condotto una ricerca sistematica sulla produzione e/o sulla ricezione dei sottotitoli forniti (2005). Tuttavia, alcune indicazioni sono proposte da ITC[33] che, nella sezione dedicata agli aspetti linguistici del documento di raccomandazioni volto a tutte le emittenti britanniche, prima introduce il concetto di "idea unit" cioè "where a proposition or key information is given" (ITC 1999) e poi suggerisce in generale di "reduce the amount of text by reducing the reading speed and removing unnecessary words and sentences; (r)epresent the whole meaning" (*ibidem*).

In particolare, il rispeaker deve assicurarsi che "subtitles should contain a reasonable percentage of the words spoken", che queste 'unità concettuali' "appear as a good percentage of the original" e che pertanto, da parte sua, "avoid 'idea units' which are unnecessary or different from the original" (*ibidem*). A fare eco a queste parole è sempre Marsh che, parlando di *editing* nel rispeakeraggio, afferma che "the hardest thing is to resist the temptation to correct the speaker's bad grammar, which is strictly forbidden" (2004: 26).

Da queste brevi ma illuminanti parole, si evince chiaramente che l'approccio del rispeaker varia anche a seconda del genere televisivo da sottotitolare. Una certa familiarità del rispeaker in questo senso sarà una discriminante del buon esito del risultato finale. Meno un rispeaker conoscerà un dato argomento, più difficile gli risulterà dare coesione e leggibilità ai sottotitoli. Questo è particolarmente vero per tutti i generi contenenti molti tecnicismi, per due ragioni fondamentali:

- il rispeaker farà molta più fatica sia nella fase di comprensione che in quella di produzione del TM rispetto al rispeakeraggio di un genere che invece conosce bene.

31 Cfr. Baaring 2006, Remael e van der Veer 2006 e Lambourne 2007.
32 Cfr. Higgs 2006.
33 ITC è una delle organizzazioni britanniche di consulenza radiotelevisiva che nel 2003 sono state sostituite da Ofcom, l'ente che attualmente supervisiona l'industria britannica delle telecomunicazioni.

Così facendo, aumenta lo sforzo che deve mettere in atto per produrre dei sottotitoli di qualità;

- il *software* potrebbe risentirne in termini di accuratezza, in quanto il processo di riconoscimento del parlato viene rallentato dalla ricerca, da parte del *software*, di un termine che non è presente nel suo vocabolario. In questo caso, verrà scelto un termine foneticamente simile a quello dettato, ma semanticamente del tutto diverso.

A proposito della familiarità del genere da sottotitolare, Marsh (2005: 28) sottolinea che:

> However well prepared a respeaker is before going on air, all manner of unexpected content can arise. If a respeaker doesn't have the necessary vocabulary trained into his or her dictionary in advance, it is impossible to use it in the subtitles. For example, if a speaker is talking about the 'Kyoto Treaty' and ViaVoice's dictionary does not contain it, it will produce something similar-sounding in its place, such as the 'key auto treaty'. A respeaker, therefore, has to find a way of communicating the message without mentioning the problematic word itself. Unfortunately, each individual respeaker has to train in each individual word into his or her dictionary – there is no way of sharing vocabulary to reduce the workload.

A questo punto, appare quindi necessario completare il quadro delle competenze professionali di un rispeaker con altre due tipologie ottenendo la tassonomia seguente:

- fonetiche: il rispeaker deve pronunciare le singole parole nella maniera più chiara possibile onde evitare 'malintesi' con la macchina;
- psico-cognitive: il rispeaker deve simultaneamente ascoltare e comprendere il TP ed elaborare e produrre il testo di mezzo nei limiti spazio-temporali dettati dalla tipologia di sottotitoli da produrre;
- diamesiche: il rispeaker deve produrre un TA scritto da un TP orale tramite la voce. Pertanto dovrà possedere competenze
  - metalinguistiche: inframezzare la produzione di testo con la dettatura della punteggiatura;

- sintetiche: nei casi di elevata velocità di eloquio del TP[34] o di necessità specifiche, il rispeaker dovrà operare una sintesi quali-quantitativa del TP in modo da garantire la leggibilità del TA nel pieno rispetto dell'aspetto multimodale in questione;

- di genere: il rispeaker deve avere una certa conoscenza del genere del programma da sottotitolare in maniera tale da evitare grossi sforzi di memoria e di essere più preciso nella resa di termini tecnici.

Resta ora da capire in che modo l'operazione di sintesi appena delineata deve essere effettuata. Dall'analisi delle migliori prassi, si potranno derivare alcune strategie che permetteranno di redigere una lista di linee guida tali da permettere alle future generazioni di rispeaker di avere delle direttive precise a cui attenersi in maniera tale da evolvere nella professione, ancora del tutto in fase embrionale.

1.4 Il rispeakeraggio in Europa

L'uso del rispeakeraggio per sottotitolare programmi televisivi in diretta o semi-diretta è cronologicamente successivo all'introduzione di sottotitoli per non-udenti nei *teletext* delle televisioni americane ed europee. Inizialmente si trattava di sottotitoli per film o altri programmi pre-registrati. In questo quadro, la britannica BBC ha da sempre svolto un ruolo pionieristico iniziando a sottotitolare programmi pre-registrati negli anni Settanta e immediatamente dopo un'edizione settimanale del TG. Hanno seguito questo *trend* i Paesi Bassi, il Belgio nederlandofono, la Germania e l'Italia negli anni Ottanta e infine la Spagna e il Portogallo all'inizio degli anni Novanta (Remael, 2007). Con l'evoluzione tecnologica, l'aumento delle richieste da parte delle associazioni in difesa degli audiolesi e la conseguente legislazione sia in ambito nazionale, sia comunitario (la prima versione della direttiva Televisione Senza Frontiere è del 1987), la sottotitolazione in tempo reale è diventata necessaria per rendere accessibili programmi importanti come i TG e altri programmi in diretta d'interesse generale. In tale contesto, l'uso della stenotipia e della 'velotipia' è stato il primo strumento per fornire il servizio di sottotitolazione in diretta. All'alba del nuovo millennio, però, per i motivi summenzionati, il rispeakeraggio è diventato lo strumento più flessibile e apprezzato dalle televisioni europee.

34 Cfr. Marsh 2005.

Nel Belgio nederlandofono, VRT ha iniziato a sottotitolare programmi in diretta già nel 1981, con un dattilografo che trascriveva un riassunto dettato da un collega, per poi passare all'uso della macchina Velotype e quindi al riconoscimento del parlato. Nel 2006 VRT ha sottotitolato tramite rispeakeraggio 20 ore settimanali di programmi in diretta e semidiretta (telecronache sportive e programmi d'informazione).

Per quanto riguarda la Germania, ARD è stato il primo *broadcaster* tedesco a sottotitolare in diretta i notiziari, con l'edizione delle 20 del celebre *Tagesschau* nel 1984. Attualmente, "…the 4 p.m., 5 p.m. and 8 p.m. news on ARD provide real-time intralingual subtitles for most bulletins in the form of closed captions on *Videotext* page 150" (Carroll *cit. in* Remael 2007: 32) tramite la tecnica del rispeakeraggio. Altre emittenti tedesche come ZDF hanno iniziato più tardi, nel 2001, ma sono in grado di fornire sottotitoli per tutti i tipi di programmi in diretta (telecronache sportive e programmi d'interesse nazionale). Stando ai dati del 2006 dell'EFHOH, ZDF ha sottotitolato in diretta 9.371 minuti di programmi nei primi tre mesi dell'anno contro i 6.433 dello stesso periodo nel 2005.

Quanto ai Paesi Bassi, invece, l'emittente di Stato *Nederlandse Openbare Omroep* fa uso sia del rispeakeraggio, sia del *velotyping*. Nel primo caso, un rispeaker produce sottotitoli in un ambiente insonorizzato e un assistente corregge eventuali errori di riconoscimento. Qualora il TP sia troppo veloce, la fase di *editing* viene soppressa a discapito della qualità del TA; nel secondo caso, un assistente riassume il TP e lo detta al 'velotipista' che lo trascrive a una velocità inferiore rispetto a quella del rispeakeraggio, ma con un'accuratezza maggiore (Lentz *cit. in* Remael 2007: 33)

Nel Regno Unito, la sottotitolazione di programmi in diretta è iniziata nel 1990, con l'impiego di resocontisti di tribunale per sottotitolare programmi come i notiziari, i *talk shows* e le telecronache sportive. Dal 2001, il rispeakeraggio rappresenta il metodo più utilizzato per sottotitolare le 650 ore mensili di programmi in diretta e semidiretta dei canali BBC1, BBC2 e BBC3 (Marsh 2006). Nel 2006, le ore sottotitolate hanno sfiorato l'80%. Forte degli ottimi risultati ottenuti dal rispeakeraggio, nel maggio del 2008, l'emittente è riuscita nell'intento di sottotitolare il 100% dei programmi trasmessi, pubblicità incluse. I rispeaker che sottotitolano per la BBC lavorano da soli, in quanto il *software* di riconoscimento del parlato da loro usato e la trasmissione di dati tramite *internet* hanno raggiunto livelli di accuratezza molto elevati. Inoltre, l'interfaccia utilizzata dai rispeaker della *RedBee Media*, *K-Live* permette di utilizzare tutti i summenzionati sistemi per il

miglioramento del riconoscimento del parlato oltre a una speciale tastiera tramite la quale si può, con una semplice operazione, spostare i sottotitoli nello schermo, cambiare colore ai sottotitoli per identificare eventuali cambiamenti di oratore, aggiungere la punteggiatura o formattare i caratteri (Marsh 2005).

In Spagna, la situazione è molto eterogenea. L'emittente di Stato TVE utilizza il rispeakeraggio, ma l'emittente catalana TVC utilizza il cosiddetto sistema semaforo per cui cinque dattilografi si alternano per tempi brevissimi alla sottotitolazione di spezzoni di testo. La successione è garantita da una specie di semaforo che indica quando il dattilografo deve prepararsi a sottotitolare (luce gialla), quando deve iniziare a sottotitolare (luce verde) e quando deve smettere (luce rossa).

In Portogallo, visto il ritardo con cui è iniziato il servizio di sottotitolazione intra-linguistica tramite *teletext*, l'accessibilità dei programmi televisivi in diretta è resa possibile dall'impiego di interpreti in lingua dei segni. Tuttavia, le emittenti sono "technically equipped to start providing live subtitling" (Neves *cit. in* Remael, 2007: 35).

In Italia, infine, la situazione è in fase molto embrionale: se da qualche anno l'accessibilità ai programmi in diretta e semidiretta (pochi TG al giorno) è garantita dall'interpretazione in lingua dei segni e dalla stenotipia, l'uso del rispeakeraggio è appena iniziato per sottotitolare il programma-contenitore di due ore su Rai tre "Cominciamo bene - estate". Entro la fine del 2009, il contratto sociale firmato con il governo, l'emittente pubblica italiana dovrà raggiungere la soglia del 60% di sottotitolazione della propria programmazione.

Allo stato attuale della ricerca, Donaldson mette in evidenza i rischi di una richiesta sempre più pressante di un aumento del numero di programmi sottotitolati dettata dai vantaggi offerti dal rispeakeraggio. In questo contesto, la qualità, già fortemente messa a repentaglio dalla non correzione di errori nella sottotitolazione in tempo reale, potrebbe ulteriormente risentirne. Facendo prova di buon senso Donaldson (*cit. in* Remael, 2007: 35) afferma che

> […] it is not enough to have a high proportion of material subtitled – such subtitles must be of an acceptable quality […]. Hearing people would not accept a soundtrack with words mispronounced, put in the wrong order, omitted or an entirely wrong word used […]. Why should the deaf and hard-of-hearing tolerate the equivalent of such errors in subtitling?

1.5 Conclusioni

Come si è visto, il rispeakeraggio è una forma di traduzione audiovisiva del tutto recente, affrontata dal punto di vista scientifico solo da pochi studiosi.[35] Eppure, si tratta di una tecnica che si sta guadagnando molto spazio nel settore della sottotitolazione, disciplina che invece gode di un maggiore interesse da parte degli studiosi europei. Le ragioni principali di questo apparente disinteresse sono da ricercare nella sua relativa novità e nel suo status: visto che si tratta di una tecnica, si tende a non considerare il rispeakeraggio come un genere traduttivo *per se*. Nel tentativo di ribaltare le sorti del rispeakeraggio, si è tentato, in questo capitolo, un approccio teorico al rispeakeraggio, che lo smarcasse dalla sua dipendenza totale dagli studi sulla sottotitolazione per non-udenti, nei quali, tuttavia, rientra per una parte della sua natura.

Per far fronte alla necessità di un quadro teorico all'interno del quale poter definire il rispeakeraggio, il ricorso a una metodologia descrittiva si è mostrato alquanto efficace. In primo luogo è stato possibile identificare il rispeakeraggio come prodotto: sottotitolazione, per non-udenti, in tempo reale. È stato così possibile isolarlo dalle altre applicazioni del riconoscimento del parlato, la tecnologia che lo caratterizza a tal punto da renderlo una forma di traduzione *sui generis*. Altrettanto immediata è risultata l'identificazione della sua macro-funzione, cioè a dire l'accessibilità ai programmi televisivi in diretta da parte di persone audiolese.

Quanto alla natura del processo e alla sua funzione, la loro individuazione non è risultata essere altrettanto semplice, in quanto il rispeakeraggio è molto complesso nelle sue sfaccettature. È stato possibile, però, indagare ugualmente gli aspetti tecnici e tecnologici del rispeakeraggio, nonché le peculiarità che lo distinguono dalle altre tecniche di sottotitolazione in diretta e di sottotitolazione per audiolesi. Infine sono state desunte le competenze che il rispeaker deve mettere in atto per poter ottenere un prodotto ottimale. Grazie a questo approccio si è potuto quindi avere una visione d'insieme e dettagliata allo stesso tempo del rispeakeraggio.

Ora che sono chiare le quattro componenti della disciplina (processo, prodotto e rispettive funzioni), è giunto il momento di costruire un quadro teorico che permetta il raggiungimento dell'obiettivo principale del presente lavoro. Partendo dai risultati

35 Cfr. Eugeni e Mack 2006.

dell'analisi descrittiva appena effettuata, sembra opportuno paragonare il processo del rispeakeraggio a quello dell'interpretazione simultanea, dal punto di vista socio-linguistico e psico-cognitivo. Questo approccio permetterà di indagare ancora più da vicino le varie sfaccettature del rispeakeraggio e fornirà spunti teorici derivanti dagli studi sull'interpretazione simultanea in caso di comprovata identità nel processo.

Successivamente, il prodotto del rispeakeraggio sarà paragonato alla sottotitolazione per non-udenti, di cui, come si è già accennato, sembra essere soltanto una tecnica di produzione. Dall'analisi contrastiva, emergerà una griglia di valutazione che sarà applicata al *corpus* di otto ore di registrazione di programmi in diretta rispeakerati dalla BBC, che come si è potuto apprezzare è *leader* in materia di rispeakeraggio televisivo. I risultati così ottenuti permetteranno, infine, di derivare delle linee guida per il buon rispeakeraggio mediante l'osservazione delle migliori prassi.

Capitolo 2 - *Shadowing* e interpretazione simultanea

2.1 Introduzione

La 'prima giornata di studi internazionale sulla sottotitolazione intra-linguistica in tempo reale', svoltasi il 17 novembre 2006 a Forlì[36], ha segnato una tappa fondamentale negli studi sul rispeakeraggio televisivo. La maggior parte degli interventi si è incentrata su questa recente tecnica di produzione di sottotitoli in tempo reale, che è stata per la prima volta analizzata sotto numerosi punti di vista: didattico, scientifico, tecnologico, professionale e sociale. Alcuni di questi interventi (Remael e van der Veer e Baaring in particolare) hanno affrontato la disciplina da un punto di vista contrastivo, paragonandola allo *shadowing* e all'interpretazione simultanea. Per amor del vero, è forse necessario chiarire che gli autori si sono concentrati sull'unico aspetto che accomuna lo *shadowing* e l'interpretazione simultanea da una parte e il rispeakeraggio dall'altra, vale a dire il processo traduttivo, cioè quello che Gottlieb (2005: 2) chiama atto del tradurre, ovvero "time, including the semantics and temporal progression of the translational process", opposto al risultato del tradurre, ovvero "space, including the semiotics and texture, or composition, of the translational product"[37]. Questa distinzione sarà utile per distinguere il ruolo dell'interprete o del rispeaker all'interno dell'atto comunicativo dall'evento in generale.

Più specificatamente, il paragone tra la produzione del TM da parte del *rispeaker* e del TA da parte dello *shadower* e dell'interprete simultaneo è stato effettuato sugli aspetti più prettamente psico-cognitivi della professione. Secondo Remael e van der Veer (2006), lo *shadowing*, l'interpretazione simultanea e il rispeakeraggio condividono esattamente le stesse fasi e in particolare:

- listening;
- understanding;
- analysing;
- and re-expressing.

Tuttavia, un elemento comune al rispeakeraggio e allo *shadowing* sembra irrimediabilmente imporsi come cruciale differenza con l'interpretazione simultanea, vale a

36 Cfr. www.respeaking.net
37 Cfr. Toury, 1995.

dire la direzionalità linguistica. Se l'interpretazione simultanea è infatti un'attività esclusivamente traducente da una lingua all'altra, il rispeakeraggio si sviluppa come tecnica concorrente alla stenotipia per la produzione di resoconti in tempi brevi e per la sottotitolazione per sordi in tempo reale. Al pari dello *shadowing*, che nasce come esercizio propedeutico per studenti di interpretazione, si tratta di un'operazione intra-linguistica, volta alla trascrizione (nel caso dello *shadowing* alla ripetizione), nella stessa lingua, di un testo orale. Secondo Treisman (1965), questo aspetto è la maggiore discrepanza tra i due processi psico-linguistici. Riferendosi esclusivamente a interpretazione simultanea e *shadowing*, Treisman imputa le differenze di resa tra un interprete e uno *shadower* a quello che definisce

> increased decision load between input and output required in translation: two selections need to be made, the first to identify the word or the phrase heard, and the second to select an appropriate response. The shadowing task is simpler if it is assured, and it is plausible, that a single central identification of the verbal unit serves for both reception and response, so that only one decision is required. (*op. cit.* in Gerver 1976)

Nonostante Treisman parli di *shadowing* come di un'operazione più semplice, Tommola *et al.* (2001) mostrano come in realtà la questione sia più complicata. In particolare, dopo aver testato il grado di accuratezza di otto interpreti finlandesi nell'espletamento di due prove di *shadowing* (in finlandese e in inglese) e di due d'interpretazione simultanea (in attiva e in passiva), gli autori parlano di migliori risultati ottenuti nella prova di *shadowing* nella propria lingua materna. Seguono, in ordine decrescente, la prova di simultanea attiva, di *shadowing* nella lingua straniera e infine di interpretazione passiva.

> The explanation for the quantitatively higher propositional accuracy score into B [language, i.e. English] is likely to be that the comprehension processes in the dominant language [i.e. Finnish] are more effective, enabling the interpreter to render more of the content, despite the fact that, qualitatively, the surface-level textual links between propositions, the lexical-syntactic formulation of output, and the fluency of delivery may not be at the same level as in SI into the dominant language […].

Un'altra sostanziale differenza sta nello *skopos*, cioè l'obiettivo dei processi traduttivi in questione. Mentre il processo traduttivo dell'interprete è volto alla ricezione da

parte del pubblico di destinazione, quello del rispeaker è, a breve termine[38], funzionale al *software* di riconoscimento del parlato. Il primo parla alle persone, il secondo alla macchina. Ecco quindi che molti aspetti divergenti che verranno sottolineati in seguito trovano una giustificazione proprio in questo aspetto. Un discorso a parte va fatto per lo *shadower*, il cui obiettivo è di svolgere un esercizio utile alla sua formazione in interpretazione simultanea. Il destinatario del suo eloquio sarà l'insegnante o lo *shadower* stesso nel caso di auto-correzione. La funzione del messaggio non è però comunicativa, come nel caso dell'interpretazione simultanea e del rispeakeraggio.

Un'ultima differenza sostanziale che emerge dalla precedente è il concetto di accessibilità. L'accessibilità ai prodotti audiovisivi, cioè la capacità di un testo di essere fruibile e utilizzabile da determinate categorie di persone, è declinabile in due grandi sottocategorie a seconda delle esigenze dell'utenza finale: accessibilità linguistica e accessibilità sensoriale[39]. Limitando il campo d'azione all'interpretazione simultanea e al rispeakeraggio, la prima è chiaramente volta all'accessibilità linguistica di quelle persone che non comprendono la lingua del TP, mentre il rispeakeraggio garantisce l'accessibilità a un testo multimodale[40] da parte dei menomati sensoriali, più esattamente i sordi. In quest'ultimo caso, la natura del prodotto finale varierà ulteriormente a seconda dei fattori di qualità considerati: nel caso si consideri più importante la fedeltà lessico-sintattica al TP, il rispeaker opterà per un processo di ripetizione *verbatim*, o quasi; nel caso si consideri preponderante la fruibilità del TA, la scelta verterà su una riformulazione sostanziale del TP.

Da queste brevi e immediate distinzioni, emerge una duplice immagine del rispeakeraggio: il rispeakeraggio *verbatim*, il cui processo, per lo scopo del presente lavoro, potrebbe essere paragonato a quello dello *shadowing*; e il rispeakeraggio *non verbatim* più simile nel processo all'interpretazione simultanea. Dopo aver fornito un'esaustiva definizione dei processi traduttivi in questione, il rispeakeraggio *verbatim* e lo *shadowing* verranno analizzati in chiave meramente contrastiva, mentre nel caso del rispeakeraggio *non verbatim* e dell'interpretazione simultanea si farà ricorso alla socio-linguistica di stampo hymesiano. Prima di entrare nei dettagli è però forse utile riassumere brevemente i risultati della ricerca scientifica concernente gli studi sull'interpretazione.

38 In seconda istanza, la trascrizione automatica effettuata dal software dovrà costituire una rappresentazione, fedele nella lingua o quanto meno nei concetti, del TP. Questo è l'obiettivo finale del rispeakeraggio.
39 Cfr. Neves, 2004.
40 Cfr. Thibault e Baldry, 2005.

2.2 Gli studi sull'interpretazione

La produzione scientifica in materia di interpretazione, intesa come tecnica traduttiva orale che implica la compresenza di più sforzi non automatici e che può essere portata a termine in diverse modalità (simultanea, consecutiva, trattativa, ecc.), è relativamente recente e risale, *grosso modo*, alla costituzione delle prime scuole di formazione per interpreti in seguito alla fioritura delle istituzioni internazionali (NATO, CECA, EURATOM, ONU, ecc.). La disciplinarizzazione degli studi sull'interpretazione si afferma soltanto nel corso degli anni Novanta. In questo periodo, si assiste al distacco degli *Interpretation Studies* dalla traduttologia, di cui fino a quel tempo l'interpretazione era considerata una variante periferica, e alla nascita di una terminologia specifica[41]. L'oggetto di studio delle varie ricerche condotte in questo ambito, infine, non è unico, ma spazia dalle scienze cognitive alla linguistica comparata, passando per le neuroscienze e la socio-linguistica. Nel tentativo di offrire un quadro il più possibile completo di tutti i contributi in materia di interpretazione, si raggrupperanno qui di seguito le varie ricerche in triplice chiave:

- in base all'oggetto di studio: il TP, l'interpretazione e il TA;
- in base alle finalità della ricerca: descrizione, evoluzione della professione, evoluzione della didattica, evoluzione della ricerca;
- in base all'approccio utilizzato (empirico, speculativo) e agli strumenti messi in atto per raggiungere gli obiettivi prefissati (materiale audio-visivo analogico e digitale, tecniche neuro- e psico- linguistiche, griglie di valutazione).

Per quanto riguarda la prima categoria, riguardante l'oggetto di studio, la sottocategorizzazione è effettuata secondo i tre principali attori del processo interpretativo: TP, interpretazione e TA. Per TP s'intende, in questo caso, non solo il messaggio prodotto dall'oratore, ma l'insieme di tutti i fattori esterni all'interprete che lo influenzano nelle summenzionate fasi di ascolto e comprensione. Molti sono i contributi dei pionieri che si sono incentrati sugli aspetti che potrebbero essere d'ostacolo alla comprensione da parte

41 La disciplinarizzazione del rispeakeraggio non è ancora avvenuta. Come lamentava infatti Remael nel corso della tavola rotonda della prima giornata di studi sulla sottotitolazione intralinguistica di Forlì (17 novembre 2006), la terminologia in uso è troppo simile a quella degli Interpretation Studies e rischia di sviare l'attenzione dei ricercatori dall'oggetto di studio.

dell'interprete e quindi influire sulla sua resa. Si dividono essenzialmente in due categorie: linguistici e non linguistici. Nella prima rientrano, tra gli altri, gli studi di Paneth (1957), Oléron e Nanpon (1965), Ilg (1959) e Lawson (1967) sulle caratteristiche precipue del TP (stilistica discorsiva, linguistica testuale, analisi di genere, ecc.) e su quelle dell'oratore (velocità d'eloquio, pronuncia, idiolettismi, ecc.). Nella seconda categoria vanno ricordati i contributi di Gerver (1974) e Barik (1971) sulle caratteristiche fisiche del canale di trasmissione del TP (rumore di sottofondo, riverbero, presenza di altri interpreti che lavorano in cabine vicine, ecc.).

Il secondo macro-oggetto di studio, l'interpretazione, va inteso come la fase di analisi operata dall'interprete prima dell'effettiva produzione del TA. Anche in questo caso sono ben distinguibili due sottocategorie che raggruppano, da una parte, i fattori esterni all'interprete e, dall'altra, quelli interni. Nella prima categoria sono da raggruppare tutti gli studi sulle tecniche interpretative e su tutto quello che rotea attorno alle singole tecniche (differenze tra consecutiva e simultanea, differenze tra simultanea e *shadowing*, implicazioni didattiche, aspetti neuro-linguistici, ecc.). Nella seconda categoria rientrano le caratteristiche del singolo interprete sia dal punto di vista linguistico (conoscenza della lingua e della cultura di partenza, grado di bilinguismo, direzionalità dell'interpretazione, conoscenze settoriali, ecc.), sia dal punto di vista tecnico (conoscenza delle strategie interpretative, resistenza allo *stress*, sistematicità nella tecnica – presa di appunti, *décalage*, ecc. –, aspetti socio-linguistici e psico-cognitivi, ecc.).

Il terzo e ultimo macro-oggetto di studio, il TA, è analizzato sia nella sua natura di oggetto della linguistica testuale (analisi degli aspetti fonetico-fonematici, morfo-sintattici, semantici e pragmatici), sia in chiave contrastiva rispetto ai fattori esterni già menzionati (caratteristiche linguistiche, velocità e comprensibilità del TP, condizioni socio-linguistiche, ecc.) e a possibili modelli di riferimento (caratteristiche fonetico-fonematiche di un normale discorso spontaneo, *standard* qualitativi della professione, ricezione del prodotto finito da parte del pubblico, ecc.).

Per quanto riguarda la seconda tipologia di ricerca in interpretazione, riguardante le finalità del ricercatore, la suddivisione segue, in qualche modo, una linea retta nell'evoluzione della disciplina. Per quanto riguarda lo scopo dei primi ricercatori, l'obiettivo era quello di determinare la disciplina in esame, cercando di definirne i meccanismi sottostanti le varie sfaccettature dell'interpretazione. Come si è già in parte

visto, i lavori di Barik, Gerver e Fabbro rientrano in questa prima categoria, quella della ricerca volta a cogliere gli aspetti identitari dell'interpretazione come disciplina *per se*. Sebbene altre ricerche siano state svolte successivamente, anche in tempi recenti, sulla comprensione del processo interpretativo, i primi ricercatori nel dominio degli studi sull'interpretazione erano particolarmente interessati a sdoganare il loro oggetto di studio dalla traduttologia, rendendolo una vera e propria disciplina accademica.

In ordine (crono-)logico, viene la ricerca volta all'individuazione degli aspetti più deboli della professione e al loro conseguente potenziamento. Vanno in questa direzione gli studi della scuola di Parigi (seguita a cadenza decennale dalle scuole di Ginevra, di Vienna e dalle SSLMIT di Trieste prima e Forlì poi), che negli anni Settanta ha svolto una lunga e fruttuosa campagna di ricerca in favore di un inquadramento scientifico e didattico delle pratiche interpretative. È stato così possibile definire le caratteristiche ottimali delle varie tecniche interpretative in vista di un'ottimizzazione della resa. Parallelamente agli istituti di formazione, anche le associazioni internazionali (fra gli altri AIIC[42] e SCIC[43]) hanno cercato, in seguito a una esigenza sempre crescente di professionalità, di dettare le linee guida dell'*optimum* in interpretazione sia dal punto di vista formale, sia contenutistico. Così facendo, le associazioni hanno anche dato il la alla costruzione di *curricula* specifici volti alla formazione di interpreti competenti e coscienti dei meccanismi alla base della loro professione. Ecco così che nasce l'esigenza di rispondere a queste aspettative. Ogni istituto superiore per la formazione di futuri interpreti sviluppa i propri corsi universitari, mirati inizialmente alla soddisfazione di mercati specifici (quelli geograficamente dominanti) e più recentemente all'acquisizione, da parte degli studenti, di competenze valide per le più svariate tipologie di mercato, in vista di una standardizzazione dei corsi in tutta l'Unione Europea.

A partire dagli anni Novanta, infine, si sviluppano i primi grandi filoni di ricerca grazie ai contributi, tra i tanti altri, di Paradis (1994), Kurz (1995), Gile (1995), Moser-Mercer (1996), De Groot e Kroll (1997) e Cowan (1995). In questo periodo sono suggerite le tendenze da seguire, avvalorati i modelli interpretativi, sviluppati e sistematizzati gli approcci cognitivi e i parametri del controllo della qualità.

42 Association Internationale des Interprètes de Conférence.
43 Service Commun d'Interprétation de Conférence.

Ora che il panorama sulla ricerca in interpretazione è stato, pur superficialmente, completato, è possibile passare all'analisi contrastiva tra i quattro processi traduttivi sopra citati, *shadowing* e rispeakeraggio *verbatim* prima e interpretazione simultanea e rispeakeraggio *non verbatim* poi.

2.3 *Shadowing* e rispeakeraggio *verbatim*

Eco (2003: 235-237) propone una tassonomia dei diversi atti interpretativi composta essenzialmente di tre macro categorie:

1) interpretazione per trascrizione;

2) interpretazione intrasistemica;

3) interpretazione intersistemica.

Mentre, sempre secondo Eco, la trascrizione è una mera sostituzione automatica, le altre due categorie si distinguono per una struttura interna più complessa. In particolare, l'interpretazione intrasistemica è da considerarsi suddivisibile in tre sottocategorie:

2.1) interpretazione intrasemiotica;

2.2) interpretazione intra-linguistica;

2.3) esecuzione.

L'interpretazione intersistemica, infine, in due, a loro volta ulteriormente ripartite al loro interno:

3.1) Interpretazione con sensibili variazioni nella sostanza

- Interpretazione intersemiotica;
- Interpretazione inter-linguistica;
- Rifacimento.

3.2) Interpretazione con mutazione di materia

- Parasinonimia;
- Adattamento o trasmutazione.

Applicando meccanicamente e superficialmente la categorizzazione di Eco, si potrebbe concludere che il rispeakeraggio *verbatim* è un semplice processo di trascrizione, una sostituzione automatica. Tuttavia, se lo si paragona allo *shadowing* sarà facile intuire

come in realtà il processo di produzione di testo mediante il riconoscimento del parlato in tempo reale sia molto più complesso di una semplice trasposizione automatizzata.

Lo *shadowing* è definibile come l'ascolto di un testo e la sua simultanea ripetizione nella stessa lingua. Secondo Lambert, lo *shadowing* è un buon esercizio per chi si cimenta nell'apprendimento della tecnica dell'interpretazione simultanea e lo definisce come "a paced, auditory tracking task which involves the immediate vocalization of auditory presented stimuli, i.e., word-for-word repetition in the same language, parrot-style, of a message" (1988: 381). Queste parole sembrano confermare che anche lo *shadowing* sia una forma di quello che Eco definisce trascrizione. Ciononostante, anche Lambert ritiene questa definizione troppo semplicistica al punto da sentire l'esigenza di contestualizzare quanto affermato citando la doppia categorizzazione dello *shadowing* proposta da Norman (1976):

- 'Phonemic shadowing': ogni suono è ripetuto senza che il significato del TP sia per forza compreso;
- 'Phrase shadowing': il TP viene ripetuto con un *décalage*[44] di una unità di senso. Contrariamente al caso precedente, il *phrase shadowing* pone come condizione indispensabile la comprensione del testo da ripetere.

Schweda Nicholson (1990) propone una terza sottocategoria dello *shadowing*, l'"adjusted lag shadowing", cioè la ripetizione del TP con un *décalage* imposto di massimo dieci parole. Anche in questo caso, Schweda Nicholson afferma che la comprensione non è una condizione essenziale. Da questa prima descrizione, due generi di *shadowing*, "phonemic" e "adjusted lag" continuano a ricadere nella categoria echiana di trascrizione, sebbene qualche dubbio emerga circa la natura prettamente automatica del processo. Il *phrase shadowing*[45], invece, sembra avere tutte le caratteristiche per rientrare nella categoria dell'interpretazione intra-linguistica, quindi del tutto assimilabile al rispeakeraggio *verbatim*.

A questo punto, sembra automatico utilizzare le parole di Lambert sullo *shadowing* anche per una generica definizione del rispeakeraggio *verbatim*. Tuttavia, malgrado l'apparente somiglianza tra i due processi, Baaring afferma che il rispeakeraggio

44 In interpretazione simultanea, il termine décalage indica l'arco di tempo che intercorre tra l'emissione del TP e la resa dell'interprete.
45 Da questo momento in poi shadowing verrà utilizzato come sinonimo di phrase shadowing.

> is not always a straightforward word-for-word repetition, parrot-style. […] The specific task requirements, including the constraints imposed by the speech recognition software, frequently force the respeaker to depart from straightforward word-for-word repetition (2006).

Per quanto riguarda un'analisi più approfondita delle due tecniche traduttive in esame, *phrase shadowing* e rispeakeraggio *non verbatim*, emerge immediatamente che lo scopo del processo differisce sensibilmente. Mentre lo *shadowing* è un esercizio propedeutico all'interpretazione, volto essenzialmente allo sdoppiamento dell'attenzione, il rispeakeraggio *verbatim* non è finalizzato alla formazione di rispeaker *non verbatim*, ma, come il rispeakeraggio *non verbatim*, alla produzione di un testo scritto per essere letto come rappresentazione del parlato.

Dal punto di vista della tecnologia necessaria all'espletamento di queste due attività, inoltre, lo *shadowing* è un esercizio svolto solitamente con cuffie e microfono per ottimizzare sia la comprensione del TP che la valutazione del TA, ma non ne è vincolato. Il rispeakeraggio invece è strettamente dipendente dalla tecnologia di riconoscimento del parlato che ne costituisce la sua ragion d'essere distinguendolo dalle altre forme di produzione di testo in tempo reale.

Dal punto di vista fonologico, lo *shadowing* richiede allo studente le stesse caratteristiche che costituiscono la qualità di un testo prodotto dall'interprete di simultanea, vale a dire una pronuncia intelligibile, ma non per forza discreta, e se possibile gradevole da ascoltare per gli eventuali futuri destinatari del TA, in vista di una situazione di interpretazione simultanea[46]. Quanto al rispeakeraggio *non verbatim*, Remael e van der Veer (2006) affermano che esso si pone al polo opposto: l'eloquio deve essere pulito e non ambiguo; ogni parola deve essere ben identificata dal *software* che, come si è visto nel capitolo precedente, non attua una vera e propria analisi morfo-sintattica del testo ricevuto in *input*, ma si basa, *grosso modo*, sulla somiglianza fonetica tra la parola emessa dal rispeaker e quella che sarà poi trascritta. Da questo punto di vista, in inglese, le difficoltà sono molteplici. Vista la sua natura fonetica di gran lunga più distante, rispetto all'italiano, dall'equazione tra fonemi e grafemi, i numerosi omofoni e gli ancor più numerosi mono-sillabi e bi-sillabi rendono di difficile decodifica il TP. Quanto all'italiano, pur presentando il vantaggio di essere una lingua fonetica, cioè con un'alta corrispondenza tra fonemi e

46 Nel caso dell'interpretazione per i film o per la TV, questo risulta essere un aspetto di capitale importanza, cfr. Mack, 2002.

grafemi, presenta lo svantaggio di avere delle sillabe molto ben distinte le une dalle altre. Qualora le sillabe di due parole 'suonassero' come un'altra parola, il *software* potrebbe fare confusione e non proiettare le parole desiderate. Un esempio di queste situazioni è il seguente:

TM: questa situazione non giova né all'una né all'altra parte;
TA: questa situazione non giovane all'una né all'altra parte.

Sempre dal punto di vista delle caratteristiche della voce, la prosodia non sempre è gradita dal *software*, che, anche in questo caso, potrebbe non riconoscere i confini tra la parole:

TM: tutti i politici
TA: tutti (…) politici

Un'altra trappola fonetica a cui il rispeaker deve fare attenzione è la pausa piena. Riempitivi come 'eeh', il respiro e gli altri eventi non-lessicali identificati da Savino *et al.*[47] possono essere riconosciuti come sillabe e quindi trascritti sotto la forma a loro più vicina:

eeh → e/è/nel/del/ecc.

Considerati questi semplici aspetti, è forse interessante sottolineare che molti rispeaker pronunciano il TM nella maniera più piatta e meno naturale possibile. Nonostante l'evoluzione della tecnologia che permette, almeno negli intenti dei programmatori, un eloquio naturale da parte di chi detta, il modo migliore per ottenere un ottimo risultato resta ancora il parlato discreto[48]. Un ultimo aspetto dipende dalla modalità in cui il *software* rilascia il TA: se il *software* rilascia il testo a blocchi di parole, il rispeaker dovrà capire il

47 Savino et al. (1999) considerano eventi non-lessicali sia quegli eventi linguistici "che sono espressione di intenzionalità comunicativa […], […] gli allungamenti in finale di parola, le pause piene con vocalizzazione e con nasalizzazione, le nasalizzazioni e vocalizzazioni caratterizzate da particolari andamenti melodici", ecc. sia "quelli non esprimenti intenzionalità comunicative, a cui appartengono fenomeni come la tosse, lo starnuto, lo schiocco di lingua, il raschiamento, ecc. (un colpo di tosse o uno starnuto non implicano necessariamente che il parlante intenda comunicare che è raffreddato)".
48 Tipologia di dettatura per cui si scandisce ogni singola parola disambiguando così eventuali omofoni.

TP e produrre frasi complete e intelligibili; nel caso di parole rilasciate una per volta invece il rispeaker non deve per forza comprendere il TP. Nella prima ipotesi si avrà un rispeakeraggio simile al *phrase shadowing*, nella seconda simile al *phonemic* o *adjusted lag shadowing*.

Per quanto riguarda l'intelligibilità del testo orale prodotto dallo *shadower* e dla rispeaker, i testi ascoltati da un orecchio umano possono contare sulla prosodia e sugli elementi extra-linguistici per disambiguare o addirittura dare senso al TA. Questo significa che il testo risultante dallo *shadowing* è comprensibile se lo *shadower* segue le pause naturali della sua lingua e ne rispetta le convenzioni tonali. Il rispeaker invece può ricorrere a un'unica strategia: dettare la punteggiatura. Vista l''innaturalità' di questa operazione, l'operatore deve continuamente fare astrazione dal TP (chiaramente privo di punteggiatura), con un continuo sforzo aggiuntivo alla sua memoria a breve termine. Una possibile conseguenza è la perdita totale (ma momentanea) di attenzione nei confronti del TP, tipica dell'interprete di simultanea che si trova a dover risolvere problemi di codifica del TA.

Per concludere, riprendendo le parole di Lambert è possibile definire il processo volto alla produzione del TM nel rispeakeraggio *verbatim* come "a paced tracking of a [spoken] text involving an immediate and phonetically accurate vocalization of auditory presented stimuli, edited when necessary for the sake of readability" (Eugeni 2008a).

2.4 Interpretazione simultanea e rispeakeraggio *non verbatim* in ottica hymesiana

Secondo la summenzionata tassonomia di Eco, l'interpretazione simultanea è un chiaro esempio di interpretazione intersistemica e inter-linguistica con variazione nella sostanza (3.1.2.), mentre il rispeakeraggio *non verbatim* varierebbe dall'interpretazione simultanea sia dal punto di vista sistemico, sia da quello linguistico. Dopo aver discusso diverse teorie sull'interpretazione simultanea, Gran (1992: 161) distingue nel processo dell'interpretazione simultanea tre fasi[49] concomitanti, che si sovrappongono, senza coincidere perfettamente:

- ascolto: l'interprete ascolta l'enunciato nella lingua di partenza;
- ideazione: l'interprete suddivide mentalmente il messaggio in unità di senso;

49 Precedentemente, si è mostrato come Remael e van der Veer suddividano il processo cognitivo alla base dell'interpretazione simultanea in quattro fasi. Anche altri autori avevano operato una simile suddivisione. Cosciente di questa variazione, Gran afferma che la differenza tra tre e quattro fasi è rilevante solo in termini teorici, non psico-cognitivi. Resta inoltre il fatto che la maggior parte dei teorici dell'interpretazione parlano di tre fasi (cfr. Russo 1999).

- produzione: l'interprete riformula ed esprime l'enunciato nella lingua di arrivo.

Da questa schematizzazione risulta evidente che la caratteristica principale dell'interpretazione simultanea è la riformulazione inter-linguistica di un testo orale di partenza prodotta in tempo reale. Rispetto al rispeakeraggio *non verbatim* scompare quindi la variazione sistemica come differenza sostanziale. L'unica che resta è il passaggio da una lingua all'altra. A colmare il divario tra i due processi traduttivi è la stessa Gran che mostra come numerosi autori suggeriscano che "scegliere un equivalente linguistico nella traduzione inter-linguistica sia psico-linguisticamente come la scelta di sinonimi o la parafrasi nella stessa lingua" (1992: 169). A questo punto, sembra chiaro che interpretazione simultanea e rispeakeraggio *non verbatim* siano, dal punto di vista psico-cognitivo, due processi assimilabili.

Tuttavia, molte sono le differenze. Come altri autori che si sono occupati di interpretazione[50], sarà qui possibile, grazie alla sociolinguistica in generale e in particolare al modello SPEAKING suggerito da Hymes (1974), definire più accuratamente le somiglianze e le differenze tra le due tecniche di trasferimento linguistico.

La teoria hymesiana

Nel suo *Foundations of Sociolinguistics: An Ethnographic Approach*, Dell Hymes (1974) studia la comunicazione da un punto di vista socio-linguistico, considerando il discorso come una serie di atti ed eventi linguistici prodotti all'interno di un contesto socio-culturale ben definito. Da questo, gli è stato possibile individuare alcune componenti comuni a ogni discorso e quindi ideare un modello di analisi valido per ogni discorso. Il nome del modello 'SPEAKING' è composto dalle iniziali delle otto macro-componenti individuate. Esse sono:

- SITUATION: è composta da *setting* e *scene*. *Setting* è "the time and place of a speech act and, in general, […] the physical circumstances" (Hymes 1974: 55); *scene* è il contesto psicologico e culturale;
- PARTICIPANTS: il mittente del TP; colui che fisicamente lo veicola; il ricevente intenzionale del messaggio; chi lo riceve in generale;

50 Per l'interpretazione di comunità cfr. Angelelli 2000; per l'interpretazione consecutiva in TV cfr. Mack 2002.

- ENDS: è composta da *purpose-goals* e *purpose-outcomes*. *Goals* sono gli obiettivi dei partecipanti e le strategie messe in atto per raggiungere lo scopo; *outcomes* sono invece i risultati ottenuti dall'evento comunicativo così come era inteso dai partecipanti;
- ACT SEQUENCES: la forma e il contenuto dell'evento comunicativo;
- KEY: il *set* di elementi che fissa "tone, manner, or spirit" (*ibidem*) dell'atto comunicativo;
- INSTRUMENTALITIES: il canale e le diverse forme e i diversi stili assunti dal discorso;
- NORMS: composta da n*orms of interactions*, che sono alla base dell'evento comunicativo e *norms of interpretation*, che costituiscono il quadro di riferimento per una corretta interpretazione dell'evento comunicativo;
- GENRES: il genere di atto o evento comunicativo strutturato secondo specifiche categorie;

2.4.1 Situation

Setting: l'interpretazione simultanea si realizza all'interno di una sala di conferenza parallelamente all'evento di cui è parte integrante: l'oratore che parla a e/o interagisce con il pubblico. Il lavoro è svolto da un interprete seduto in una cabina di simultanea che ascolta il TP e, con un ritardo di pochi secondi, ma in maniera sovrapposta, produce il TA.

L'interpretazione simultanea può anche non avere luogo nello stesso posto dell'evento principale, ma svolgersi in teleconferenza, vale a dire con una trasmissione di dati audio, via telefono o via computer in rete, fra i partecipanti all'interazione comunicativa, che possono trovarsi fisicamente in luoghi molto lontani gli uni dagli altri; o in videoconferenza, cioè con una trasmissione di dati sia audio sia video tramite computer collegati in rete. In entrambi i casi, l'interprete non condivide lo stesso spazio fisico degli altri partecipanti all'evento e il ritardo tra il momento della pronuncia del TP e la fruizione del TA da parte del partecipante alla conferenza che si trova fisicamente più lontano dal punto in cui parte il segnale è sensibilmente più ampio. La ragione è da ricercarsi nei limiti fisici della trasmissione del suono da parte di computer e soprattutto telefoni. Dal punto di vista situazionale, quest'ultima è la modalità di interpretazione simultanea che più si avvicina al rispeakeraggio in esame, quello televisivo, in quanto l'evento rispeakerato si

produce in più luoghi. In tutti i casi in cui il rispeakeraggio viene utilizzato, possono essere annoverati tra i luoghi che contribuiscono alla produzione dei sottotitoli per sordi:

- il posto da cui proviene l'*input* video (campo di tennis, emiciclo, studio televisivo, chiesa, strada, ecc.);
- il posto da cui proviene l'*input* audio (il medesimo da cui proviene l'*input* video o lo studio in cui viene effettuato il commento o in cui è registrato l'audio per poi essere sovrapposto al video);
- gli studi televisivi in cui è effettuata la regia e può essere effettuato il montaggio;
- gli uffici del *teletext* (nel caso di *rispeaker* remoto anche la postazione dove si trova fisicamente la cabina di rispeakeraggio – azienda appaltatrice, casa del *rispeaker*, ecc.) in cui si produce e il sottotitolo e lo si sovrappone al TP;
- i numerosi televisori sintonizzati sul canale in cui è proiettato il programma in questione con i relativi sottotitoli.

Anche in questo caso, il prodotto traduttivo è sincronicamente inglobato, nella sua versione grafica, nell'evento principale, ma le varie fasi di produzione e trasmissione di suoni e immagini al rispeaker, sommate al ritardo fisiologico della produzione, elaborazione e trasmissione dei sottotitoli, ritardano la ricezione del TA da parte del pubblico a casa.

In maniera del tutto simile a quanto è stato appena affermato per l'interpretazione simultanea, il rispeakeraggio come processo ha luogo in una cabina, simultaneamente, ma non del tutto sincronicamente, all'emissione del TP. Da questo punto di vista, una curiosa differenza risiede nella proprietà della cabina. Mentre gli interpreti di simultanea si muovono solitamente nel luogo dove l'evento si svolge e lavorano quindi in cabine pre-esistenti o appositamente installate, i rispeaker conoscono meglio le loro cabine, visto che nella maggior parte dei casi sono installate a casa propria o nel proprio ufficio. In questi casi, il rispeaker è avvantaggiato dal fatto di conoscere le caratteristiche fisiche della propria cabina, mentre l'interprete deve continuamente adattarsi alle caratteristiche fisiche della cabina in cui si trova a lavorare, che può essere poco insonorizzata, senza ricambio d'aria, con una scarsa visione dell'oratore e della sala o con un'apparecchiatura sconosciuta. Tuttavia, da qualche tempo si è iniziato a proporre il rispeakeraggio come forma di sottotitolazione intra-linguistica per le conferenze in cui partecipano oratori e/o pubblico sordi. Un'eventuale crescente esigenza di rispeaker *free-lance* in questo settore

comporterebbe la nascita di una tradizione di 'rispeaker di conferenza'. Da un punto di vista situazionale, questo implicherebbe un ulteriore avvicinamento della figura del rispeaker a quella dell'interprete di simultanea.

Scene: Parlando di *scene* in riferimento all'interpretazione simultanea, Angelelli (2000) sostiene che lo "speaker generally shares it [la stessa *scene*] with the listener since both belong to the same speech community. It might be not as accessible or evident to the interpreter". La stessa Angelelli riconosce comunque che c'è "little possibility to explore and discover it". Questo sarebbe perlopiù dovuto al fatto che la "situation does not always allow for clarification". Parlando di interpretazione in televisione, Mack (2002: 206) afferma che la "scene is mainly determined by the transmission genre […] and by the specific roles and statuses of participants". Vista la natura variegata del rispeakeraggio e l'indefinibilità della *scene* in un contesto talmente mutevole, lo stesso può valere per una generica definizione del rispeakeraggio. Tuttavia, va sottolineata la differenza linguistica e culturale tra coloro che normalmente 'fanno televisione' e gli spettatori sordi segnanti.

2.4.2 Participants

La conferenza con interpretazione simultanea è un evento comunicativo in cui i ruoli tra le varie categorie di partecipanti si possono scambiare continuamente. Cercando di schematizzare il quadro, Remael e van der Veer (2006) propongono una triplice categorizzazione della conferenza con servizio di interpretazione simultanea:

- 'the source text', composto dall'oratore, che in genere simboleggia sia il mittente del TP che produce personalmente (a meno che non legga qualcosa scritto da altri), sia colui che fisicamente lo veicola a quella parte di pubblico che comprende la lingua del TP;
- 'the target audience', costituito dal pubblico sia della lingua di partenza, sia di quella di arrivo e che generalmente svolge il ruolo dell'utente che intenzionalmente fruisce del TA. Nel caso di una presa di parola da parte di una persona del pubblico o semplicemente di *feedback* fornito all'oratore (risata, applausi, borbottii, smorfie di consenso o incomprensione, ecc.), il suo ruolo smette di essere del tutto passivo, invertendo così la direzionalità della comunicazione;

- 'the interpreter', che riveste sia il ruolo del ricevente del TP, talvolta non intenzionale, sia di colui che produce e veicola il TA, indipendentemente dal suo intervento sul contenuto.

Per spirito di completezza è forse utile riportare anche due altre tipologie di partecipanti all'evento comunicativo e cioè quelli che Pöchhacker (1991 *op. cit.* in Russo 1999: 94-97) definisce 'the Translations-Initiator/Bedarfsträger', cioè l'istituzione o l'organizzazione che promuove la conferenza multilingue e il 'Besteller', cioè l'organizzatore materiale della conferenza, sia esso un'agenzia di professionisti o meno.

Quanto al messaggio *per se*, nel caso tipico dell'oratore che produce e veicola il TP, questo non è percepibile soltanto dalla parte di pubblico che comprende la lingua di partenza e dall'interprete, ma anche da quella parte di pubblico che, ascoltando gli interpreti, riesce a integrare la naturale percezione delle componenti extra- e para-linguistiche del TP. Si ottiene così l'illusione da parte dei fruitori del TA di comprendere appieno il TP e di essere quindi posti allo stesso livello del pubblico che comprende la lingua di partenza. Un caso particolare è rappresentato da quella porzione di pubblico che non comprende a fondo la lingua di partenza ma che riesce a riconoscerne alcune strutture. Queste persone godono del cosiddetto effetto *background*[51]. Visto che capiscono in parte il TP, ascoltare il TA permette loro di integrare quest'ultimo con delle informazioni che potrebbero non essere state veicolate appieno o per nulla dall'interprete[52].

In nessun caso quindi l'oratore è mittente di un messaggio solo per il pubblico della lingua di partenza, ma anche per coloro che la comprendono solo in parte o che non la capiscono affatto. Un'eccezione a quanto appena affermato è la teleconferenza, durante la quale la mancanza della componente prossemica, cinetica ed extra-linguistica e la sovrapposizione pressoché totale del TA sul TP rendono praticamente impossibile l'effetto *background*. In questo caso, il TP sarà esclusivo appannaggio del pubblico della lingua di partenza.

Quanto alle interazioni tra i partecipanti, l'oratore produce il TP in maniera monologica, salvo i casi in cui vengono poste domande dal pubblico. Similmente, anche gli interpreti producono il TA in maniera monologica e senza aspettarsi *feedback* dal pubblico.

51 Esso viene definito da Gottlieb 2007 come il carico cognitivo costituito dalle informazioni ricevute dal TP indipendentemente dal TA prodotto dall'interprete.
52 Cfr. Pöchhacker, 2004.

Tuttavia, grazie alle reazioni di quest'ultimo, gli interpreti potranno avere un'idea generale della bontà o meno del proprio lavoro. In generale, essendo utenti a cui l'oratore non si rivolge direttamente, gli interpreti non svolgono affatto un ruolo passivo. In particolare, compiono uno sforzo cognitivo maggiore rispetto al pubblico a cui il TP è rivolto, in quanto devono comprendere, con solo una parte della loro attenzione[53], un testo di cui, a differenza del pubblico, potrebbero non essere esperti o interessati. Inoltre, qualora gli oratori o il pubblico non facessero niente per agevolare il lavoro dell'interprete, quest'ultimo, oltre a fare meglio che può, potrà in qualche modo uscire dal suo ruolo di intermediario. Se fisicamente presente nella sala di conferenza, potrà trovare una soluzione a situazioni che impediscono il naturale espletamento del proprio compito[54], nel rispetto di una deontologia professionale peraltro non istituzionalizzata:

- usando con moderazione il tasto 'rallenta' (se disponibile nella *consolle* di interpretazione);
- attirare tramite gesti l'attenzione dell'oratore;
- comunicare verbalmente all'oratore di rallentare;
- chiedere al pubblico di comunicare all'oratore di rallentare o di parlare più vicino al microfono.

Un caso che comporta la presenza di ulteriori partecipanti all'evento comunicativo è la conferenza con più di due lingue ufficiali. Qualora un interprete non conoscesse una delle lingue che viene parlata, sarà costretto a fare affidamento a quello che in gergo si chiama *relais*, vale a dire il collegamento a una cabina in cui gli interpreti comprendono la lingua in questione. In questo caso, i parlanti in generale e gli utenti che ricevono, senza per forza esserne i diretti interessati, uno dei testi disponibili sono raggruppabili in tre categorie: i parlanti saranno l'oratore che produce oralmente il TP, l'interprete *pivot* (cioè quello che lavora dalla lingua di partenza) e l'interprete che prende in *relais* l'interprete *pivot*; i riceventi del testo, invece, saranno l'interprete *pivot*, l'interprete che prende in *relais*

53 Cfr. Gile 1995.

54 È esperienza comune a molti professionisti incontrare un oratore che parla troppo velocemente, fuori microfono, con espressioni linguistico-culturali tipiche di una determinata area geografica, non considerando il divario temporale necessario alla ricezione del TA, ecc. Comune è anche il caso di un pubblico che si aspetta una resa dell'interprete all'altezza delle aspettative che si è costruito intorno all'oratore e che non contempla eventuali errori nel TA, imputandone la colpa all'incapacità dell'interprete.

l'interprete *pivot* e il pubblico, suddiviso a sua volta tra chi riceve il TP, chi riceve il TA dall'interprete *pivot* e chi riceve il TA dall'interprete che prende in *relais* l'interprete *pivot*.

Per quanto riguarda il rispeakeraggio, invece, i produttori del testo saranno sempre due: il parlante del TP e il rispeaker; i riceventi del messaggio saranno, a seconda dei casi, gli utenti del TP, non sottotitolato (composto da coloro che non necessitano dei sottotitoli intra-linguistici per sordi e il rispeaker), e gli utenti del TA, sottotitolato (sordi, stranieri, studiosi della sottotitolazione per sordi, ecc.).

Come si può intuire, anche in questo caso, il rispeaker si trova in una situazione simile a quella sopra descritta per l'interprete di simultanea, cioè nella duplice veste di ricevente del TP e produttore del TA, ma con una sostanziale differenza: l'impossibilità di controllare o influenzare in alcun modo il TP. Indipendentemente dalla velocità di eloquio del produttore del TP, dalla sua pronuncia e da tutte le difficoltà che possono essere incontrate dal rispeaker, questi non potrà chiedere al mittente di rallentare, ripetere o parlare più vicino al microfono.

Inoltre, secondo Baaring, il rispeaker non ha nemmeno controllo sul TA. Nonostante sia il rispeaker stesso a produrre il TM che poi passerà al vaglio automatico del riconoscitore, non bisogna dimenticare che, in realtà, il TA è il frutto del passaggio dal *software* di riconoscimento a quello di sottotitolazione. In caso di errori di resa, mentre "[t]he interpreter has limited possibilities to repairing infelicities[, t]o the respeaker that door seems to be completely closed" (Baaring 2006). Il rispeaker si troverà in realtà di fronte a un bivio di difficile soluzione. Nonostante l'ente di controllo della qualità dei sottotitoli per sordi delle emittenti britanniche, Ofcom (2003), raccomandi di "[s]end an apology caption following any serious mistake or a garbled subtitle; and, if possible, repeat the subtitle with the error corrected", la velocità di eloquio del TP impedisce nella maggior parte dei casi una simile operazione[55]. Ecco quindi che sarà necessario decidere se l'errore comporta oppure no una scorretta interpretazione dei sottotitoli. In caso affermativo, la correzione dovrà avvenire a discapito della memoria a breve termine del rispeaker o del rispetto dell'eventuale obiettivo di trascrivere il 100% del TP.

Un altro aspetto su cui soffermarsi è l'identità del produttore del TP. Come fanno infatti notare Remael e van der Veer (2006), nei casi tipici rispettivamente della

55 In realtà, esistono dei software sviluppati proprio con l'obiettivo di produrre sottotitoli tramite riconoscimento del parlato che consentono una correzione del TM prima della sua messa in onda.

sottotitolazione televisiva in diretta e dell'interpretazione di conferenza, "[f]or the respeaker, the speaker and his/her audience are not directly visible, while a conference interpreter for example has a clear view of the speaker and of his/her audience". In altre parole, mentre in una conferenza l'interprete riesce a individuare facilmente l'oratore grazie alla sua presenza fisica nell'aula di conferenza, le immagini che invece compaiono sullo schermo televisivo non sempre consentono al rispeaker di identificare il vero mittente del TP. Questo aspetto è di fondamentale importanza perché, come fanno notare gli stessi autori, similmente all'interpretazione simultanea in video- o tele- conferenza, "there is a similarity in the notion of presence, the feeling of being there, which is common to other activities where the tasks are performed in a virtual environment. If you are virtually present, do you feel that you are there or do you feel that you are not there?".[56]

Nel caso di rispeakeraggio utilizzato per la sottotitolazione di programmi sportivi in diretta, a essere inquadrato per la maggior parte del tempo è l'evento che viene commentato a discapito dei telecronisti, il cui volto raramente appare sullo schermo. Sorte simile spetta ai tele-giornalisti che sono solitamente inquadrati solo all'inizio e alla fine del servizio in diretta. Questa complementarietà tra le parole e le immagini non sempre è di aiuto al rispeaker che, a causa della rapidità con cui si succedono le immagini, tipica dei generi in questione, è costretto a non restare troppo indietro rispetto al TP, evitando così di produrre dei sottotitoli "appearing late and therefore under the head of the following speaker rather than the current one" (Remael e van der Veer 2006). Un caso interessante riportato da Eugeni (2008b) è avvenuto durante la sottotitolazione in diretta del dibattito tra Prodi e Berlusconi, il 3 aprile 2006, in vista delle elezioni per la costituzione del nuovo Parlamento italiano. Nonostante il programma non presentasse grosse difficoltà dal punto di vista dei cambi di scena

> (s)ome [...] problems could not be solved […]. When the TV presenter introduced the two candidates, shot changes were fast and some subtitles appeared under images relating to the other candidate or to the journalists. Some of the deaf viewers have remarked that and laughed after this ambiguity. (Eugeni, 2008b: 198)

L'ambiguità in questione si riferisce al momento della presentazione dei due candidati. Il moderatore ha appena finito di presentare il presidente Berlusconi e si appresta

56 Cfr. Lombard e Ditton 1999 e Mouzourakis 2003.

a presentare il leader della coalizione opposta. La telecamera si sposta su Romano Prodi, ma i sottotitoli non hanno ancora finito di riportare quanto detto su Berlusconi. Il risultato è il seguente:

Figura 13: i sottotitoli riferiti a Berlusconi, sotto l'immagine di Prodi

Quanto all'interazione tra i partecipanti all'evento comunicativo, questa non è semplicemente possibile, visto che l'autore del TP, il rispeaker e i telespettatori non condividono lo stesso ambiente e non possono nemmeno comunicare tra di loro. Tuttavia, il rispeaker può esercitare un controllo sulla produzione del TM sia in anticipo, accedendo qualora possibile al TP nel caso di rispeakeraggio in semi-diretta; sia durante, leggendo i suoi stessi sottotitoli comparire sullo schermo; sia alla fine del suo lavoro, grazie ai commenti dei telespettatori che gli uffici del *teletext* ricevono e prendono in considerazione.

Una figura raramente presente in interpretazione, ma che è diffusa nel rispeakeraggio è l'*editor*. Si tratta di un professionista incaricato di correggere eventuali errori del software o del rispeaker, prima che il TM venga trasmesso al *software* di sottotitolazione e quindi al pubblico. In altri casi, l'*editor* è un censore che detta al sottotitolatore in diretta la versione ufficiale del testo dei sottotitoli.[57]

Per concludere il quadro dei partecipanti al rispeakeraggio come processo e che influenzano il TA, il responsabile della programmazione potrebbe esigere un certo stile comunicativo rispetto a un altro, il regista potrebbe preferire alcune immagini ad altre, il

57 Cfr. De Seriis 2006.

responsabile del *teletext* potrebbe richiedere il rispetto di certe linee guida e infine i portavoce delle associazioni in difesa dei sordi potrebbero imporre alcune linee guida preferendole ad altre. In tutti questi casi, il rispeaker dovrà attenersi ad alcune variabili per il soddisfacimento delle aspettative create attorno al sottotitolo che dovrà risultare dal suo lavoro. Si dovrà quindi annoverare anche una o più delle summenzionate tra i partecipanti al processo comunicativo.

Un'ultima osservazione è forse necessaria per rendere esaustiva questa panoramica. Il rispeaker potrebbe trovarsi a operare in una conferenza multilingue. In questo caso, il rispeaker e l'interprete di simultanea interagiscono nella maniera più collaborativa possibile: l'interprete ottiene una maggiore comprensione del TP grazie ai sottotitoli intra-linguistici, che possono anche fungere da supporto alla memoria; dal canto suo, il rispeaker beneficia dell'interpretazione simultanea nel caso di un oratore di lingua straniera. Come riporta Mack (2006)

> (u)no degli esempi meno consueti di trasformazione dell'orale in scritto, accompagnato dal passaggio tra varie lingue, si è verificato proprio in occasione della […] Giornata di studi di Forlì, dove gli interventi degli oratori di lingua inglese sono stati interpretati simultaneamente in italiano, e su questa base sia interpretati in Lingua Italiana dei Segni (LIS), sia trascritti da una stenotipista e visualizzati come testo continuo di 15 righe che scorrevano mano a mano verso l'alto su uno schermo.

Un'eccezione a questa quadro è rappresentato dai rispeaker gallesi dell'agenzia che produce sottotitoli in tempo reale per la BBC, RedBee Media, o dai sottotitolatori in diretta dell'olandese NOB che producono rispettivamente sottotitoli inter-linguistici dal gallese in inglese e dall'inglese in nederlandese senza passare per un interprete.

2.4.3 Ends

Hymes (1974) distingue due tipi di scopi: "purpose-goals", cioè gli obiettivi che si sono posti i partecipanti all'evento comunicativo e le conseguenti strategie messe in atto per raggiungerle; "purpose-outcomes" sono, invece, il risultato dell'evento comunicativo nei termini previsti dai partecipanti. Anche in questo caso non è necessario distinguere tra i due visto che, in un *setting* professionale, corrispondono.

Per quanto riguarda l'interpretazione di conferenza, Mack presenta un ventaglio delle ragioni d'essere di questo servizio. Innanzitutto, l'interpretazione simultanea è "the most immediate (and often the cheapest) way of granting verbal communication between people speaking different languages" (2002: 208). Anche se una delle prime obiezioni potrebbe riguardare il costo di un tale servizio, che non sempre risulta essere economico agli occhi del finanziatore di una conferenza, è indubbio che allo stato attuale del progresso tecnologico, l'interpretazione simultanea costituisce il mezzo più rapido e flessibile per garantire il buon esito della comunicazione in un contesto multilingue. Oltre a questi aspetti, l'obiettivo di dell'interprete è anche quello di svolgere la sua attività nel rispetto della norma del "neutral mediator and honest spokesperson, loyal both to the speaker and to the listener, aiming at mutual understanding, equivalence and possibly completeness" (*ibidem*). Nel tentativo di approfondire ulteriormente quest'aspetto, l'*Allgemeine Translationstheorie* (ATT)[58] afferma che il buon esito di un'interpretazione simultanea non è tanto da ricercarsi nel rapporto tra i due testi, di partenza e di arrivo, ma proprio nel raggiungimento dello *skopos* determinato dal pubblico e dal contesto situazionale e socioculturale in cui si svolge la conferenza. In altre parole, il TA deve soddisfare appieno le aspettative dei partecipanti[59].

Quanto all'etica dell'interpretazione, è forse necessario fare ricorso al *Code of Ethics* dell'AIIC che fissa il principio per cui ogni interprete deve garantire "an optimum quality of work performed with due consideration being given to the physical and mental constraints inherent in the exercise of the profession" (AIIC, 2006). Da questo, viene ricavata una serie di *Professional Standards* "of integrity, professionalism and confidentiality" (AIIC, 2006). Il *purpose goal* dell'interprete, e conseguentemente anche il *purpose outcome*, sarà quindi "to adhere to these criteria or to any other set of judicious criteria for the success of the communicative event in which s/he is taking part" (Eugeni 2008a: 369).

Per concludere, altri obiettivi secondari considerati da Mack sono "to earn one's living" e "to preserve one's face in front of an invisible but large audience, not to mention critical clients and colleagues" (2002: 208).

Quanto al rispeakeraggio, il quadro non è ben definito in quanto, grazie agli insegnamenti della *Allgemeine Translationtheorie*, è possibile comprendere come la

58 L'ATT è una teoria funzionalista che concilia la Skopostheorie e la Theorie über translatorisches Handeln. Cfr. Pöchhacker 2004.
59 Cfr. Mack 2002.

tipologia del lavoro da svolgere vari a seconda del genere da sottotitolare e soprattutto del tipo di pubblico per cui si lavora. Queste due variabili determineranno le caratteristiche tecniche del TA. C'è da sottolineare, però, che a breve termine il compito del rispeaker non è di produrre un testo a immediato uso del pubblico di destinazione, ma di produrre un *output* che si adatti alle esigenze del software di riconoscimento del parlato in uso, permettendogli così di riconoscere le parole pronunciate e di trascriverle correttamente. Solo dopo la fase di riconoscimento e trascrizione, auspicabilmente esatti, il testo potrà essere usufruito dall'utenza finale. Nonostante le due fasi siano cronologicamente separate in maniera molto distinta, non sembra teoricamente valido accettare una suddivisione così netta tra i due passaggi. In altre parole, non si può sostenere che l'obiettivo del rispeaker sia di parlare al *software* e quello del *software* di parlare agli utenti finali. Una soluzione di sintesi accettabile potrebbe essere invece la seguente: il rispeaker parla agli utenti finali attraverso il *software*.

Alla luce di questa che potrebbe sembrare un aspetto secondario, ma che rischierebbe di confondere le idee in un quadro teorico così affollato di passaggi psico-cognitivi, è finalmente possibile affermare che, nel caso del rispeakeraggio *non verbatim*, il *purpose-goal* non sarà di riprodurre il massimo numero di parole del testo originale possibile (come avviene per il rispeakeraggio *verbatim*[60]), disattendendo così la richiesta della maggior parte delle associazioni in difesa degli audiolesi (ma soddisfacendo le loro necessità). Bensì sarà veicolare il messaggio a un pubblico con esigenze linguistiche specifiche. Secondo i risultati della ricerca in esame al capitolo 5, "a target text mirroring the grammar of LIS, if it respects the Italian grammatical rules, is the best way to satisfy the needs of the target audience" (Eugeni, 2008a: 370). Ne consegue che, in questi casi, la maggiore preoccupazione di un rispeaker sarà quella di rendere il TP nella maniera più comprensibile possibile.

2.4.4 Act sequences

Gli *act sequences* sono la forma e il contenuto del messaggio. A tal proposito, uUna sostanziale differenza con l'interpretazione è da ricercarsi nella co-occorrenza dei testi di partenza e di arrivo nel supporto audiovisivo. L'effetto *background* sugli utenti sarà ancora maggiore che in interpretazione, in primo luogo perché se i due testi sono nella stessa

60 Cfr. Marsh 2005.

lingua, ogni utente che sia in grado sia di sentire abbastanza bene, sia di leggere il testo scritto noterà inevitabilmente ogni sostanziale differenza tra i due testi. Inoltre, stranieri che fruissero dei sottotitoli intra-linguistici per compensare la loro insufficiente comprensione della lingua orale, avranno la possibilità di guadagnare maggiore controllo sulla lingua di partenza. Per completare il quadro determinato dall'effetto *background* è forse interessante notare che anche i diretti interessati del rispeakeraggio potranno utilizzare questa co-occorrenza dei due testi per rafforzare le loro competenze nella labiolettura con i sottotitoli intra-linguistici.

Tuttavia secondo Kalina (1992), la differenza tra una situazione comunicativa monolingue e una mediata tramite interpretazione simultanea sta, oltre che nella summenzionata contemporanea presenza del TP e del TA, nella mancanza di autonomia semantica da parte dell'interprete, ossia l'impossibilità di quest'ultimo di cambiare contenuto o registro nella fase di produzione del TA. In interpretazione simultanea infatti sia la forma (aspetto traduttivo escluso), sia il contenuto sono stabiliti dal produttore del TP. Tuttavia, visto che è possibile affermare senza essere smentiti che non esistono lingue orali isomorfe, l'interprete è chiamato a intervenire nella forma in maniera tale da poter veicolare quello che Halliday e Hasan (1985) chiamano 'mode of discourse', vale a dire il ruolo che i partecipanti all'evento comunicativo si aspettano che la lingua svolga nella situazione comunicativa in cui si trovano, l'organizzazione simbolica del testo, lo status del testo, la sua funzione nel contesto in cui viene prodotto e il cosiddetto 'rhetorical mode'. Ciò significa che la forma del testo può avere sia una rilevanza estremamente importante (discorso politico o retorico in genere), sia più marginale (discorso tecnico). In genere, però, la semplificazione deliberata del TP non è in alcun modo accettabile.

D'altro canto, un rispeaker, come sintetizza Baaring (2006), "must be able […] to simultaneously pay attention to the form and content of the source message and to the form and content of the target communication". La forma e il contenuto acquistano quindi una valenza particolarmente complessa, perché se l'obiettivo finale del rispeakeraggio è di andare incontro alle esigenze degli utenti finali, il rispeaker che produce sottotitoli *verbatim* riuscirà a svolgere appieno il suo compito solo aderendo alla forma e al contenuto del TP. Delle difficoltà potrebbero sorgere nella resa di discorsi altamente retorici, in cui numerosi fattori contribuiscono alla forza illocutoria del messaggio. Per rendere nella sottotitolazione questi aspetti sarà necessaria un'intelligente trasposizione di tutte le componenti para- ed

extra-linguistiche. Nel caso di testi troppo veloci, inoltre, il rispeaker dovrà prestare attenzione, oltre che alla sempre presente necessità di dettare correttamente la punteggiatura, anche alla delicata strategia della compressione sintattica, che diventa indispensabile e che va ad accomunare il rispeakeraggio *verbatim* a quello *non verbatim* e all'interpretazione simultanea.

Nel caso del rispeaker chiamato a produrre sottotitoli *non verbatim*, alla responsabilità dell'equivalenza e dell'accuratezza del TP si aggiungono anche quelle di adeguatezza e fruibilità del TA, indispensabili aspetti che definiscono la qualità in interpretazione simultanea[61]. In altre parole, seppur riformulato e quindi ridotto nella forma, il TA deve essere fruibile sia linguisticamente, sia culturalmente dal pubblico a cui è destinato. Chiaramente, questo non implica uno snaturamento della forza illocutoria. Così come nel caso di sottotitoli *verbatim*, sarà necessario anche in questo caso attuare delle strategie per rendere il più possibile la componente pragmatica del TP tramite la summenzionata strategia della resa delle componenti para- ed extra-linguistiche. In ordine di importanza, seguono il rispetto della componente semantica, di quella lessicale e infine di quella sintattica.

Per concludere, è forse opportuno nominare un altro aspetto importante che diversifica il rispeakeraggio in genere dall'interpretazione simultanea e che risiede nella sua natura diamesica. Come è stato già detto, se interpretazione simultanea e rispeakeraggio *non verbatim* si assomigliano per la co-occorrenza del TP e del TA, i due prodotti sono uno destinato direttamente all'ascolto e l'altro al software di riconoscimento del parlato e in seconda battuta alla lettura. Risulta quindi chiara, nel rispeakeraggio in generale, oltre l'importanza della punteggiatura e della riformulazione, anche quella della correttezza fonetica del TM e conseguentemente la realizzazione grafemica del TA. Mentre in interpretazione una parola non pronunciata correttamente può essere comunque compresa dal pubblico o comunque immediatamente corretta dall'interprete, una parola pronunciata male dal rispeaker può comportare uno scorretto riconoscimento da parte del *software*. Questo risulterà in una parola diversa da quella pensata che andrà a impattare negativamente sulla *legibility* del TA. Nel caso di proiezione *pop-on* del TA, sempre che ci si renda conto immediatamente dell'eventuale errore e che quindi la correzione avvenga immediatamente

61 Cfr. Viezzi 1999.

dopo la parola pronunciata male, il sottotitolo risulterebbe comunque incomprensibile come nell'esempio seguente:

TM: Il ragazzo aveva un auricolare rosso
TA: Il ragazzo aveva un auricolare rotto rosso

In questo caso, l'oratrice sta difendendo il metodo oralista nella comunicazione dei bambini sordi e in particolare sta spiegando che un amico sordo di suo figlio aveva un auricolare rosso di cui andava fiero perché tifoso della Ferrari, valore aggiunto questo che si sommava agli altri vantaggi dell'auricolare, tra cui il suo funzionamento rispondente alle esigenze del ragazzino sordo. L'errore di riconoscimento non risulta immediatamente evidente, sembra anzi essere un'informazione in più che, oltre a non essere veritiera, potrebbe inficiare non solo la corretta comprensione del passaggio in questione, ma anche l'interpretazione di tutto il testo. Nell'altro caso qui riportato, l'errore di riconoscimento è invece meno grave anche se un po' bizzarro:

TM: i due ospiti avranno due minuti e mezzo per un ultimo appello
TA: i due ospiti avranno 2 min e mezzo per un ultimo bello appello

In questo esempio l'informazione aggiunta non compromette la natura del testo sebbene alteri il registro del testo[62]. Nel caso di errore in generale, come si è già visto precedentemente, Ofcom suggerisce di notificare l'avvenuto errore tramite un sottotitolo *ad hoc* e di correggere il sottotitolo errato. Questo implica un'espansione, in termini quantitativi, del TM rispetto al TP, ma almeno ne riporta correttamente il significato.

2.4.5 Key

Key è l'insieme di tutti quegli elementi che definiscono il "tone, manner, or spirit" (Hymes, 1974: 57) dell'atto comunicativo. Per raggiungere questo obiettivo, l'interprete simultaneo "will focus on the tone, manner or spirit of the speaker" (Angelelli, 2000). Comunque, ricostruire la stessa *key* creata dalla lingua di partenza non è sempre possibile. Per esperienza personale è possibile affermare senza temere di essere smentiti che molti

62 Cfr. Pirelli 2006.

professionisti si rifiutano di 'recitare' in cabina preferendo a una resa equivalente un TA più 'neutro'. In altre occasioni è proprio impossibile rendere la stessa *key*. Tuttavia, l'oratore principale è generalmente fisicamente presente nello stesso *setting* comunicativo e alcuni aspetti che l'interprete non riesce a rendere vengono comunque trasmessi all'utente finale tramite la prossemica o i tratti sopra-segmentali del TP. Infine, il TP rimane presente in sottofondo e quindi "contributes to the sense of authenticity in the translation and prevents a degree of mistrust from developing" (Luyken *et al.* 1991: 80 *op. cit.* in Mack 2002: 209).

Anche nel rispeakeraggio, il sottotitolatore 'will focus on the tone, manner, or spirit of the speaker', tuttavia, la semplice variazione diamesica impedirà il buon esito di una resa equivalente della *key* senza un'esplicitazione delle componenti prossemiche ed extra- e para-linguistiche nel TA. Per quanto riguarda l'effetto 'background', è interessante sottolineare che, mentre in interpretazione di conferenza un sottofondo ben udibile crea il senso di autenticità di cui parla Luyken, nel rispeakeraggio questo non è sempre vero: se da un lato coloro che hanno un residuo di udito riescono a compensare le loro carenze sensoriali con un sottotitolo fedele ed eventualmente con la labiolettura, dall'altro questo non è possibile se il sottotitolo presenta un sostanziale cambiamento nella forma. In particolare, nel caso del rispeakeraggio *non verbatim*, la lettura labiale non può andare di pari passo con la lettura del sottotitolo perché i due testi (di partenza e di arrivo) non solo non sono sincronizzati, ma divergono nella forma. Inoltre, la sola lettura labiale non permette comunque una cognizione totale del TP in quanto troppo complessa se protratta nel tempo e se il TP è particolarmente veloce. Nel caso gli spettatori siano sordi profondi, infine, la situazione è aggravata dalla mancanza totale di una percezione acustica del TP. L'effetto *background* risulta quindi inutile se non fuorviante. Alcuni sordi potrebbero infatti risentirsi dalla mancata corrispondenza totale tra il TP e il TA. Come già dimostrato infatti il senso di essere vittima di paternalismo è forte e notare una tale differenza potrebbe essere interpretata come un'ingiustizia.

2.4.6 Instrumentalities

Instrumentalities è l'insieme costituito dal canale e dalle diverse forme e stili del discorso. Ai fini del presente lavoro, l'espressione *instrumentalities* viene quindi utilizzata per comprendere sia le varie forme assunte dal discorso (per es.: monologo/dialogo), sia l'insieme degli elementi che ne permettono la produzione (ad es.: orale/scritto), la

trasmissione (ad es.: carta/radio/TV) e infine la ricezione (ad es.: grafico-visivo/fonico-acustico/fonografico-audiovisivo). In interpretazione simultanea, la comunicazione può avvenire sia in maniera monologica che dialogica. Nel primo caso, l'oratore leggerà un testo scritto per essere letto o parlerà seguendo o meno una scaletta precedentemente preparata. Inoltre, potrà far uso di diapositive o di presentazioni animate. Altre componenti comunicative coinvolgeranno il linguaggio del corpo, eventuali segnali subliminali o i tratti soprasegmentali. Infine, il discorso monologico è raramente interrotto e l'oratore mantiene la stessa forma per tutta la durata dell'atto comunicativo. Nel secondo caso, invece, la forma del discorso varierà a seconda dell'idioletto e della competenza linguistica e/o in materia degli oratori, ognuno dei quali potrebbe improvvisarsi tale previo stimolo a intervenire. Il testo prodotto tenderà quindi più all'oralità, con un conseguente abbassamento di registro rispetto al primo caso. Indipendentemente da tutto questo, l'unico canale di trasmissione che l'interprete può utilizzare è quello orale/acustico e la forma sarà il più rispettosa possibile dell'originale e comunque corrispondente in termini di registro, varietà e altro ai "different registers, varieties, etc. used by the speaker" (Angelelli 2000).

Nel rispeakeraggio, la distinzione tra processo e prodotto risulta essere determinante: sempre riferendoci alla tassonomia di Gottlieb, come processo, il rispeakeraggio è una traduzione isosemiotica, che usa quindi "exactly the same semiotic channels as the original" (2005: 4). In particolare, come fanno notare Remael e van der Veer (2006) "the speech is meant to be written and the written speech, in the end, is meant to be read by the viewers at home". C'è quindi un'ibridizzazione del canale e del registro sia del testo prodotto dal rispeaker, sia dei sottitotoli che il pubblico legge. In altre parole, il rispeaker ascolta il TP, che è prodotto oralmente, e produce oralmente il TM. Il risultato finale sarà un testo scritto, o meglio la trascrizione del TM che rispecchia auspicabilmente in maniera esatta i diversi registro, varietà, ecc. del rispeaker.

È proprio per questo motivo che il rispeakeraggio, come prodotto per sordi cofotici, dovrebbe essere considerato come una forma di traduzione diasemiotica, caratterizzata "by its use of different channels, while the number of channels (one or more) is the same" (Gottlieb 2005: 4). I sottotitoli sostituiscono infatti completamente la componente acustica del TP. Come prodotto per persone con un residuo di udito, invece, i sottotitoli per sordi diventano una forma di traduzione supersemiotica, cioè a dire "the translated texts display more semiotic channels than the original" (*ibidem*), in quanto vanno ad aggiungersi, seppur

in maniera lieve, alle componenti acustiche del TP, come nel caso dei sottotitoli per normoudenti. È forse necessario notare, però, che la forma del TA non può rappresentare soltanto la componente audio verbale, visto che il rispeaker produce un testo che sa che deve assolvere a funzioni specifiche in relazione con altre componenti di significato.

Per quanto riguarda la struttura del discorso, infine, il rispeaker produrrà un testo equivalente a livello di unità di significato, ma occasionalmente diverso nella struttura sintattica e semantica, in modo che l'equivalenza non vada a discapito della fruibilità.

2.4.7 Norms

Sebbene, il concetto di norme si sia evoluto nel corso degli ultimi decenni[63], la definizione data da Hymes sembra essere ancora pertinente ai fini dell'analisi in corso. Hymes distingue tra *norms of interaction*, che sono alla base dell'evento comunicativo, e *norms of interpretation*, che stabiliscono un quadro di riferimento per una corretta interpretazione dell'evento comunicativo.

Norms of interaction

Per quanto riguarda l'interpretazione, le 'norme di interazione' sono "highly ritualized" (Mack, 2002: 210). In questo contesto, un interprete di simultanea non ha quasi nessun controllo sul TP, ma ha in teoria un controllo assoluto sul TA. In altre parole, l'interprete è colui che materialmente veicola il TA ed è, eccezion fatta per la prossemica, i tratti soprasegmentali e i supporti multimodali, l'unico garante della comunicazione tra l'oratore e il pubblico della lingua di arrivo. Tuttavia, non può decidere né la presa di parola, né altre norme di interazione. In particolare, per quanto riguarda la produzione del TA, l'interprete non viene considerato come il produttore del testo, ma come una 'seconda voce' dell'oratore principale, la cui presenza è nota, ma la cui influenza sul testo che ricevono viene eclissata dal già menzionato effetto *background*. Quanto allo svolgersi della conferenza, la sua influenza può essere percepita in caso di difficoltà più o meno evidenti nel processo traduttivo.

Se nelle conferenze l'interazione è ritualizzata e le conseguenze della presenza dell'interprete sono implicitamente accettate da tutti i partecipanti, i testi audiovisivi non permettono un'interazione tra l'oratore principale e il rispeaker da una parte e i telespettatori

63 Cfr. Toury 1986 e Hermans 1999.

dall'altra o anche soltanto tra il rispeaker e i telespettatori. L'unica interazione possibile avviene tra l'oratore principale e i primi ricevitori del TP (in un quiz, tra il conduttore, i concorrenti e il pubblico in studio, al TG tra l'inviato e il giornalista in studio, ecc.). Un aspetto degno di nota riguarda la natura dell'intervento del rispeaker sul TA. Se il produttore di sottotitoli *non verbatim* si concede la libertà della riformulazione in nome del rispetto delle unità concettuali espresse nel TP[64]), il rispeaker *verbatim* non può nemmeno intervenire per migliorare il testo che sta sottotitolando. A tal proposito, Marsh (2004) dice chiaramente che "the hardest thing is to resist the temptation to correct the speaker's bad grammar, which is strictly forbidden".

Norms of interpretation

In materia di norme interpretative, malgrado gli studi in tale campo siano numerosi, non esiste una teoria che vada per la maggiore su come interpretare il TP. Ciò nondimeno, la deverbalizzazione, cioè "the ability of the translator/interpreter to perceive the meaning […] of a text/an utterance in its proper context and thus convey its underlying message as distinct from mere transcoding […], or word-for-word translation" (Mouzourakis 2005), è l'oggetto di due delle teorie sull'interpretazione più quotate, la *Théorie du sens* e la *Skopostheorie*. Nonostante siano molto elastiche e facilmente applicabili a molti contesti, un abuso di queste norme porterebbe a una resa forse troppo *target oriented* del testo, snaturandolo inevitabilmente. A tal proposito, parlando di traduzione inter-linguistica, Toury (1995) asserisce che

> whereas mainstream *Skopos*-theorists see the ultimate justification of their frame of reference in the more 'realistic' way it can deal with problems of an *applied* nature, the main object being to 'improve' (i.e. change!) the world of our experience, my own endeavours have been geared primarily towards the *description* and *explanation* of whatever has been regarded as translational within particular target cultures, the ultimate object being to formulate a series of interconnected laws of a probabilistic nature, along with their conditioning factors.

Un altro aspetto normativo concernente l'interpretazione simultanea, in particolare il processo interpretativo, è il discorso attorno alla qualità del TA. Se "concepts such as accuracy, clarity, or fidelity are invariably deemed essential" (Pöchhacker 2002: 96), in

64 Cfr. Ofcom 1999.

realtà non esiste un consenso altrettanto unanime sul significato di questi termini. Cerca di sopperire a questa mancanza Viezzi (1999: 146-151) che definisce quattro obiettivi e parametri volti a dare un quadro generale che definisca la qualità in interpretazione simultanea:

- *equivalenza*: secondo lo stesso Viezzi "(i)l concetto di equivalenza è probabilmente il concetto più discusso e contestato nel campo degli studi sulla traduzione e sull'interpretazione" e che addirittura "viene rifiutato da valenti studiosi" (*ibidem*). Si tratta essenzialmente dell'identità di valore tra il TP e il TA, in termini di funzione comunicativa, valenza socio-comunicativa, significato lessico-grammaticale ed effetto perlocutorio;
- *accuratezza*: è "la trasmissione del contenuto informativo di un testo, o meglio, delle singole informazioni contenute nel TP" (*ibidem*);
- *adeguatezza*: sia culturale (che permette una comunicazione interculturale tra l'oratore principale e il pubblico a cui si rivolge), sia linguistica (il TA deve soddisfare le esigenze linguistiche del pubblico in termini di registro e di genere;
- *fruibilità*: il TA deve essere immediatamente comprensibile "in modo da facilitarne la ricezione e l'elaborazione" (*ibidem*).

Per quanto riguarda il rispeakeraggio, niente di tutto questo è disponibile. Non esiste una tradizione di studi sull'interpretazione che si è evoluta in questo senso. È per questo motivo che ispirarsi a modelli scientifici già esistenti potrebbe essere una soluzione per colmare il divario. Prima di entrare nei dettagli però, è forse necessario richiamare brevemente la differenza tra la produzione di sottotitoli *verbatim* e sottotitoli *non verbatim*. La prima forma di rispeakeraggio richiede poche ma chiare e ben definite regole sull'uso della punteggiatura e della riduzione del TP. Tuttavia, nonostante il rispeakeraggio *verbatim* sia la soluzione richiesta da molte associazioni di sordi per qualsiasi contesto (TV, conferenze, lezioni, ecc.) e quella offerta dalla maggior parte di produttori di sottotitoli in tempo reale, pochi sono gli studi in materia per definire il concetto di qualità. Per quanto riguarda il Regno Unito, l'agenzia che offre sottotitoli in tempo reale per la BBC si pone l'obiettivo della trascrizione fedele di quello che viene detto nel testo audiovisivo con il

limite auto-imposto(si) di 300 parole al minuto[65]. Dopo questo limite si può procedere alla compressione. Paradossalmente, però, "the news and the parliamentary sessions, being particularly fast, you have to go along with them. While, if you subtitle sport, the idea is that you describe the action you can see on the screen so you do not need to speak all the time" (Marsh 2005). La natura della compressione dipende quindi, più che dalla velocità del TP, dal genere del testo da sottotitolare.

In materia di sottotitolazione *non verbatim*, oltre a trarre profitto dagli studi sulla sottotitolazione per sordi di programmi pre-registrati[66], il rispeakeraggio è stato recentemente l'oggetto di alcune ricerche presso le sedi provinciali dell'ENS (Ente Nazionale Sordi) della regione Emilia Romagna[67]. Da questi studi è chiaramente emersa innanzitutto l'inadeguatezza di una sottotitolazione *verbatim* dei notiziari e la conseguente necessità di una riformulazione del TP. Nello specifico, è risultato evidente che la riformulazione rende molto più comprensibile il testo[68] rispetto alla semplice trascrizione. In particolare, la riformulazione sintattica è quella che, forse più di tutti gli altri tipi, permette una fruizione più agevole del TP senza snaturarlo delle sue precipuità linguistiche (idiomatismi, ricchezza lessicale, precisione concettuale). Da questi risultati è emersa una serie di linee guida[69], complementari a quelle offerte da Ofcom, a cui la BBC si ispira e tenta di aderire. In particolare, Ofcom richiede oltre a una resa quantitativa e qualitativa delle unità concettuali ('idea units'), anche l'eliminazione di "idea units which are unnecessary" (Ofcom 2003). C'è da sottolineare, però, che mentre le linee guida di cui sopra si sono dimostrate utili alla produzione di sottotitoli di alta qualità, in quanto compresi dalla maggior parte degli utenti segnanti, nella pratica, la grande varietà linguistica e cognitiva dei possibili utenti dei sottotitoli intra-linguistici rende pressoché impossibile un intervento massiccio del rispeaker sul TP.

65 In realtà, come si vedrà successivamente, il limite reale sembra essere 180 parole al minuto. 300 parole al minuto è da considerarsi come il picco massimo raggiungibile in un arco ristretto di pochi secondi.
66 Per l'Italia, cfr. Volterra 1986. Per il Regno Unito, cfr. Ofcom 1999. Per il Portogallo in particolare e per avere una panoramica della sottotitolazione per sordi di programmi pre-registrati, cfr. Neves 2005.
67 Cfr. Eugeni 2007.
68 In seguito a una riformulazione sintattica del TP, il 61.42% del campione ha risposto positivamente alla maggior parte delle domande incluse in un test di comprensione contro il 13.7% di una riformulazione lessicale e il 6.6% della trascrizione. Il dato che si riferisce alla riformulazione semantica (78.17%) non è identificativo di una maggiore comprensione di una riformulazione semantica rispetto a una sintattica, ma di una maggiore comprensione di una riformulazione semantica e sintattica rispetto alla sola riformulazione sintattica.
69 Cfr. Eugeni 2007, in cui viene proposto un decalogo, che si è dimostrato efficace, e che propone, tra l'altro il rispetto dell'ordine sintattico di base, la coordinazione, la disambiguazione e l'esplicitazione dei tropi.

2.4.8 Genres

Il genere è il tipo di atto o evento comunicativo a cui appartiene il TP e le diverse categorie in cui esso è strutturato. Secondo Bhatia (2002: 4), il genere è la serie di "instances of conventionalised or institutionalised textual artefacts in the context of specific institutional and disciplinary practices, procedures and cultures" costruite, interpretate e utilizzate dai membri di una *discourse community* con l'obiettivo di "achieve their community goals" (*ibidem*). Grazie a questa definizione onnicomprensiva della categoria genere è possibile definire l'interpretazione come un artefatto testuale che, come si è visto, è altamente istituzionalizzato, in quanto riformulazione inter-linguistica e in tempo reale di un TP in un TA, all'interno di un contesto che segue procedure istituzionali e disciplinari specifiche. Per definire tale concetto, Russo (1999: 92) spiega che una conferenza, luogo corrispondente al *setting* in cui si svolge l'interpretazione simultanea, può essere strutturata in vari modi, a seconda della tipologia a cui appartiene:

- la classica assemblea parlamentare comporta il susseguirsi di formalità d'apertura, presentazione di rapporti, discussione, votazione e, di conseguenza, i tipi di testi probabili saranno: discorso cordiale d'inizio, relazioni scritte che verranno invariabilmente lette (di norma, assai velocemente), domande o interpellanze, ecc.;
- il convegno specialistico di medicina prevede una sessione d'apertura con lettura magistrale e sessione di chiusura inframmezzate da sessioni di lavoro dove gli unici tipi di testo presentati sono relazioni scientifiche (articolate in: introduzione, materiali e metodi, risultati, discussione e conclusioni), accompagnate da diapositive e seguite da un breve (di solito) scambio di domande e risposte;
- la tipologia "*forum* politico" prevede gli interventi dei relatori invitati e il dibattito; testi tipici saranno discorsi a braccio, magari propagandistici, sarcastici o ironici e ricchi di riferimenti a fatti e personaggi del Paese in questione.

A questa categorizzazione di contesti si aggiunge la video-/tele-conferenza, l'interpretazione per i festival del cinema e in particolare per la TV. L'interpretazione simultanea, infine, 'è costruita, interpretata e utilizzata' da tutti i partecipanti (interpreti e membri delle singole *speech community* che compongono il consesso – rispettivamente parlamentari; autorità politiche e/o accademiche, scienziati, ricercatori, studenti, professionisti, esperti, appassionati, curiosi, ecc.; e politici, rappresentanti di spicco della

società, elettori, ecc.; gli stessi ma in video-/tele-conferenza; esperti di cinema, registi, attori, giornalisti, spettatori, ecc.; tutta la gamma di oratori presente nella TV e i telespettatori).

Per concludere il quadro, è forse opportuno citare anche la categorizzazione fatta da Pöchhacker (1994 *op. cit.* in Russo 1999) che distingue la conferenza, più che in generi, in vari tipi di ipertesto:

- *Versammlung einer internationalen Organisation*, ovvero l'assemblea di un'organizzazione internazionale, paragonabile alla prima tipologia proposta da Russo;
- *Fachkonferenz*, cioè la conferenza specialistica, affine alla seconda tipologia di Russo, ma più generale;
- *Seminar und Schulung*, che abbraccia il campo più vasto dei seminari e delle lezioni universitarie frontali;
- *Verhandlung*, ossia il dibattito nelle sue forme più varie, anche all'interno di un altro ipertesto;
- *Aktuelles Forum*, simile al forum di Russo ma allarga il campo anche ad altre tipologie di *forum*;
- *Pressekonferenz und Präsentation*, letteralmente la conferenza stampa e ogni tipo di contesto situazionale in cui uno o più oratori prendono la parola di fronte a un pubblico più o meno omogeneo per presentare qualcosa o qualcuno (un libro, un giocatore appena acquistato, un film, ecc.);
- *Gastvortrag,* simile alla *Schulung* nella forma, ma assai diversa nello *skopos*, in quanto il rapporto tra oratore e pubblico non si iscrive in un contesto specifico (reciproca conoscenza, circolarità della lezione, attinenza a un programma predefinito, ecc.), ma somiglia a un convegno specialistico.

Per quanto riguarda il rispeakeraggio, soprattutto il rispeakeraggio *non verbatim*, la sua caratteristica principale, forse addirittura la sua *raison d'être*, è la flessibilità che lo contraddistingue nel produrre sottotitoli accettabili in maniera rapida, dinamica e soprattutto economica. Proprio per quest'ultima ragione, il rispeakeraggio è utilizzato in particolare dalle emittenti televisive per sottotitolare sia programmi in diretta (sedute parlamentari, collegamenti in diretta, telecronache, ecc.), sia in semi-diretta (TG e altri programmi

registrati a ridosso della messa in onda), sia in pre-registrato per velocizzare le procedure di scrittura nei *software* di sottotitolazione.

Come risulta ovvio da una prima rapida analisi, generi diversi determinano discorsi diversi. Nel caso specifico dei notiziari, il *pattern* che viene seguito è abbastanza omogeneo per tutti i paesi: un giornalista legge le notizie del giorno dapprima sottoforma di riassunti, o 'titoli', poi in maniera più approfondita, facendo ricorso anche a supporti audiovisivi pre-registrati o in diretta che vedono la presenza o meno di *reporter* e di altri oratori (passanti, esperti, testimoni, ecc. nella stessa lingua del TG o interpretati), le cui competenze linguistiche, insieme allo stile della conduzione del tele-giornale e all'argomento trattato, determinano il discorso nella sua globalità.

Nel caso di sedute parlamentari, il discorso è invece assolutamente non determinabile. Se esistono formule di apertura e chiusura dei lavori oltre che per la turnazione e il registro è plausibilmente elevato, non sempre è possibile definire in maniera chiara il registro dei parlanti o il *mode of discourse*, che dipenderanno dall'argomento trattato e dalle competenze linguistiche dei singoli parlanti, siano essi parlamentari o esperti chiamati a intervenire nelle audizioni. A tal proposito, un aspetto forse non di immediata considerazione è l'idioletto di ogni singolo parlante, che può variare enormemente sia in termini puramente fonetici e prosodici, sia in termini lessico-grammaticali.

Per quanto riguarda le telecronache sportive, infine, gli inviati hanno solitamente una velocità di eloquio molto elevata (più dei giornalisti che intervengono al TG e dei politici chiamati a intervenire sotto rigide limitazioni di tempo), utilizzano un gergo molto specifico (ma anche circoscritto), molti nomi propri (che si riferiscono alle squadre, ai nomi degli sportivi, a città, competizioni e altro ancora) e tendono a utilizzare una struttura sintattica semplice proprio per far fronte alle esigenze dei telespettatori che assistono alla visione dell'evento commentato.

In tutti i casi analizzati, il buon esito del rispeakeraggio dipenderà essenzialmente dall'esperienza dell'operatore, dalla sua conoscenza della materia e infine dalla possibilità di aver potuto addestrare il *software* preventivamente. In questo ultimo caso, la differenza sarà di peso. Oltre a non riconoscere termini fondamentali, la mancata possibilità di addestrare preventivamente il *software* implica uno stress maggiore per il rispeaker, che deve sempre stare attento a dover correggere il testo o a cercare sinonimi per le parole in questione, e

conseguentemente peggiori risultati anche di riconoscimento dovuti a una voce sotto costante tensione.

2.5 Cenni psico-cognitivi

Come si è potuto ampiamente vedere, l'interpretazione simultanea e il rispeakeraggio sono due processi che implicano la contemporanea ricezione di un TP, la sua comprensione e l'elaborazione e la produzione di un TA 'di qualità' "under the time pressure imposed by the speed with which the source text is delivered" (Remael e van der Veer 2006). Dal punto di vista psico-linguistico, questa serie di operazioni compiute contemporaneamente implica da parte del rispeaker, così come anche dell'interprete di simultanea, una sapiente gestione della sua attenzione cercando di trovare un equilibrio tra la quantità di attenzione da dedicare all'ascolto e quella da dedicare alla riformulazione del TP, nel pieno rispetto dei criteri di qualità e cercando di gestire sia il carico cognitivo proveniente dal controllo dell'*output* sia quello proveniente dai limiti temporali imposti dal TP.

Per raggiungere questo obiettivo, una caratteristica che accomuna i due processi è la capacità di attuare una segmentazione del TP[70]. Come fanno notare Remael e van der Veer (2006), sia il rispeakeraggio, sia l'interpretazione simultanea "depend on segmentation of the source text with this difference that an interpreter is looking for units of meaning for which an equivalent can be found in the target language, whereas segmentation in respeaking aims at formulating text that is both correct and screen-ready". Non bisogna dimenticare, infatti, che se il rispeaker intra-linguistico, rispetto all'interprete, viene generalmente facilitato dall'assenza di cambiamento di lingua, il suo compito è complicato dalle imposizioni dettate dal *software* oltre che dalle norme che sono alla base della creazione del TA per cui il rispeakeraggio viene richiesto. Nel caso specifico della produzione di sottotitoli in diretta, il rispeaker dovrà prestare attenzione a rispettare i criteri di *readability* e *legibility* esposto da Gambier. Il rispeaker dovrà quindi innanzitutto eliminare dal TP tutte le caratteristiche del parlato (false partenze, pause piene, informazioni ridondanti, ecc.), rendendo il TA leggibile dallo schermo oltre che coerente, coeso e conciso grazie alla punteggiatura e a riformulazioni puntuali e immediate.

70 Cfr. Lederer 1981.

Per cercare di capire più a fondo quali siano le attività che regolano il processo del rispeakeraggio, è forse opportuno fare ricorso alla letteratura inerente gli studi sull'interpretazione simultanea. In particolare, visto che uno sguardo alla *Skopostheorie*, o meglio alla sua applicazione nell'interpretazione simultanea proposta da Pöchhacker, è stato già dato, restano da esaminare le restanti tre aree d'interesse teorico che potrebbero essere applicate anche al rispeakeraggio: il *modèle d'efforts*, in cui vengono descritte le operazioni svolte dall'interprete, la *théorie du sens*, che si concentra maggiormente sulla resa a cui dovrebbe tendere l'interprete e infine le teorie sulle strategie che ogni interprete esperto mette in atto per raggiungere tale scopo. Così facendo si avranno a disposizione tutti gli strumenti necessari allo studio del rispeakeraggio come attività cognitiva e al posizionamento dello stesso all'interno degli studi sull'interpretazione.

2.5.1 Il Modèle d'efforts

Il *Modèle d'efforts*, sviluppato da Daniel Gile nel corso degli anni Ottanta e Novanta, è uno dei contributi maggiori che gli studi sull'interpretazione offrono alla costruzione di un quadro teorico all'interno del quale cogliere l'essenza del rispeakeraggio. Gile parte dall'osservazione degli errori commessi da qualsiasi interprete nel corso del suo lavoro sia dal punto di vista del contenuto (omissioni, alterazioni del TP, perdita di informazioni, ecc.), sia dal punto di vista della forma (alterazione dei tratti distintivi della voce come pronuncia, tono, prosodia, accento, ecc.), per poi passare al vaglio le possibili ragioni (scarsa conoscenza dell'argomento e della terminologia, velocità o complessità dell'eloquio originale, cattivo funzionamento del mezzo di trasmissione delle componenti audio e video del TP, ecc.). A conclusione di tale disamina, Gile osserva che forse gli errori commessi dagli interpreti non sono di natura ambientale o (para-/extra-)linguistica, altrimenti non si spiegherebbero errori in testi obiettivamente facili da rendere. La ragione, secondo Gile, è da ricercarsi nella natura psico-cognitiva dell'azione dell'interpretare, vale a dire la presenza attiva e concomitante di due lingue nella mente dell'interprete. Sulla base di questa ipotesi e alla luce dei risultati dei suoi studi, Gile propone un modello per l'analisi delle attività che l'interprete deve portare avanti nel corso del suo mestiere e postula la teoria degli sforzi, secondo la quale l'interprete, sotto pressione a causa dei limiti temporali a cui è vincolato, deve coordinare i singoli sforzi, necessari per poter portare avanti proficuamente ogni attività che compone il processo interpretativo, in maniera tale da dare il

giusto peso a ognuno di essi. Ne consegue che ogni scarto da questo equilibrio ideale si traduce in uno sforzo aggiunto da parte dell'interprete a detrimento di uno degli altri sforzi con conseguente produzione di errori specifici (Gile 1985).

Quanto al modello, Gile parte dall'assunto, derivato dalle scienze cognitive, secondo cui esistono due tipi di operazioni mentali, automatiche e non automatiche. Le prime comportano un dispendio di energia minimo in quanto meccaniche, mentre le seconde necessitano di un'attenzione particolare da parte di chi le compie. Nel caso di due o più operazioni non-automatiche e concomitanti, il dispendio di energia cresce esponenzialmente sovraccaricando la *capacité de traitement*, ossia la capacità totale di ogni singola persona di operare diverse attività contemporaneamente. Nel caso dell'interpretazione simultanea, le operazioni da portare avanti simultaneamente sono tre e corrispondono ai tre sforzi che compongono il modello proposto da Gile:

- *sforzo di ascolto e analisi*: viene compiuto relativamente alla fase di percezione e comprensione del messaggio. Aumenta con l'aumentare delle difficoltà di ascolto (scarsa qualità dell'*input*, velocità di eloquio, ecc.) e di comprensione (densità delle informazioni, tecniciscmi, idiolettismi, ecc.);
- *sforzo di memoria*: concerne in particolare quella fase in cui i diversi tipi di memoria (ecoica, a breve-medio termine e a medio-lungo termine) interagiscono per permettere l'immagazzinamento di stringhe fonetiche, lessico-grammaticali e concettuali più o meno lunghe finalizzato alla ricerca della migliore soluzione possibile nella fase di produzione del TA;
- *sforzo di produzione*: si compie nel produrre il TA nel pieno rispetto degli standard di qualità[71]. Aumenta in caso di difficoltà nel trovare i giusti corrispettivi e si traduce in pause non naturali (per lo più piene) o soluzioni zoppicanti.

La *capacité de traitement* totale deve essere sempre superiore o uguale alla somma dei tre sforzi appena descritti compiuti contemporaneamente. Nel caso di aumento dello sforzo necessario all'espletamento di uno o più dei tre sforzi o per altri fattori esterni, le conseguenze si tradurranno in *défaillances*. Dallo studio di questi scarti dalla migliore delle interpretazioni possibili, si possono derivare matematicamente le ragioni per cui l'interprete si è trovato in difficoltà.

71 Cfr. Viezzi 1999.

Quanto alla 'simultaneità' dell'operazione, Gile si riferisce a segmenti diversi che vengono trattati nello stesso momento. In altre parole, l'interprete segmenta il TP in unità di significato e vi opera prima uno sforzo e poi i successivi in maniera sequenziale. Mentre attua ognuno di questi sforzi sull'unità di significato in questione, una parte del cervello è intenta a operare lo sforzo cronologicamente precedente sull'unità di significato successiva a quella in corso di elaborazione. Siano queste unità di significato 1, 2, 3, ecc. l'interprete attuerà lo *sforzo di ascolto e analisi* sul segmento 1. Non appena questa operazione si conclude, l'interprete inizierà a produrre contemporaneamente il medesimo sforzo sull'unità 2 e lo *sforzo di memoria* sull'unità 1. Successivamente e per tutta la durata dell'interpretazione, l'interprete produrrà tre sforzi simultaneamente su tre unità diverse e in particolare lo *sforzo di produzione* sull'unità 1, lo *sforzo di memoria* sull'unità 2 e lo *sforzo di ascolto e analisi* sull'unità 3[72].

Da questa breve disamina è possibile intravedere un applicazione del modello degli sforzi di Gile al rispeakeraggio, in cui sembrano essere presenti tutti e tre gli sforzi non automatici: ascolto e analisi, memoria e infine produzione. Tuttavia, nel rispeakeraggio, il processo di analisi o di produzione risultano essere intuibilmente inferiori rispetto a quanto descritto da Gile per l'interpretazione simultanea (soprattutto se verso una lingua straniera), in quanto l'operazione è intra-linguistica. D'altro canto però c'è da ricordare che nel rispeakeraggio, sia *verbatim*, sia *non verbatim*, è presente uno sforzo che potrebbe essere definito automatico: parlare al *software* in maniera chiara e 'riconoscibile'. L'automaticità dell'operazione deriva dall'intuizione che si tratta di uno sforzo che soggiace all'intera operazione traduttiva e al quale non può essere conseguentemente dedicata troppa attenzione, altrimenti si rischia di compromettere il resto delle operazioni. Rimane però un'operazione supplementare rispetto all'interpretazione simultanea, in cui non viene richiesta un'assoluta pulizia dell'emissione fonica.

2.5.2 La Théorie du sens

Da quanto emerge dal *modèle d'efforts* e dalla *ATT*, il rispeakeraggio e l'interpretazione simultanea sembrano essere non solo socio-linguisticamente, ma anche

72 Come già accennato, il triplice sforzo simultaneo continuerà fino alla fine dell'interpretazione nella maniera appena schematizzata. Chiaramente, nella realtà, questo schema non viene rispettato meccanicamente per tutta la durata dell'interpretazione perché altri fattori intervengono nella resa, come la difficoltà di traduzione o la velocità di eloquio dell'oratore. Le naturali conseguenze di questi fenomeni sono l'alterazione o addirittura l'interruzione della catena di montaggio del TA e la conseguente necessità di riassumere od omettere alcune unità di significato.

psico-cognitivamente, del tutto assimilabili, in quanto vengono meno due delle maggiori differenze esistenti tra i due processi, cioè il carico cognitivo e la funzionalità. Tuttavia, resta il fattore della direzionalità linguistica a pesare contro l'applicabilità delle teorie appena descritte al rispeakeraggio. A tal proposito, uno strumento che potrebbe essere d'aiuto è la *Théorie du sens*, che affronta l'unico aspetto non prettamente verbale dei tre sforzi, la concettualizzazione del TP.

Sviluppata alla fine degli anni Sessanta e nel corso degli anni Settanta da un gruppo di ricercatori e insegnanti coordinati da Danica Seleskovitch dell'*École Supérieure pour Traducteurs et Interprètes* (ESIT) di Parigi, la *théorie du sens*, dopo aver premesso i presupposti necessari a un interprete per svolgere appieno il suo lavoro (ottima conoscenza delle due lingue di lavoro, capacità traduttive, buona conoscenza dell'argomento da trattare durante la conferenza, ecc.), definisce la strategia interpretativa dalla quale dipende il lavoro dell'interprete, cioè a dire la resa del 'senso' del TP nella massima libertà linguistica dalle strutture lessico-grammaticali della lingua di partenza (Seleskovitch 1968). Seppur semplice e intuitiva, la *théorie du sens* focalizza l'attenzione sull'interprete e non sull'attività di semplice *transcodage* o resa meccanica delle *significations linguistiques*. Concentrandosi in particolar modo sul processo di analisi operato dall'interprete, che coglie, elabora e rende *le sens des énoncés* del TP, si vede chiaramente come l'interpretazione sia un'attività altamente intellettuale e basata sulle competenze situazionali e cognitive dell'interprete oltre che sulla sua conoscenza linguistica.

Sulla base di questa distinzione tra resa *mot-à-mot* e resa concettuale del TP, Lederer (1981) afferma che l'attività di comprensione del testo non è la somma di tante piccole attività di comprensione delle micro-componenti di un testo, ma un'attività olistica le cui micro-componenti si influenzano a un livello tale da costituire una *unité de sens* proprio in virtù del testo e del contesto in cui vengono espresse[73]. Ecco quindi che la resa del 'senso' diventa un'attività indipendente dalle lingue coinvolte in quanto le parole che comporranno il TA servono solo a esprimere un concetto che si possiede già nella propria mente, non sotto forma di *significations linguistiques*, ma di unità concettuali esprimibili indipendentemente dalla lingua di partenza.

2.5.3 L'interpretazione come attività strategica

73 Cfr. Chernov 2004.

Dall'introduzione socio-linguistica approntata precedentemente e dai brevi cenni psico-cognitivi dei paragrafi precedenti emerge la chiara corrispondenza tra l'interpretazione simultanea e il rispeakeraggio, sia *verbatim* che *non verbatim*, dal punto di vista del processo traduttivo. Nessuna differenza che li contraddistingue sembra essere influente nella possibilità di applicare le teorie comprovate per l'interpretazione simultanea al rispeakeraggio. A questo punto, risulta naturale concludere il quadro, attingendo ulteriormente agli studi sull'interpretazione simultanea. L'obiettivo sarà quello di definire il rispeakeraggio come attività cognitiva durante cui vengono operate delle scelte strategiche sulla base delle opzioni disponibili.

A tal proposito, oltre ai contributi di Færch e Kasper (1983), Dam (1993), Gile (1995) e Riccardi (2003), quello di Kohn e Kalina (1996) offre, forse più di tutti, un quadro esaustivo delle strategie utilizzate dall'interprete simultaneo nella produzione del TA. Partendo dalla convinzione che l'interpretazione simultanea è un *strategic discourse processing* (Kohn e Kalina 1996) con molti aspetti in comune con la comprensione e la produzione di un testo in lingua straniera, i due ricercatori tedeschi sviluppano un modello che muove le fila dallo *strategic model of discourse comprehension* di Kintsch e Van Dijk (1978) e van Dijk e Kintsch (1983) secondo cui la comprensione di un testo dipende dall'attuazione di sei diverse strategie:

- *propositional strategies*: l'*input* fonetico e morfo-sintattico si traduce nella mente di chi ascolta in comprensione lessico-grammaticale della struttura superficiale del TP;
- *local coherence strategies*: le singole proposizioni costruiscono nella mente di chi ascolta dei nessi logici che favoriscono la comprensione lessico-grammaticale del testo;
- *macrostrategies*: le macrostrategie utilizzate dall'oratore costruiscono la macrostruttura del TP che permette a chi ascolta di comprenderne l'evoluzione nel tempo e nello spazio;
- *schematic strategies*: il genere testuale a cui appartiene il TP, la cui conoscenza *a priori* permette a chi ascolta di anticipare o inferire elementi testuali;
- *production strategies*: le strategie generali attuate dall'oratore per veicolare il messaggio che possono essere selezionate sulla base della conoscenza condivisa del

mondo e della lingua e del contesto comunicativo all'interno del quale il testo viene prodotto;

- *other strategies for comprehension and production*: tutta una serie di strategie (stilistiche, retoriche, non-verbali e conversazionali) che fanno parte della conoscenza condivisa della lingua e che aiutano il pubblico a comprendere l'essenza di un testo.

Sulla base di questo modello e grazie ai contributi della pragmatica e della linguistica testuale (tra i tanti Grice 1957, De Beaugrande e Dressler 1981), Kohn e Kalina definiscono un modello che ogni interprete mette in atto in virtù della sua acquisita professionalità per superare possibili ostacoli legati al processo traduttivo che deve portare avanti (compresenza del TP, vincoli temporali, assenza di autonomia semantica oltre che di forma e contenuto, difficoltà/assenza di comprensione del TP, ecc.). Il modello in questione si concentra in particolar modo sul ruolo importante svolto dall'interrelazione tra la conoscenza condivisa della lingua e del mondo. Grazie a questa condivisione, che comprende anche le varie convenzioni discorsive che sono oggetto degli studi sulla pragmatica, la conoscenza del contesto comunicativo, la conoscenza del genere a cui appartiene il TP, ecc., è possibile attuare due tipi di approccio al processo traduttivo, ossia *bottom-up* e *top-down*, rispettivamente l'inferenza dai dati e l'interpretazione *a priori* sulla base delle aspettative e delle conoscenze pregresse. Sfruttando appieno il potenziale offerto da queste due metodologie interpretative, l'interprete potrà attuare diverse *processing strategies* come:

- *elaborative inferencing*: che permette l'anticipazione di elementi lessico-grammaticali o concettuali e conseguentemente una migliore resa stilistica;
- *memorising*: che permette una migliore resa stilistica attraverso la posticipazione di elementi lessico-grammaticali o concettuali memorizzati;
- *monitoring strategies*: l'interprete tiene sotto controllo la propria resa per tutto il corso dell'interpretazione così da correggere eventuali errori lessico-grammaticali o concettuali tramite diminuzione del *décalage*, riformulazione sintattica, segmentazione, correzioni *a posteriori*, ecc.;
- *adaptation strategies*: nel caso di divario lessico-grammaticale o concettuale tra le due lingue o culture l'interprete adatta linguisticamente il TA tramite

disambiguazioni, spiegazioni, riformulazioni, generalizzazioni, sostituzioni, parafrasi, riproduzione fonetica o naturalizzazione di un elemento del TA, ecc.;

- *neutralisation and evasion strategies*: in caso di dubbio, l'interprete non si impegna in affermazioni potenzialmente compromettenti, ma assolve al suo compito tramite generalizzazioni od omissioni, nel pieno rispetto della coerenza e della coesione testuali.

A queste strategie si aggiungono quelle discorsive, dipendenti dalla presentazione del testo, come la segmentazione, *strategies of repair* (quando ci si accorge di un errore commesso nell'elaborazione dell'*input*) o infine l'uso della prosodia e dei tratti sopra-segmentali. A tal proposito è forse interessante aprire una parentesi riguardante il rispeakeraggio. Se fino a questo momento le varie strategie descritte per l'interpretazione sono applicabili in varia misura anche al rispeakeraggio, lo stesso non vale, almeno in parte, per la segmentazione e per nulla per l'uso della prosodia e dei tratti soprasegmentali. Queste strategie infatti non sono, allo stato attuale della ricerca sul riconoscimento del parlato, applicabili come nell'interpretazione simultanea in quanto non producono l'effetto desiderato. In particolare, per quanto riguarda la segmentazione, le soluzioni saranno diverse a seconda del *software* in uso, o meglio della tecnica di proiezione del sottotitolo da parte del sottotitolatore. Qualora si voglia presentare i sottotitoli a blocchi, la segmentazione dell'*input* vocale è, accompagnata da un uso della punteggiatura, la migliore soluzione per ottenere sottotitoli comprensibili. Nel caso di proiezione *roll-up* o *scrolling*, l'unica strategia possibile è invece la punteggiatura. Quanto all'uso strategico dei tratti para- ed extra-verbali della voce, invece, non esiste ancora un software che possa tradurre questi tratti in simboli grafici. L'unica soluzione possibile è quindi un uso intelligente della punteggiatura e un'eventuale riformulazione disambiguante.

Per concludere l'aspetto delle strategie adottate in interpretazione, un caso particolare è rivestito dalle *emergency strategies* dovute a cause esogene (scarsa capacità di eloquio dell'oratore, eccessività del carico cognitivo, cattiva ricezione del TP, altri elementi 'esterni' di disturbo) o endogene (effetto *background*, difficoltà traduttive, scarsa conoscenza del testo, stanchezza, stress, ecc.). Nei limiti dei vincoli deontologici, alcune soluzioni possibili per la compensazione di problemi di comprensione del testo sono la riduzione del *décalage* (per evitare un carico cognitivo eccessivo o per tentare una

traduzione solo momentaneamente letterale) o la sua dilatazione (in attesa di informazioni disambiguanti); la selezione dell'informazione e la conseguente omissione di elementi non altamente informativi; il recupero di informazioni precedenti per compensare quantitativamente il vuoto lasciato da un'intera unità di significato; la generalizzazione; l'approssimazione; ecc. Nel caso di difficoltà traduttive dettate puramente da problemi linguistici, le strategie disponibili sono la semplificazione linguistica (lessicale, sintattica, o semantica), la parafrasi, la riformulazione[74], ecc.

2.6 Conclusioni

In questo capitolo si è visto come il rispeakeraggio condivida molti aspetti dell'interpretazione simultanea e dello *shadowing*. Tuttavia, per capire quanto veramente ci si possa spingere nei paragoni è stato necessario rivolgersi alla socio-linguistica di stampo hymesiano. Grazie a questo approccio contrastivo è stato possibile far affiorare tutte le sfaccettature che costituiscono la peculiarità del rispeakeraggio, nella sua duplice versione, *verbatim* e *non verbatim*. Considerate le grandi differenze che sono emerse tra rispeakeraggio e interpretazione simultanea (assenza di una situazione di bilinguismo, minore possibilità di intervento sul TP, supersemioticità del prodotto finale, diversa finalità, maggiore attenzione agli aspetti fonetici, ecc.), si è verificata la necessità di approfondire ulteriormente l'analisi dei due processi traduttivi, ricorrendo alla disamina dei tratti psico-cognitivi che contraddistinguono l'interpretazione simultanea.

Da questo studio, che si è soffermato sulle peculiarità dell'interpretazione simultanea come processo traduttivo, è stato possibile ricavare una serie di informazioni che possono servire a costruire un quadro teorico di riferimento per inquadrare il rispeakeraggio come processo. In particolare, l'*Allgemeine Translationstheorie* ha permesso di definire la ragion d'essere del rispeaker: la soddisfazione delle esigenze linguistico-cognitive degli utenti finali. Una volta determinato lo scopo del processo, il *modèle d'efforts* contribuisce alla determinazione delle tappe che portano al raggiungimento dello stesso da parte del rispeaker: sforzo di ascolto e analisi, sforzo di memoria e sforzo di produzione. La certezza della possibilità di applicare questo quadro teorico anche al rispeakeraggio deriva dalla *théorie du sens* che supporta quanto teorizzato anche da Gran (1992) alla luce dei suoi studi

74 Per riformulazione si intende qui l'operazione che Prandi (2004: 45) definisce meta-discorsiva in ragione della differenza tra significato (linguisticamente dato per scontato tra parlanti della stessa lingua) e contenuto del messaggio (che deve continuamente essere inferito hic et nunc).

psico-linguistci sugli interpreti di simultanea. In particolare, Lederer (1981) dice chiaramente che la fase di produzione del TA è indipendente dalla direzionalità linguistica in quanto avviene quasi in maniera meccanica una volta che si sono incamerate le unità concettuali sviluppate dal TP.

Chiarito quel che sembrava essere il maggiore ostacolo al raffronto tra interpretazione simultanea e rispeakeraggio, è stato possibile completare il quadro teorico attraverso la definizione delle strategie messe in atto dal rispeaker per raggiungere il suo obiettivo, passare dal TP al TA attraverso la serie di sforzi già menzionata. Il modello è stato preso in prestito da quello messo a punto da Kohn e Kalina (1996), che conferma che l'interpretazione simultanea, e quindi il rispeakeraggio, è un'operazione strategica grazie alla quale è possibile adempiere al proprio compito nel massimo equilibrio psico-cognitivo. Resta ora da vedere come il rispeakeraggio come prodotto si colloca all'interno degli studi sulla traduzione audiovisiva per poter così completare il quadro teorico all'interno del quale sarà possibile operare l'analisi strategica di un prodotto rispeakerato che è l'obiettivo di questo lavoro.

Capitolo 3 - La sottotitolazione per sordi di programmi pre-registrati

3.1 Premessa terminologica

Prima di iniziare a trattare la sottotitolazione intra-linguistica per sordi come disciplina a sé stante e ad analizzarne gli insegnamenti che possono essere tratti in vista della definizione di un quadro teorico di riferimento per l'analisi del rispeakeraggio come prodotto, è forse necessario fare qualche considerazione sulla terminologia in uso relativa alla sottotitolazione per sordi. Come fanno notare De Linde e Kay (1999: 1), dal punto di vista linguistico, la sottotitolazione può essere suddivisa in due tipi distinti: intra-linguistica (per sordi e audiolesi in generale) e inter-linguistica (di film in lingua straniera). La differenza tra le due tipologie starebbe, secondo gli stessi autori, in "the different requirements of deaf and hearing viewers" (*ibidem*). Gli autori vedono quindi la sottotitolazione inter-linguistica come una sottotitolazione volta esclusivamente agli udenti e quella intra-linguistica diretta all'accessibilità di un qualsiasi testo agli spettatori sordi. Quest'ultima descrizione è tendenzialmente vera visto che la sottotitolazione intra-linguistica è nata proprio con questo obiettivo e continua tutt'oggi a essere offerta proprio per rendere un testo audiovisivo accessibile ai sordi, malgrado tra i potenziali utenti di un servizio simile rientrino anche gli stranieri o persone che lavorano in ambienti rumorosi. Tuttavia, da un punto di vista teorico-terminologico, la prima definizione è forse un po' troppo circoscritta. Se, infatti, la sottotitolazione inter-linguistica ha come caratteristica principale quella di tradurre un testo da una lingua all'altra, non per forza l'utente di un prodotto audiovisivo straniero deve essere normoudente. Qualora un film straniero debba essere reso accessibile a un pubblico non-udente, la sottotitolazione, pur mantenendo intatta la sua natura inter-linguistica, avrà le caratteristiche tipiche della sottotitolazione intra-linguistica per sordi. Tenderà cioè a rappresentare anche la componente extra- e para-linguistica del TP.

Alla luce di queste considerazioni, si propone una diversa suddivisione della sottotitolazione: per udenti e per non-udenti conformemente alla terminologia usata negli Stati Uniti[75]. Questa terminologia ha il vantaggio di dire chiaramente che l'obiettivo, e

75 Visto che la traduzione audiovisiva è nata interessandosi di doppiaggio e di sottotitoli 'normali' (cioè interlinguistici, cioè per normoudenti) e che solo di recente è emersa un'attenzione particolare alla 'special needs subtitling', nella

conseguentemente il risultato, dei due processi traduttivi è diverso indipendentemente dalla lingua di partenza.

Inoltre, sebbene l'errata convinzione di molte persone che i sottotitoli inter-linguistici possano soddisfare sia le esigenze degli udenti (per cui sono prodotti) che quelle dei sordi sia stata da tempo abbandonata, è forse bene ribadire che, a parte la velocità di lettura (che varia a seconda delle singole persone sorde e a seconda dei paesi), i sottotitoli 'per normoudenti' traducono, nel senso letterario del termine, soltanto la componente linguistica della traccia audio di un prodotto audiovisivo. Come fa notare Gottlieb (2005), la componente audio di un film veicola sia informazioni prettamente linguistiche che extra- e para-linguistiche che contribuiscono in varia misura alla comprensione del TP. È evidente quindi che mentre l'udente riuscirà facilmente a cogliere tutti gli aspetti extra-linguistici e in parte anche quelli para-linguistici (alcuni sono indissolubilmente legati alla lingua di partenza e non possono pertanto essere compresi dall'utente del TA o trasferiti nei sottotitoli), l'utente audioleso non potrà percepirli se non attraverso una sistematica trasposizione da parte del sottotitolatore.

In sintesi, quindi, se da un lato la sottotitolazione per udenti e quella per non-udenti si assomigliano perché entrambe traducono lo stesso tipo di testo, dal canale orale a quello scritto, riducendolo quantitativamente in maniera da rispettare i limiti spazio-temporali imposti dal mezzo di trasmissione del TA oltre che dalle differenze nella ricezione di un testo orale rispetto a uno scritto; dall'altro, l'obiettivo della sottotitolazione inter-linguistica per udenti è di tradurre un prodotto audiovisivo per soddisfare delle carenze linguistiche, quello della sottotitolazione (sia inter- che intra-linguistica) per non-udenti è di soddisfare delle carenze sensoriali.

Appurato che i termini intra-linguistico e inter-linguistico risultano ambigui ai fini del presente lavoro, saranno sostituiti dalle espressioni sottotitolazione per udenti (o per normoudenti) e sottotitolazione per non-udenti (o per sordi o per audiolesi).

---

comunità scientifica si parla comunemente di Subtitling for the Deaf and the Hard-of-Hearing (SDH), ma raramente la sottotitolazione per udenti viene definita con un'espressione diversa da sottotitolazione interlinguistica o sottotitolazione tout court. Negli Stati Uniti e in Canada e in altri paesi che utilizzano il sistema della Line 21, invece, si parla di captions per definire i sottotitoli per sordi e di subtitles per i sottotitoli per udenti.

3.2 Cenni storici

La storia della sottotitolazione intra-linguistica per sordi inizia con i primi anni di vita del cinematografo. A partire dalla geniale invenzione del cinematografo, attribuita ai fratelli Lumière nel 1895, l'unica fonte audio a cui si era esposti nelle prime sale deputate alla proiezione delle pellicole era quella del piano, suonato *in loco*, che accompagnava lo scorrere delle immagini. In quegli anni, visto che la tecnologia non permetteva la riproduzione del suono, le pellicole venivano concepite per essere viste e non per essere ascoltate. La necessità di comprendere i dialoghi originali era pertanto intuitivamente assente. Tuttavia, emerge nei registi dell'epoca la necessità di arricchire sempre più le proprie rare produzioni. Ecco, quindi, che nell'Europa del 1903 si assiste, per la prima volta al mondo, all'introduzione, direttamente sulla pellicola, di *intertitoli*, cioè didascalie (De Linde 1996: 173) che alternavano il susseguirsi delle scene e che aggiungevano testo (spiegazioni, descrizioni o brevi dialoghi) a quella che fino a quel momento era considerata "un'arte, notoriamente la settima, prettamente 'visuale'" (Perego 2005: 34).

Con l'avvento del sonoro, la situazione è cambiata notevolmente. Visto che le nuove tecnologie permettevano di far parlare gli attori, i dialoghi sono aumentati in maniera esponenziale e la comprensione del TP si è fatta immediatamente più complessa. Per i sordi, questo ha comportato un duplice svantaggio: da un lato, si sono trovati a fruire dei film in posizione di svantaggio rispetto agli udenti, mentre fino a quel momento erano sullo stesso piano; dall'altro, molti sordi, che durante il periodo del cinema muto erano apprezzati attori per via della componente mimica delle lingue dei segni, si sono ritrovati senza lavoro. Uno di questi attori, il cubano Emerson Romero, decide di utilizzare la tecnica dell'intertitolo per riprodurre i dialoghi dei primi film sonori. Come fa notare Neves, nonostante l'idea fosse ottima, il risultato non era altrettanto buono in quanto "this meant that text and image would alternate rather than co-exist as came to happen later" (2005: 107-108) raddoppiando il tempo di proiezione. Visto che la tecnologia permetteva la triplice riproduzione degli effetti sonori, della voce umana e delle musiche, infatti, i registi progettavano le loro pellicole anche in funzione della componente audio facendo un uso sempre più massiccio dei dialoghi. L'uso dell'intertitolo, in questi casi, comportava un sensibile prolungamento dei tempi di esposizione alla pellicola oltre che una frammentazione costante della stessa. Poco dopo, nel 1949, Arthut Rank sviluppa ulteriormente questa tecnica e inventa un meccanismo

in grado di offrire una sottotitolazione perfettamente in sincronia con i dialoghi sullo schermo. Tuttavia, anche in questo caso, la funzionalità del meccanismo non permetteva agli utenti un'agevole fruizione del film in quanto obbligava gli spettatori a dover volgere lo sguardo dallo schermo che proiettava il film a un secondo schermo posizionato in basso a sinistra. Con i progressi tecnologici, la sottotitolazione per sordi si evolve sempre più fino ad assumere, con la nascita del *teletext*, le forme che oggi conosciamo.

3.3 Aspetti tecnici

Indipendentemente dalle sue origini, la sottotitolazione per non-udenti ha ottenuto particolare visibilità grazie alla televisione dove fa la sua prima apparizione nel 1972, in un episodio di *French Chef*. Risulta quindi necessario approfondire maggiormente gli aspetti tecnici che in varia misura influenzano la produzione dei sottotitoli per sordi in televisione. Verranno quindi presi in esame non solo i due sistemi più diffusi per la trasmissione dei sottotitoli televisivi, il *teletext* e la *Line 21*, ma anche le modalità di proiezione e le varie applicazioni dei sottotitoli per sordi che vincolano la sottotitolazione dal punto di vista sia della produzione (vincoli spazio-temporali) che del prodotto (modalità di visualizzazione).

3.3.1 Servizi di informazione televisiva

Il *teletext* è un servizio di fornitura di informazioni tramite la televisione analogica sviluppato nel Regno Unito dalla BBC nei primi anni Settanta con il duplice obiettivo di diffondere notizie in maniera alternativa al telegiornale e di diffondere i neonati sottotitoli per non-udenti. La grande flessibilità del *teletext* permetteva di immagazzinare informazioni nel televisore per essere selezionate in un secondo momento. Questo è stato l'elemento di successo di tale sistema che si è diffuso in tutto il mondo adattandosi ai diversi sistemi di trasmissione del segnale analogico televisivo (PAL, SECAM, ecc.).

Oltre a servizi come notizie dell'ultim'ora, previsioni meteo, programmazione radiotelevisiva, estrazioni del Lotto e molto altro ancora, il servizio di *teletext* è di fondamentale importanza per le persone audiolese in quanto permette loro di avere accesso al servizio di sottotitolazione dei programmi di ogni singola emittente che offre tale servizio.

Come è stato appena accennato, il segnale del *teletext* viene trasmesso insieme al resto del segnale televisivo analogico in quello che viene chiamato *Vertical Blanking Interval*, una serie di righe del segnale televisivo che non contengono informazioni visive e che pertanto restano 'invisibili' benché presenti nella memoria del televisore e attivabili su richiesta. A seconda dell'ampiezza della banda sfruttata dal segnale analogico per trasmettere le informazioni televisive, le righe che di solito vengono utilizzate dal *teletext* occupano l'intervallo tra la 6 e la 22 e tra la 318 e la 335. In queste righe vengono inviate dall'emittente le pagine del *teletext* una alla volta a flusso continuo. Quando viene richiesta una pagina il *decoder* aspetta che venga inviata la pagina richiesta e la visualizza sullo schermo. Moderni televisori sono dotati di una memoria specifica all'interno della quale, fin dal momento in cui il televisore riceve il segnale dall'emittente, vengono registrate tutte le pagine del *teletext* che possono essere così immediatamente visualizzate senza dover aspettare che la pagina richiesta venga proiettata a flusso continuo dall'emittente.

Attualmente, grazie agli sviluppi tecnologici nel settore della televisione digitale, il sistema *teletext* sta cedendo il passo al sistema di trasmissione digitale, ma il termine *teletext* continua a essere utilizzato per definire i moderni sistemi come il britannico *MHEG-5* o la più diffusa *Multimedia Home Platform*.

Benché il sevizio di *teletext* e la sua evoluzione in digitale siano ormai diffusi in una vasta zona del mondo, i limiti di questo servizio hanno portato in altre epoche e in altre regioni del mondo coperte da altri sistemi di diffusione del segnale analogico (NTSC, SECAM, ecc.) a sviluppare servizi di informazione televisiva con una grafica più complessa e un numero di pagine maggiore. In Francia, verso la fine degli anni Settanta, è nato *Antiope* che sfruttava il sistema SECAM e che, grazie a una potenza maggiore del servizio, permetteva una grafica più dinamica e interattiva. Malgrado questi vantaggi, nel 1991 è stato sospeso in favore dell'essenziale *teletext* standard. In Canada, il sistema *Telidon* ha subito un percorso simile ad *Antiope*. Dal 1983 al 1986, è stato utilizzato perché permetteva di utilizzare una maggiore potenza di trasmissione per una risoluzione grafica maggiore, ma ha finito per cedere al sistema più semplice.

In Nord America, in alcune regioni del Sud America, in Giappone, in Corea del Sud e in altri paesi dell'Africa e dell'Asia, dove il sistema di diffusione del segnale analogico è l'NTSC (*National Television System Committee*), sono stati sperimentati vari sistemi (*World System Teletext*, NABTS, *Electra*, *WaveTop*, *Guide Plus*, *Star Sight*, ecc.), spesso

contemporaneamente, ma che hanno finito per cedere agli alti costi dei *decoder*, alla maggiore diffusione del sistema concorrente o, più semplicemente, alla non standardizzazione dei sistemi proposti. In tutti questi paesi, gli unici sistemi di trasmissione delle informazioni che sono riusciti a sopravvivere alla concorrenza del *teletext* sono, attualmente, i servizi di *closed captioning*, *TV Guide On Screen* e l'*eXtended Data Services*[76], attivabili a richiesta e gestiti da *decoder* diversi. I *closed captions*[77] sfruttano il sistema analogico EIA-708 e, come il servizio di *teletext*, sono trasmessi nel *Vertical Blanking Interval*, più precisamente nella riga 21, donde l'altro nome con cui sono conosciuti, *Line 21*.

Soprattutto in Canada e Stati Uniti, visto l'ambiente multilingue della regione, i *closed captions* non sono pensati esclusivamente per persone audiolese, ma anche per persone che vogliono imparare una lingua, che non la conoscono o semplicemente che vivono o lavorano in ambienti rumorosi. Secondo il *National Captioning Institute*, infatti, la maggior parte dei fruitori del servizio di *closed captioning* sono udenti che hanno l'inglese come seconda lingua. Per questa ragione, il concetto di accessibilità è, in Canada e negli Stati Uniti, sinonimo di trascrizione il più possibile *verbatim*.[78]

A tal proposito, una curiosa differenza tra la *Line 21* e *il teletext* sta nella modalità di accesso a due o più diversi tipi di sottotitoli per lo stesso programma. Nel caso di una doppia sottotitolazione (una intra-linguistica e l'altra inter-linguistica), infatti, il *teletext* proietta su pagine diverse le due versioni dello stesso programma. La *Line 21*, invece, dispone di quattro canali diversi noti come CC1, CC2, CC3 e CC4. Nei primi due canali vengono trasmessi i sottotitoli intra-linguistici[79] e negli altri due una o due versioni inter-linguistiche (nel caso degli USA spagnolo e portoghese o francese e nel caso del Canada francese ed eventualmente spagnolo). Tutte e quattro le versioni possono essere visualizzate contemporaneamente.

76 La TV Guide On Screen è il comune servizio di guida ai programmi TV offerto anche da teletext. L'XDS è invece un servizio simile a quello offerto dai canali satellitari e che visualizza il giorno, l'ora e il nome dell'emittente, del canale e del programma in corso. Può anche offrire contatti con l'emittente (e-mail, numero di telefono e di fax).

77 Negli Stati Uniti e in Canada, il termine closed significa che i sottotitoli sono visualizzabili solo se attivati (a differenza dei sottotitoli inter-linguistici per udenti che sono solitamente open, cioè incisi sulla pellicola e visibili da tutti). Captioning, come si è già detto, designa i sottotitoli per sordi.

78 Anche nel Regno Unito, i sottotitoli per sordi sono seguiti da ben sei milioni di persone audiolese su circa sette milioni e mezzo di utenti regolari. I sottotitoli intralinguistici sono il 97% del testo audiovisivo di partenza.

79 CC2 viene solitamente utilizzato come canale di scorta qualora ci dovessero essere problemi di trasmissione tramite CC1. Viene anche talvolta sfruttato per proiettare versioni semplificate per bambini.

3.3.2 Sistemi di proiezione

Indipendentemente dalla tecnologia sfruttata per la diffusione delle informazioni televisive, la fase di preparazione e la visualizzazione dei sottotitoli per sordi presentano molte somiglianze. Sia il sistema analogico che quello digitale infatti permettono alle pagine di comparire sullo schermo sostituendo o sovrapponendosi alle immagini del programma diffuso dall'emittente che gestisce il servizio in questione. Per motivi di visibilità del TP, i sottotitoli vengono generalmente proiettati su una, due o talvolta tre righe e possono essere spostati proprio per evitare di nascondere parti salienti delle immagini sullo schermo.

Per quanto riguarda la modalità di proiezione e conseguentemente di visualizzazione sullo schermo, esistono tre diversi stili:

- *roll-up* (o *scroll-up* o *scrolling*): secondo la descrizione del *National Captioning Institute* (2002), il *roll-up* è un sistema di proiezione di titoli in cui, questi ultimi, invece di comparire e scomparire, sfilano, su tre righe (in Europa tendenzialmente due), dal basso verso l'alto dello schermo. Il titolo successivo compare nella parte bassa dello schermo e sale spingendo le altre righe verso l'alto finché quella più in alto scompare. In alcuni casi, le parole della riga più in basso scorrono una per una da sinistra verso destra fino al riempimento dell'intera riga. In altri, la riga viene proiettata per intero[80]. Questo sistema viene privilegiato per il dinamismo e il ritmo che gli sono tipici. Il sistema *roll-up* permette inoltre di guadagnare tempo, e quindi denaro, in quanto offre la possibilità di non passare per una fase di sincronizzazione con la pronuncia delle battute perché una didascalia rimane sullo schermo anche una volta che la battuta corrispondente è stata pronunciata. Inoltre, nel caso di spettatori sordi, se qualcuno è un po' più lento nella lettura può contare su quest'ultimo aspetto per poter terminare la fase di lettura dei sottotitoli. Curiosamente, questa peculiarità è anche quella maggiormente criticata dai detrattori del *roll-up*. Secondo gli studi condotti da Sancho-Aldridge e IFF Research ltd. (1996) e Captionmax (2002), infatti, il pubblico trova macchinoso e ingombrante il movimento continuo dei titoli che li distrae dal filo del discorso. Se basta concentrarsi su una soltanto delle tre righe proiettate sullo schermo per seguire il discorso (visto che le righe salgono incessantemente), il pubblico poco

80 La prima modalità viene solitamente adottata nei casi di sottotitolazione dal vivo mentre la seconda per la proiezione di sottotitoli di programmi pre-registrati.

allenato a questo tipo di proiezione viene distratto dalle altre due righe e finisce inevitabilmente per rileggere più volte la medesima riga. Sempre secondo gli stessi studi, sarebbero gli anziani ad avere maggiori difficoltà a seguire questo sistema mentre i giovani, abituati maggiormente per motivi storici a farne uso, riescono a servirsene in maniera migliore;

- *pop-on* (o *pop-up* o *block*): questo sistema prevede che i sopratitoli si presentino sullo schermo in blocco. Tendenzialmente si tratta di una o due righe che compaiono insieme nel momento esatto in cui viene pronunciata la battuta, restano sullo schermo il tempo necessario alla lettura per poi scomparire insieme sostituite da un'altra didascalia di uno o due righe di sottotitoli o vuota. Questo metodo è molto diffuso nel mondo e viene utilizzato per tutte le produzioni pre-registrate. Per quanto riguarda la semi-diretta è utilizzato in molti paesi tra cui anche l'Italia. Il *pop-on* ha il vantaggio di essere un sistema molto preciso perché permette la sincronizzazione delle didascalie con l'inizio delle battute, di garantire la visualizzazione di un'unica unità concettuale nonché di poter posizionare la didascalia nella parte sinistra, destra o centrale dello schermo così da agevolare l'individuazione di chi sta parlando. Riferendosi in particolare al sopratitolaggio, Le Du e Petit (*op. cit.* in Desmedt 2002) sono concordi nel sostenere che il *pop-on* ha il vantaggio di facilitare la lettura del sottotitolo da parte dello spettatore sordo che riesce a capire bene la scansione sintattica del testo audiovisivo e quindi a comprenderne il significato;
- *paint-on*: incorporato nel *roll-up* da emittenti rinomate come la BBC, questo stile di proiezione prevede la formazione del sottotitolo parola per parola o lettera per lettera, da destra verso sinistra fino a coprire un'intera riga. La modalità *paint-on* prevede che una volta ultimata questa operazione, la riga appena prodotta scompaia in modalità *pop-on*. In alternativa, è questo il caso della BBC, la riga di sottotitoli appena formata sale verso l'alto spinta dalle parole componenti il sottotitolo successivo.

A seconda dei paesi o delle emittenti viene utilizzato il *pop-on* o il *roll-up* (anche combinato con il *paint-on*) o entrambi. In Italia, la RAI fa uso delle due modalità per sottotitolare i telegiornali. I sottotitoli per i servizi già montati e per il testo che il giornalista

in studio legge dal telesuggeritore sono preparati precedentemente la messa in onda del TG e proiettati in diretta in modalità *pop-on*, perché ritenuta di più agevole lettura. Nei servizi in diretta, sia la produzione che la proiezione dei sottotitoli avvengono in tempo reale. Per evitare di ritardare troppo la comparsa del sottotitolo sullo schermo, le singole parole vengono proiettate in modalità *paint-on*.

3.3.3 Il futuro della sottotitolazione per non-udenti

Oggi, grazie a numerosi provvedimenti legislativi europei e nazionali (non da ultima la direttiva europea *Television Without Frontiers*), la sottotitolazione per sordi si è diffusa enormemente sia negli Stati Uniti che nella maggior parte dei paesi europei. Ad oggi, come afferma la *European Federation of the Hard-of-Hearing* (EFHOH), il numero di ore di programmi televisivi sottotitolati è aumentata sempre più nel corso degli ultimi anni. Secondo i dati in loro possesso, nel 2006, le nazioni che più hanno sottotitolato le loro trasmissioni sono state la Danimarca (49%), il Belgio (50%), la Norvegia (50%), la Svezia (50%), i Paesi Bassi (75%), l'Irlanda (80%) e il Regno Unito (80%).

In futuro, la sottotitolazione per non-udenti continuerà ad allargarsi ad altri ambiti, diversi dalla televisione. A parte applicazioni *una tantum*, la sottotitolazione di programmi pre-registrati trova infatti sempre più spazio al cinema e nel mercato dei DVD. Per quanto riguarda quest'ultimo, c'è una tendenza da parte delle case produttrici a riservare almeno una delle 32 tracce all'interno di ogni disco alla produzione di sottotitoli per sordi. Secondo uno studio condotto da Neves (2005) su un campione di 250 DVD prelevati in vari video-noleggi portoghesi, i sottotitoli per sordi sono, nel 25% dei casi, in inglese e nel 6% dei casi in altre lingue (tedesco, 4,8% e italiano, 1,2%) (2005: 117). Quanto al cinema e al teatro, l'offerta è sempre più ampia anche in ragione delle soluzioni proposte. Fra le numerose soluzioni provenienti quasi tutte dal Nord America, Desmedt (2002) cita ben dodici sistemi diffusi in molti teatri e sale cinematografiche sia americane che europee.

Sempre secondo Neves (2005: 118), la digitalizzazione delle comunicazioni e degli impianti elettronici porterà all'interazione tra i videoregistratori, i televisori, i DVD, il cinema e il *personal computer* con conseguenze rilevanti per gli utenti sordi. Se, fino a oggi, gli utenti devono infatti accontentarsi di avere dei sottotitoli nelle modalità decise dai committenti e dai sottotitolatori, la tecnologia digitale permetterà ai sordi di scegliere non solo la modalità di visualizzazione dei sottotitoli e il *layout*, ma anche la versione che riterrà

più adeguata alle sue competenze linguistiche e alla sua velocità di lettura. In particolare, non sarà difficile aumentare o diminuire le dimensioni dei caratteri, selezionare il tipo, lo stile e il colore del carattere che più agevolano la lettura, eliminare o inserire didascalie che mettono per iscritto la componente non verbale della colonna sonora del film (effetti sonori, canzoni e musiche di sottofondo, tono, intonazione, accento di chi parla, ecc.), decidere, infine, se seguire i sottotitoli in 'versione trascrizione' o in 'versione riformulazione'. A tal proposito, c'è da chiarire che la scelta tra trascrizione fedele e riformulazione dipenderà dalla disponibilità delle due versioni. Preparare due versioni dello stesso TP potrebbe infatti implicare costi aggiuntivi. In realtà, nel caso di film da sottotitolare intra-linguisticamente (perché la sottotitolazione viene effettuata nella stessa lingua dell'originale o nella lingua della versione doppiata), la presenza *a priori* di una traccia scritta del TP, e quindi della versione trascrizione, riduce le operazioni alla sola riformulazione.

3.4 Aspetti traduttivi

Dal punto di vista linguistico, la sottotitolazione intra-linguistica per sordi è volta alla soddisfazione di esigenze specifiche, nel pieno rispetto dei vari sistemi semiotici integrati (audio e video sia verbali che non verbali). Per questo motivo, il sottotitolatore per sordi, rispetto al sottotitolatore per udenti, dovrà tenere in considerazione l'impossibilità da parte degli utenti finali di fruire di due componenti semiotiche fondamentali, il sistema audio verbale e quello audio non-verbale. A loro volta, i due sistemi audio veicolano due tipi di informazione: fonetica e sonora, ovvero informazione fonetica sia grammaticalizzata (informazione linguistica) che non (extra-linguistica) e informazione non fonetica che può contribuire (para-linguistica) o meno (extra-linguistica) al sistema linguistico generale. Queste componenti e loro suddivisioni interne devono poter essere 'percepite' dagli utenti sordi. Ma quali strategie debbono essere attuate per poter far percepire al pubblico queste due componenti? Bisogna limitarsi a descrivere o bisogna interpretare? Trascrivere o semplificare? Come riassume bene Neves (2005: 205), una domanda ancora inevasa riguarda proprio questo aspetto:

> Do Deaf people really benefit from subtitles that read like written speech or will they benefit more from having subtitles that read like written text to which they relate through usage?

Per rispondere a questa domanda, si tenterà, grazie ai contributi nel settore di De Linde e Kay (1999) e della stessa Neves (2005), di raggruppare le strategie utilizzate nella sottotitolazione per sordi per veicolare tutte queste informazioni, che verranno suddivise a seconda del sistema a cui fanno riferimento, il sistema verbale e quello non verbale.

3.4.1 Componente verbale

Per sottotitolare la componente verbale del prodotto audiovisivo originale, i professionisti tendono a impiegare diverse tecniche. Per trasmettere informazioni circa gli accenti e le lingue straniere, visto che "a phonetic representation of a speaker's foreign or regional accent may slow up the reading process and possibly ridicule the speaker" (De Linde e Kay 1999: 13), la soluzione solitamente adottata per trattare questo tipo di informazioni è di indicare tra parentesi o in corsivo la tipologia di accento[81] o la lingua straniera. Senza voler propendere per una soluzione o l'altra, Neves (2005: 216-217) si limita a sottolineare come la tradizione britannica sia di "keep the flavour of orality and transcribe personal idiosyncrasies" mentre

> most analysed guidelines[82] […] show that caution is needed when dealing with this issue by proposing that only one instance of "deviant language" should be included in any one subtitle or that marked language be limited to the bare minimum to get the message across. (*ibidem*)

Quanto alla trasposizione dello *humour*, qualora siano da riportare casi di *verbally expressed humour* (VEH), Chiaro (2006) dice chiaramente che "no matter how complex issues concerning the translation of written and spoken instances of VEH may be, they are relatively simple when compared to the intricacy of having to translate them when they occur within a text created to be performed on screen". In ogni caso, qualora un equivalente

81 Oltre a provenienza geografica, età e genere, l'accento può rivelare, a seconda dei paesi, anche l'estrazione sociale e l'istruzione del parlante.
82 Come parte dei suoi studi di dottorato Neves ha analizzato 15 manuali di sottotitolazione per sordi di 15 diverse emittenti europee

sia difficile da trovare, Chiaro (2006) suggerisce che "equivalence will need to be relinquished in favour of skopos". Tale affermazione non vale soltanto per il doppiaggio e per la sottotitolazione inter-linguistica per udenti, ma anche per la sottotitolazione intra-linguistica per sordi. La compenetrazione tra la componente video e i sottotitoli deve essere infatti il più collaborativa possibile nei casi di *humour* derivante da un'interrelazione tra le immagini e le parole, tra i suoni e le parole o tra la lingua scritta e la lingua parlata. In quest'ultimo caso, come fanno notare De Linde e Kay "Homophones, for example, cause particular problems for deaf viewers as the oral component is not recoverable from the sound track. One way of preserving both meanings in a pun is to spell out the word according to the less obvious meaning" (1999: 13). Tuttavia, in seguito a un lungo studio svolto in Portogallo, Neves (2005: 210) fa notare come persone sorde segnanti

> enjoy telling jokes and laugh heartily at the jokes they tell each other. However, it was often the case that they did not react in the same way to the jokes they read. […] Further probing showed that they had not arrived at the implied meanings, and once explained, they did understand the pun, but they did not see the fun of it.

Esiste, dunque, una via di uscita? Solitamente si cerca di trovare una soluzione intermedia, che si ponga tra la semplice trascrizione e la spiegazione della battuta. Nel caso di un doppio senso che si basi su omofonia, abbiamo visto come la soluzione più diffusa sia la trascrizione del termine meno comune. Ma una panacea sembra impossibile. Riprendendo le parole di Chiaro, si può concludere dicendo che per quanto difficile possa essere la traduzione di VEH nel doppiaggio o nel sottotitolaggio, risulta essere relativamente semplice se paragonata alla sottotitolazione per sordi segnanti, in quanto implica una trasposizione da un codice all'altro (dall'orale allo scritto) per un pubblico che nella sua quotidianità non solo non condivide la lingua del TP, ma nemmeno quella del TA oltre che il codice dei sottotitoli (quello scritto, estraneo alla tradizione segnica).

Per concludere, le parole o espressioni volgari sembrano essere considerate *taboo* da molte televisioni, ma sempre più film ne fanno uso. Allo stesso modo, la maggior parte dei ricercatori continua a considerare la questione come un problema che il sottotitolatore non deve porsi dato che il valore informativo di ogni espressione volgare è di indiscutibile importanza. D'altra parte però è altrettanto inoppugnabile la considerazione che, in forma scritta, queste parole acquisiscono una forza maggiore che, a seconda della tradizione

letteraria oltre che dei valori condivisi di ogni singolo paese, può rasentare l'illegalità tanto da giustificarne la censura. Nel Regno Unito così come in Italia e in altri paesi europei esiste una fascia protetta durante la quale la televisione si considera responsabile del contenuto dei propri programmi nei confronti dei minori. Come si deve comportare dunque un sottotitolatore? A tal proposito, già Baker *et al.* (1986: 40) consigliavano ai professionisti di non censurare il TP a meno che non vi fossero dei vincoli di tempo. Questa affermazione è ancor più valida se applicata alla sottotitolazione intra-linguistica per sordi. Se un programma non è stato censurato nella forma perché ritenuto idoneo al potenziale pubblico (udente), il pubblico non-udente dello stesso programma non può essere considerato meno capace di intendere una determinata espressione.

3.4.2 Componente non-verbale

Malgrado l'importanza di veicolare la componente verbale sia grande ai fini della comprensione del testo audiovisivo, bisogna ricordare che questa non è l'unico fattore che contribuisce alla composizione e alla comprensione di un qualsiasi testo. Anche la componente non-verbale concorre, in varia misura, al raggiungimento di questo duplice scopo. A tal proposito, Poyatos (1997: 17-47) ci dimostra come le tre componenti della comunicazione (lingua, para-lingua e cinetica), siano distribuite in varia misura all'interno di un qualsiasi testo svolgendo ruoli diversi a seconda della funzione del testo in questione. Nel caso di prodotti audiovisivi, l'importanza della componente prettamente verbale sembra essere intuitivamente inferiore rispetto ad altri ambiti come il giornalismo su carta stampata o addirittura i romanzi non illustrati. Se è vero che i sordi sono abituati a trarre più informazioni dalla componente cinetica di un film rispetto a un normoudente, è anche vero che non percepire alcuni suoni 'invisibili' intralcia la comprensione del prodotto finale. Nonostante sia così rilevante, sia De Linde e Kay (1999), sia Neves (2005) sottolineano come quest'aspetto sia stato trascurato

> under the assumption that viewers will be receiving, through image and sound, the information that comes encoded in para-linguistic or non-linguistic signs. When actual speech is relayed in [verbatim] subtitles […], these often become deficient in cohesion, grammatically incorrect or as unstructured strings of words that would be quite incomprehensible if not accompanied by image and sound (Neves 2005: 206).

Da queste breve disamina, risulta chiara l'importanza del componente non-verbale nella comprensione del testo finale. Prima di entrare nei dettagli è forse opportuno indicare chiaramente che all'interno della componente non-verbale rientrano sia i tratti para-linguistici (tono, intonazione, volume, ritmo velocità, prosodia, ecc.), sia quelli extra-linguistici (effetti sonori, musica di sottofondo, cinetica, mimica, prossemica, ecc.).

Componente para-linguistica

Tra i due tipi di tratti, quelli para-linguistici, essendo più strettamente collegati alla lingua, sono stati oggetto di studio già a partire dai primi lavori sulla traduzione audiovisiva in generale. L'enfasi viene solitamente segnalata tramite l'uso di corsivo o di colori diversi. Un aumento di volume invece tramite maiuscole (Baker ***et al.*** 1984: 29). L'esitazione infine tramite l'uso di punti di sospensione o di interruzione di riga (BBC 1994 ***op. cit. in*** De Linde e Kay 1999: 13):

No…

…But I don't dislike him.

Il tono di voce può essere reso tramite un uso equilibrato dei sistemi semiotici. Qualora le immagini non suggeriscano l'intento comunicativo di chi parla, di solito vengono impiegati i punti esclamativi, i punti interrogativi o entrambi per designare sarcasmo, stress, ironia o ambiguità. Quando risulta difficile il ricorso a tali strumenti si può indicare tra parentesi il chiaro intento illocutorio. Tuttavia, anche in questo punto è forse giusto porsi una domanda: descrivere o inferire? Se una voce trema, chiaro segno di nervosismo, ma anche di forte emozione, bisogna indicare tra parentesi che la voce trema o che il personaggio è nervoso o emozionato? Se da un parte la trascrizione risulta essere la soluzione più comoda per il sottotitolatore e meno scorretta nei confronti dei sordi, dall'altro è anche vero che non sempre è facile decifrare uno stato d'animo di un personaggio tramite la semplice descrizione delle caratteristiche fisiche della sua voce. Come afferma Gambier (2006), "AVT is useless, if it is not understandable". Nel tentativo di trovare una soluzione a questo dilemma, Crystal (2001: 36-39) propone di sviluppare e diffondere l'uso degli *emoticon* utilizzati nelle *chat* e nelle *e-mail* per rendere di più immediata comprensione un sentimento o un'emozione. All'interno di un progetto portato avanti in Portogallo con numerosi collaboratori sordi, Neves (2005: 225-231) ha approfondito ulteriormente la

questione sperimentando l'uso di *emoticon* nella sottotitolazione tramite *teletext* della telenovela *Mulheres Apaixonadas*. Dopo vari tentativi, è giunta a redigere una lista di otto *emoticon* per veicolare le emozioni più diffuse che è stata pubblicata nella pagina del *teletext* portoghese (2005: 230):

:) for "happy";
:(for "sad";
:-/ for "angry";
:-s for "surprise";
:-& for "confusion";
;-) for "irony";
:-o for loud speech/screaming;
:-° for "soft speech/whispering;

Civera (2005) oltre a ribadire la necessità di non lasciarsi frenare dalla 'sacralità' del mezzo scritto e quindi di introdurre gli *emoticon* come ottima soluzione allo spreco di spazio necessario per introdurre le didascalie esplicative, ha addirittura proposto, nei supporti digitali, l'uso di *smiley* e di altra iconografia utile alla comprensione immediata di un determinato evento. Anche se quest'ultima soluzione non è ancora stata adottata in maniera diffusa, le soluzioni appena descritte sono adottate in molti paesi europei "even if sparsely" (Neves 2005: 222).

Componente extra-linguistica

Quanto alla componente extra-linguistica, come una canzone romantica, l'ululato di un lupo, lo scricchiolio di una porta che si apre o una musica che si fa sempre più insistente, essa contribuisce alla creazione del giusto contesto all'interno del quale si sta per svolgere la scena che si sta guardando. In tutti i casi sopraccitati, l'effetto viene garantito proprio dalla memoria enciclopedica di ogni singolo spettatore che riconoscerà l'effetto sonoro in questione e reagirà predisponendo il proprio stato d'animo in base al genere filmico in cui l'effetto sonoro è utilizzato. Come può il sottotitolatore adempiere al suo compito in simili situazioni visto che "(e)ven the most accurate representation of a sound is likely not be as evocative as the sound itself" (De Linde e Kay 1999: 14)? Baker *et al.* optano comunque per una didascalia esplicativa giustificando la loro scelte spiegando che "though

BLOODCURLING SCREAM

will not curdle the blood, the viewer at least knows the intensity of the sound that SCREAM alone would not convey" (1984 *op. cit.* in De Linde e Kay 1999: 14). A tal proposito, Neves (2005: 244-245) suggerisce:

> If the simple description of sound is sufficient to convey the intended effect, then the best option will be to simply indicate the presence of sound (e.g. ambulance siren); however, if further to the siren, other sound effects interrelate to build an atmosphere, then it may be more economical and relevant to describe the resulting effect (e.g. [tension mounts])

Come già accennato da Neves, per quanto riguarda effetti sonori che intendono semplicemente indicare la presenza di qualcuno o qualcosa o giustificare un'azione sullo schermo, la soluzione più semplice è quella di indicare tra parentesi o in altro colore l'effetto sonoro qualora non sia intuibile dalle immagini. Nel caso contrario, l'unica soluzione da ammettere è l'omissione. Sottotitolare che un cane abbaia se il muso è inquadrato in primo piano costituirebbe una grave mancanza di considerazione per l'intelligenza degli utenti finali. A tal proposito, "don't patronise us, please" è l'appello più volte lanciato da associazioni in difesa dei diritti dei sordi e da ricercatori sordi[83] in riferimento all'atteggiamento paternalistico delle emittenti che optano per una talvolta esagerata semplificazione del TP.

Simili difficoltà sono riscontrabili nella trasposizione della musica. Prima di iniziarne l'analisi è però forse necessario fare un distinguo tra due diversi tipi semiotici di musica: quella le cui parole sono importanti ai fini della comprensione del testo, come nei casi dei *musical* o dell'opera lirica, e quella che serve da accompagnamento o sottofondo a un'azione in corso. Mentre per il primo tipo di musica la soluzione comunemente adottata è di trascrivere (o tradurre) per intero il testo preceduto e seguito dal simbolo #, per il secondo tipo, malgrado sia difficile immaginare che l'effetto ottenuto in un utente sordo sia lo stesso di quello di un utente normoudente, la soluzione è quella adottata nel caso di effetti sonori: evitare ogni tipo di sottotitolo se inutile alla comprensione del testo o se interferirebbe con la sottotitolazione di parti importanti del TP; in caso contrario, scrivere tra parentesi o con diversi colori musica di sottofondo, oppure il titolo, l'autore ed eventualmente il genere della musica in questione.

83 Cfr. Donaldson 2004.

Altri aspetti

Altri aspetti della sottotitolazione per sordi che esulano dal binomio verbale/non-verbale riguardano l'identificazione del parlante, la sincronizzazione con le immagini e l'uso dei *font*. Quando si sottotitola per udenti, l'identificazione di un parlante non viene tenuta in considerazione perché si ritiene giustamente ridondante, quindi superflua. Anche se sta guardando una pellicola in lingua straniera, un udente riesce a capire chi sta parlando, se chi parla è fuori scena o se la voce proviene dalla radio o dal telefono grazie alla sua esperienza del mondo e alle informazioni sonore che la voce fornisce (tono, volume, timbro, prosodia, accento, mezzo di comunicazione, ecc.). Questo però non è il caso di un pubblico non-udente che troverà problematico eseguire questo compito nei casi in cui:

> characters are talking off screen;
> a narrator is speaking;
> a group of people are talking;
> there are unknown off-screen voices;
> characters are moving on screen, or a group of people are talking against a shot change. (De Linde e Kay 1999: 14).

In questi casi le soluzioni adottate sono numerose. Si possono indicare il nome di chi parla, la fonte della voce, che le persone stanno parlando allo stesso tempo, muovere i sottotitoli in maniera da farli comparire al di sotto della persona che sta parlando, utilizzare colori diversi o un trattino prima di ogni frase pronunciata da persone diverse, ricorrere all'uso di *add-on*[84] nei cambi di scena, al simbolo >> per dire che qualcuno sta parlando fuori scena, ecc. Sebbene queste soluzioni offrano la possibilità di individuare il parlante, possono esserci situazioni in cui è impossibile trovare un'opzione che riesca a compensare la mancanza di udito dello spettatore. Nel caso di più persone che parlano allo stesso tempo e che dicono cose diverse, indicare chi sta dicendo cosa diventa un'impresa impossibile per il sottotitolatore ed eventualmente di impossibile decodifica per l'utente. Una situazione particolare è quella in cui una persona sta muovendo le labbra ma in realtà non sta dicendo niente. Vedere delle bocche silenziose muoversi causa sempre una certa frustrazione a meno che la situazione non venga disambiguata. Neves (2005: 242) ricorda che alcune *guidelines* chiedono ai sottotitolatori di fornire spiegazione anche in questi casi, ma l'applicazione nella

84 Porzioni di sottotitoli che vengono aggiunti a sottotitoli già presenti sullo schermo per completare una frase che viene a trovarsi a cavallo tra due scene.

realtà non sembra essere regolare. Una soluzione a tutti questi problemi può essere offerta ancora una volta dalla tecnologia digitale. Con sottotitoli digitali sarà infatti possibile fare uso di una tecnologia più flessibile in grado di assolvere a maggiori funzioni.

Quanto alla sincronizzazione, mentre per il doppiaggio sembra essere una condizione indispensabile alla qualità del prodotto finale, nella sottotitolazione per udenti si tende a far coincidere la comparsa e la scomparsa del sottotitolo con l'inizio e la fine della battuta a cui si riferisce, lasciandolo sullo schermo nel rispetto dei limiti massimi di esposizione per evitare che venga riletto una seconda volta. Quando si sottotitola per non-udenti, invece, il garante per la televisione britannica, Ofcom, propone di lasciare sullo schermo i sottottitoli "for a sufficient time for viewers to read them" (ITC 1999: 11) vale a dire due secondi per riga fino a un massimo di 2,5 secondi qualora si faccia uso di *add-on*. In generale, a seconda della tradizione nel paese in cui e per cui viene svolto il sottotitolaggio, la tendenza è di lasciare i sottotitoli per più tempo quando si sottotitola per sordi rispetto a quanto avviene per gli udenti. L'importante è che si rispetti il principio della riduzione della frustrazione "caused to hearing-impaired viewers by being faced with silent moving mouths" (*ibidem*).

3.5 La leggibilità dei sottotitoli

La leggibilità dei sottotitoli, nel senso espresso dal concetto della *legibility*, è una componente fondamentale al fine di una piena accessibilità del testo audiovisivo. Ciononostante, Neves (2005: 186) sottolinea come alcuni televisori non permettano ancora una corretta visualizzazione dei sottotitoli proiettati tramite *teletext*, "for screens often have poor resolution and letter contours are not always sharp and clear. Moreover, viewers don't always sit at the most adequate distance from the television screen so as to achieve optimal viewing conditions". Se, da una parte, la tecnologia attualmente disponibile già permette una migliore qualità dell'immagine, resta da risolvere la questione della distanza e della posizione da cui si guarda lo schermo. Secondo Ivarsson e Carroll (1998), la distanza dalla quale si dovrebbero guardare le immagini per non sovraccaricare l'occhio è del triplo rispetto all'altezza dello schermo. Quanto alla posizione, lo spettatore dovrebbe sedere di fronte allo schermo onde evitare possibili distorsioni dell'immagine.

Oltre allo spettatore e ai produttori di televisori, anche l'emittente ha una responsabilità, forse la maggiore, nella leggibilità dei sottotitoli. In particolare, sviluppando

ulteriormente quanto già affrontato nel paragrafo precedente, ogni professionista che produce sottotitoli per il piccolo (o il grande) schermo, soprattutto se la proiezione avviene in analogico, dovrà tenere in considerazione due componenti fondamentali: la visualizzazione del carattere e quella del sottotitolo.

Per quanto riguarda la visualizzazione del carattere, quattro aspetti devono essere presi in considerazione: tipo, stile, dimensione e punteggiatura. Il tipo del carattere dipende fortemente dalle tradizioni di ogni paese. Sebbene i sottotitoli inter-linguistici per udenti tendano a utilizzare un tipo di carattere *Sans Serif*, il tipo dei caratteri del *teletext* o dei *closed captions* varia da paese a paese. Nel progetto di ricerca da loro condotto, Silver *et al.* (2000) hanno tentato di diffondere *Tiresias*, proponendolo come tipo di carattere standard. Come loro stessi hanno osservato, l'accettazione di questo carattere dipende dalla familiarità dello spettatore con lo stesso e che l'uso di un qualsiasi altro *font* diverso da quello usuale necessiterebbe un po' di tempo prima di essere accettato.

Lo stile del carattere invece non presenta queste difficoltà. Vista la limitata gamma delle opzioni e considerata la tradizione della lingua scritta, la maggior parte dei *teletext* proiettano i sottotitoli in stile normale minuscolo anche se talvolta il maiuscolo può essere utilizzato per la sottotitolazione dei telegiornali (cfr. Televideo RAI), per indicare un aumento del volume della voce dell'oratore, per sottotitolare le canzoni o per segnalare elementi extra-linguistici tra parentesi. Quanto al corsivo, può essere utilizzato in maniera concorrente alle maiuscole per sottotitolare le canzoni, o per indicare sorgenti acustiche particolari (una voce che viene dal telefono, dall'aldilà, dalla radio, dall'interno di un baule, ecc.) o che l'oratore sta parlando in una lingua straniera. Tuttavia, così come il grassetto o il sottolineato, il corsivo è raramente impiegato perché la sua lettura non risulta essere immediata.

La dimensione del carattere è un aspetto fondamentale. Se i vincoli spazio-temporali impongono al sottotitolatore di condensare enunciati troppo lunghi entro i 30-40 caratteri per riga, sarà necessario che la dimensione del carattere permetta al tipo di rendere ben visibile il carattere utilizzato compatibilmente con queste restrizioni. Un altro ruolo svolto dalle dimensioni del carattere è quello della spaziatura. Più piccolo è il carattere, maggiore sarà la difficoltà di distinguere non solo ogni singolo carattere, ma anche gli spazi tra le parole. A tal proposito, un curioso episodio viene riportato da Neves (2005: 189) che dimostra quanto la spaziatura tra le parole, e quindi le dimensioni del carattere, siano

importanti per una lettura ottimale dei sottotitoli. Durante la proiezione di una puntata della telenovela *Mulheres Apaixonadas*:

> while tidying up a bedroom, a character asked another to pass her a handkerchief. The subtitle on screen read [Passa-me o lenço!]. As the second character passed her a handkerchief, two Deaf viewers reacted with surprise. […] Even though the television set in use showed very clear subtitles, these viewers had confused [lenço!][85] with [lençol][86].

Da quanto riportato risulta chiaro come anche la punteggiatura concorra alla leggibilità dei sottotitoli. Solitamente, compatibilmente con la struttura delle frasi all'interno di ogni sottotitolo, l'uso della punteggiatura è ridotto al minimo. In particolare, i segni più frequenti sono:

- il punto, che serve a segnalare la fine di ogni frase, anche se generalmente ogni sottotitolo viene progettato con l'obiettivo di riportare un intero enunciato;
- i punti di sospensione, che servono a indicare che la frase contenuta nel sottotitolo non è ancora compiuta e che continua nel sottotitolo successivo. Un uso particolare dei punti di sospensione è fatto dalla RAI che ne utilizza due prima di ogni sottotitolo che contiene la continuazione della frase interrotta nel sottottitolo precedente;
- la virgola, che indica la suddivisione concettuale all'interno di ogni sottotitolo. Tuttavia, nei casi in cui si opta per una linearità sintattica, raramente un sottotitolo conterrà una frase incidentale e l'uso della virgola sarà limitato ad alcuni avverbi e preposizioni e alle liste. In questo caso un'eventuale condensazione lessicale comporterebbe nondimeno una riduzione del numero degli elementi di una lista rendendo pertanto inutile il ricorso alla virgola;
- il punto esclamativo, che può indicare un ordine, rabbia o altri intenti;
- il punto interrogativo, che è generalmente utilizzato per le domande. Insieme al punto esclamativo può essere utilizzato per indicare stupore, ironia, ambiguità o altro;
- le parentesi, che servono esclusivamente a contenere una didascalia esplicativa;

85 In portoghese significa fazzoletto.
86 In portoghese significa lenzuolo.

- i due punti, che vengono utilizzati talvolta per indicare che le parole che seguono fanno riferimento al nome che precede (es: […] me lo chiede proprio stasera? VESPA: certo […]);
- il trattino indica un nuovo oratore (es: -chi è lei? -il tenente Stone);
- il punto e virgola e le virgolette singole e doppie non vengono quasi mai usate perché non abbastanza utili. Tuttavia, le virgolette doppie sono talvolta utilizzate per indicare un uso non standard di una parola data.

Quanto alla visualizzazione dell'intero sottotitolo, due sono i parametri che più devono essere considerati dall'emittente televisiva: il colore dei caratteri e dello sfondo del sottotitolo e la tecnica di proiezione. Per rendere gradevoli e leggibili i sottotitoli, numerosi sono stati gli studi sull'uso dei colori da parte delle emittenti (BBC 1998, RAI 2002, AENOR 2003) e di eminenti ricercatori (Baker *et al.* 1984, De Linde e Kay 1999, Silver *et al.* 2000, Neves 2005). Secondo Baker *et al.* (1984: 9) esistono due ragioni principali per utilizzare colori diversi all'intero di uno stesso programma: "for emphasis and phrasing, and sound effects" e per identificare i singoli personaggi. Per quanto riguarda il primo aspetto, Neves (2005: 193-194) fa notare che "most broadcasters show preference for using […] other colour combinations [than coloured letters over black background] for comments or information about sound effects". Tuttavia, "[n]ot very many television channels use different colours within the same subtitle for emphasis", mentre "[c]reative usage of colour might be seen, for instance, in the subtitling of songs, where the karaoke technique helps people to keep up with the singing". Risulta quindi chiaro che, *de facto*, i colori sono utilizzati principalmente per identificare i diversi personaggi sullo schermo e che altri usi sono limitati a situazioni puntuali.

Quanto all'identificazione del parlante, Baker *et al.* (1984: 8-9) affermano che il *teletext* permette di utilizzare sette colori diversi per i caratteri e otto per lo sfondo. Secondo le loro ricerche, la migliore combinazione di colori risulta essere il carattere bianco su sfondo nero. In alternativa, si possono usare in ordine di leggibilità il giallo, il ciano e il verde su schermo nero. I colori meno visibili sono il magenta, il rosso e il blu. Quanto all'uso di altri colori, Neves (2005: 195) afferma chiaramente che soprattutto "in very colourful films, coloured subtitles may be problematic for they will not have sufficient contrast to guarantee adequate legibility". Un altro problema risulta essere il numero dei

personaggi sullo schermo. Come afferma sempre Neves (2005: 195), l'uso di colori diversi per identificare personaggi diversi secondo il criterio un colore-un personaggio per tutto il corso del film è

> cause for confusion rather than an asset. This [is] particularly the case when, due to the high number of characters on screen, it [becomes] necessary to switch from the pre-established colour codes to distinguish between two characters to whom the same colour had been attributed.

Vista la vasta gamma di linee guida diverse, una ricerca più approfondita risulta quindi necessaria sulla praticabilità, a seconda dei programmi, dell'uso di diversi colori per differenziare i singoli turni, prima ancora che sulla maggiore o minore visibilità di un colore su un determinato sfondo piuttosto che un'altra combinazione.

Per quanto riguarda il secondo aspetto importante della visualizzazione del sottotitolo è necessario forse distinguere la modalità di proiezione dei sottotitoli dalla loro presentazione sullo schermo. Visto che la modalità di proiezione è già stata ampiamente discussa, sarà oggetto di discussione soltanto la presentazione dei sottotitoli all'interno del testo audiovisivo. A tal proposito, Neves (2005: 201) propone di distinguere tra il numero di righe, il posizionamento del sottotitolo e l'allineamento. La scelta tra sottotitoli di due o una riga[87] dipende chiaramente dalla tipologia di proiezione dei sottotitoli, dal ritmo del testo audiovisivo da sottotitolare, dalla necessità o meno di sincronizzare il sottotitolo con le battute e soprattutto dalla quantità di testo da sottotitolare. A sua volta, questi ultimi tre aspetti dipenderanno dal genere da sottotitolare. A seconda del peso di ognuna di queste componenti all'interno del testo, la scelta per una o due righe risulta evidente. Nel caso di sottotitoli *roll-up* risulta impossibile la proiezione di due righe, mentre i *pop-on* possono contenere sottotitoli di una o due righe. Quanto al genere del testo, a generi diversi corrispondono testi diversi, in cui la componente verbale conta più o meno a seconda della ridondanza determinata dalla componente video. Ad esempio, il telegiornale, che ha un ritmo superiore alle 150 parole inglesi al minuto, in cui la componente verbale interagisce in maniera referenziale e compensativa con le immagini, e il cui scopo è informare il più possibile nel minor tempo possibile, necessita chiaramente di una sottotitolazione che corrisponda al suo scopo. In questo caso, due righe piene sembrano essere inevitabili. Al

87 Cfr. D'Ydewalle (1991), Jensema (1999) e Neves (2005: 201).

contrario, un documentario, in cui l'oratore non compare sullo schermo e ha un eloquio sicuramente più lento rispetto al TG, in cui la componente verbale descrive le immagini e il cui scopo è di accompagnare lo spettatore all'interno delle immagini fornendo informazioni talvolta ridondanti, la sottotitolazione può essere anche su una riga.

Quanto alla posizione dei sottotitoli sullo schermo, le opzioni sono molteplici e il loro uso è dettato ancora una volta dalla componente video del testo audiovisivo di partenza. Se, convenzionalmente, i sottotitoli compaiono centrati in basso allo schermo, dove le immagini sono meno importanti rispetto al resto dello schermo, è anche possibile trovare sottotitoli che, per le più svariare ragioni, si possono trovare allineati a sinistra o a destra e in alto allo schermo o a metà. Quando in un telegiornale compare il nome e il titolo della persona intervistata, obbligo del sottotitolatore sarà veicolare questa informazione al pubblico a casa. Se compensare nel sottotitolo la perdita dovuta alla sovrapposizione del sottotitolo su questa informazione risulta essere molto dispendioso in termini di economia testuale, più facile risulterà muovere il sottotitolo, contenente la restituzione dell'intervista, in maniera da evitare che si sovrapponga all'informazione data sullo schermo e da non perdere tempo e spazio nel veicolare l'informazione sull'identità dell'oratore. Un altro esempio potrebbe essere la presenza in basso allo schermo di immagini importanti alla comprensione generale del testo. Anche in questo caso, muovere il sottotitolo in cima allo schermo, piuttosto che in mezzo, sembra essere la soluzione migliore. Un ultimo esempio proviene dall'esigenza, nei film, di identificare la fonte di dialoghi o di effetti sonori. Proprio per veicolare questo tipo di informazioni alcune emittenti potrebbero optare per un allineamento a destra o a sinistra a seconda che l'oratore o la fonte sonora si trovi a destra o a sinistra.

Un altro aspetto di sicuro interesse per il sottotitolatore è la distribuzione del testo sulle due righe. Se di solito la prima riga è più corta della seconda "so as to leave the important visual information as visibile as possible" (Neves 2005: 203), c'è anche da considerare che per ragioni sintattiche non è possibile separare i sintagmi. Qualora risultasse più lungo il sintagma nominale rispetto a quello verbale e visto che un'eventuale riformulazione sintattica porta sovente a un ordine dei componenti frastici di tipo lineare (soggetto-verbo-oggetto), allora la seconda riga sarà necessariamente più breve rispetto alla prima. Per concludere con questo aspetto, Ivarsson e Carroll (1998: 77 *op. cit.* in Neves 2005) sostengono che:

> (i)f the subtitles are centred, it makes no difference which of the two lines is the longer or shorter, since the distance of the eye to travel is always the same, *i.e.* half the length of both lines […]. But when the lines are left-justified it is quite a different story […]. Here it is obvious that the upper line should be the shorter of the two.

Come affermato precedentemente, ancora una volta risulta evidente che il metro più vincolante resta l'accessibilità. Per una maggiore consapevolezza dell'importanza di questo aspetto, si cercherà nel prossimo paragrafo di affrontare la questione dell'accessibilità in maniera più dettagliata.

3.6 La qualità dei sottotitoli per sordi

Un aspetto fondamentale della traduzione audiovisiva in generale su cui si sta lavorando da anni sia nel mondo accademico[88], sia in quello dei *broadcaster*[89] è il concetto di qualità del trasferimento linguistico. Dopo aver analizzato a fondo le norme sottostanti la traduzione in generale e la traduzione audiovisiva in particolare, Gambier (2003) individua sei standard di qualità che ogni tipologia di traduzione audiovisiva deve rispettare per poter svolgere appieno la sua funzione inclusiva. Queste sono:

- *acceptability:* ogni testo deve essere 'accettabile' dal punto di vista grammaticale e più in generale dal punto di vista linguistico (terminologia, stile, registro, ecc.);
- *legibility:* deve poter essere possibile per lo spettatore 'leggere' il testo, in termini sostanzialmente di definizione dei grafemi o dei fonemi e di velocità e praticità della fruizione del testo;
- *readability:* nei limiti della fedeltà al TP e del rispetto della forza allocutiva del testo audiovisivo di partenza, il TA deve presentare una semplicità di fruizione in termini di carico semantico, coerenza e coesione testuale e densità delle informazioni;
- *synchronicity:* le componenti verbali e quelle non verbali, siano esse audio o video, devono poter entrare in sintonia tra di loro in modo da rispettare l'equilibrio del TP. Per le tipologie traduttive che implicano il canale acustico (doppiaggio, *voice over*, commento, ecc.), significa rispettare le pause e laddove possibile il movimento delle labbra dell'originale. Nel caso della sottotitolazione e affini, invece, questo

88 Cfr. Subtitle Project www.subtitleproject.net
89 Cfr. Ofcom www.ofcom.org.uk

concetto è vincolato ad altri aspetti come il rispetto dei cambi scena e la coerenza con la componente video non verbale;

- *relevance:* ogni tipo di traduzione audiovisiva, in particolar modo quelle tipologie che richiedono un utilizzo diverso dei canali di trasmissione del testo rispetto a quello di partenza, deve tenere in considerazione il carico cognitivo dell'utente finale selezionando l'informazione più pertinente da trasmettere;
- *domestication strategies:* come avviene in maniera ancor più dettagliata nella localizzazione, ogni testo audiovisivo deve rispettare una serie di aspetti culturali (valori, pregiudizi, comportamenti, ecc.) che variano a seconda delle società. Queste ultime possono non recepirli alla stesa maniera del pubblico per cui è pensato il TP. Compito del sottotitolatore sarà quindi di adeguare il TP alla cultura di arrivo, intervenendo anche sostanzialmente laddove ritiene necessario.

In riferimento proprio a quest'ultima componente socio-linguistica, il pubblico a cui il TP è destinato, Neves (2005: 121-131) sottolinea che nel caso della sottotitolazione per sordi c'è da aggiungere un altro criterio fondamentale, estendibile a tutta la gamma di traduzioni per scopi speciali: "*adequacy* to the special needs of Deaf and hard-of-hearing receivers". La questione ulteriore che si pone con quest'ultimo criterio di qualità è quindi quella inerente la modalità con cui si rendono accessibili accettabilità, leggibilità (nei due sensi di cui sopra), sincronia e pertinenza del TA e 'addomesticamento' del TP in modo da soddisfare le legittime esigenze degli utenti finali. A tale proposito, Nord (2000: 195) sottolinea come nell'ambito di ricerca sulla linguistica testuale e discorsiva siano rari gli studi che si concentrano

> on the assumption that the addressee or rather: the idea of the addressee the author has in mind, is a very important (if not the most important) criterion guiding the writer's stylistic or linguistic decisions. If a text is to be functional for a certain person or group of persons, it has to be tailored to their needs and expectations.

Tuttavia, nonostante sia vero che lo stesso Gambier (1998), oltre che Gottlieb (1997) e più recentemente Díaz-Cintas (2004) abbiano segnalato questa esigenza di produrre un testo che sia veramente adeguato alle esigenze dell'utenza finale, Neves (2005: 122) sottolinea che "not much empirical research has been carried out to provide reliable data that

might shed light on the profile of actual receivers" e che se studi sull'accessibilità vengono effettuati, questi "derive from marketing efforts that aim at characterising audiences for the sake of shares and advertisement campaigns". Per di più, "such data rarely feeds into other departments, such as those where subtitling is provided".

Ecco quindi che risulta indispensabile una conoscenza approfondita da parte di chi produce i sottotitoli per i sordi delle reali esigenze di questa tipologia di utenti nell'accedere al prodotto audiovisivo. Per ovviare a questa necessità c'è bisogno di una ricerca approfondita in questo settore in ogni paese. È chiaro infatti che, a seconda dei paesi, ogni comunità di sordi ha esigenze diverse indipendenti dall'ovvia constatazione che, per questioni legate al grado e alla tipologia di sordità, all'età, all'istruzione, alla lingua madre (lingua dei segni o lingua verbale) e all'esposizione alla lingua parlata nel paese d'appartenenza, ogni singola persona sorda ha le sue esigenze specifiche e che, proprio in quest'ottica, [a]n "elastic" text intended to fit all receivers and all sorts of purposes is bound to be equally unfit for any of them, and a specific purpose is best achieved by a text specifically designed for this occasion" (Nord 2000: 195). Si possono infatti riscontrare, per ogni paese, delle tendenze generali che possono accomunare le persone sorde nel loro insieme. I fattori che determinano questo profilo 'nazionale' dell'utente medio sordo sono legati all'abitudine audiovisiva preponderante nel paese in cui vivono (sottotitolazione vs. doppiaggio), all'attenzione riservata da parte del sistema sociale ai cittadini sordi (sistemi volti all'accessibilità dei sordi ai servizi più prettamente uditivi – oltre alla televisione, il telefono, le comunicazioni per altoparlante come gli annunci delle ferrovie e simili, ecc. –, la fornitura del servizio di interpretazione in lingua dei segni per i sordi segnanti, istruzione mirata o indifferenziata, copertura dei costi medici per la prevenzione della sordità o per il recupero dell'udito, ecc.) e, conseguentemente, all'integrazione delle comunità o dei singoli sordi nella società in cui vivono.

Sulla questione dell'esigenza di una conoscenza approfondita da parte dei sottotitolatori delle necessità e delle aspettative dell'utenza sorda, torna con insistenza e convinzione Neves (2004):

> if we are to consider subtitling as "translational action" (Vermeer, 1989: 221), serving a functional end, its *skopos* needs to be perfectly understood by all those involved in the commission. Quite often, the commissioners of SDH, and the subtitlers themselves, are not completely

> aware of the particular needs of their "clients" for not much is given to them in terms of audience design or reception analysis. In fact, only by knowing the distinctive features of the target audience will people be reasonably aware of the possible effects of their work on their receptor. Only then can anyone aim at the utopic situation where the "new viewer's experience of the programme will differ as little as possible from that of the original audience (Luyken, 1991: 29)".

Nel tentativo di offrire un valido contributo alla questione che si sta affrontando, Gambier (2003a: 185) cita i quattro fattori che Kovačič (1995: 376) individua per la creazione di un modello di riferimento volto alla ricerca nell'ambito dei *reception studies* applicati alla sottotitolazione audiovisiva e cioè:

- il contesto socio-culturale in generale che influenza la ricezione dei sottotitoli;
- la preferenza dei telespettatori per il doppiaggio o la sottotitolazione;
- le strategie percettive che ogni spettatore mette in atto nel decodificare tramite la vista un prodotto audiovisivo sottotitolato;
- l'impatto psicologico che l'ambiente cognitivo ha sulla comprensione dei sottotitoli.

Secondo Gambier questi fattori possono essere utilizzati anche e soprattutto quando si sottotitola per una porzione di pubblico che ha delle caratteristiche precise e che pertanto ha delle esigenze specifiche, difficili da comprendere se non si condivide o addirittura si ignora la questione della sordità[90]. Tuttavia, un lavoro di questo genere risulta inutile se, nell'era della globalizzazione, non si opta, a livello internazionale, per una standardizzazione delle tecniche e delle prassi di produzione e proiezione dei sottotitoli. Nel prossimo paragrafo si cercherà di affrontare la questione della standardizzazione sia in termini generici, sia riportando un caso specifico, che illustra bene le tappe da seguire nel tentativo di raggiungere questo obiettivo.

3.5 La standardizzazione

90 Nell'ambito del workshop "Live Subtitling. How to respeak and for which audience? Theory and practice", tenutosi a Berlino il 25 ottobre 2006 (http://www.languages-media.com/lang_media_2006/programme.php ultimo accesso 28 gennaio 2007), alla domanda iniziale "quanti di voi hanno avuto a che fare con o si sono interessati alla sordità?" solo quattro sottotitolatori per sordi su tredici hanno risposto in maniera affermativa seppur con gradi diversi di conoscenza.

Nonostante in Europa ci siano stati dei tentativi da parte del CENELEC e di altri organismi internazionali di standardizzare la pratica della sottotitolazione per sordi, le emittenti dei vari paesi sono restie a cambiare le proprie abitudini nel settore. Quelli che vengono considerati i quattro strumenti per garantire accessibilità, nel senso largo del termine, ai media (doppiaggio, sottotitolaggio per udenti, sottotitolaggio per sordi e audiodescrizioni) non sono ancora oggetto di una vera e propria normativa tecnica internazionale. Secondo Clark (2006), nonostante ci siano delle leggi sull'accessibilità emanate dai garanti per le radiotelevisioni[91], ci siano stati dei tentativi di scambio di *file* tra emittenti[92], siano stati pubblicati manuali volti agli operatori del settore da parte di alcune emittenti[93], alcuni ricercatori[94] e alcune agenzie[95], persiste una diffusa mancanza di 'collaborazione' tra le varie emittenti mondiali che comporta una inaccessibilità dei vari prodotti audiovisivi agli spettatori oltre che dei costi aggiuntivi di programmazione. Per tutti i paesi anglofoni, infatti, ma anche all'interno di ogni singolo paese tra le varie emittenti che compongono il panorama televisivo, uno scambio di *file* risulterebbe essere non solo più pratico ed economico per la sottotitolazione dei programmi di cui si ricevono i sottotitoli già preparati da altre agenzie, ma permetterebbe a ogni singola emittente di concentrarsi sulla sottotitolazione di programmi che non sono ancora stati sottotitolati, alimentando così lo scambio di *file* e garantendo ai propri spettatori un servizio più vasto. Di conseguenza, le quattro tipologie di accessibilità appena menzionate variano in termini tecnici a seconda della lingua, del paese, della tecnologia utilizzata e soprattutto a seconda dei clienti e dei committenti. Come fa notare Robson (1997: 48-51), all'interno dei paesi che sfruttano la tecnologia NTSC, i *closed captions* sono diversi tra agenzie concorrenti. Inoltre, in Canada, i sottotitoli in francese differiscono da quelli in inglese a causa della "French-language preference for mixed-case typography and the use of the accented-character set available in Line 21". Clark (2006) invece fa notare come "[s]ome clients demand only the cheapest captioning available, while other clients use higher-cost captioning, up to and including

91 Cfr. RCQ 1983, CRTC 1995, FCC 1997, Australia 1998 (dove però la parola 'standards' non riguarda la qualità ma la quantità) e HMSO 2003. In nessuno dei casi citati si richiede standardizzazione ma solo accessibilità in termini generali.
92 Cfr. W3C 2003.
93 Cfr. ITC 1999, 2001, 2002, CAB 2003 e Auscap 1999.
94 Cfr. Verlinde e Schragle 1986 e CFV 1996.
95 Cfr. Carlson et al. 1990 e Dittman et al. 1989.

multiple sets of captions on a single program (as with near-verbatim and easy-reader versions, or English- and Spanish-language captions)".

La necessità di una standardizzazione è quindi quanto mai necessaria. Prima di addentrarci nel merito della questione è forse necessario fare una distinzione tra standardizzazione tecnica e standardizzazione delle prassi. Come si è visto, la standardizzazione tecnica è una fattore abbastanza diffuso, almeno all'interno delle zone che utilizzano lo stesso sistema di diffusione dei dati televisivi[96]. In generale, si può affermare che, a parte le regole interne a ogni singola agenzia o emittente, alcuni criteri sono trasversalmente applicabili (ma non sempre applicati) a tutti i sistemi come ad esempio il numero massimo consentito di righe e di caratteri per riga, il tipo, lo stile, la dimensione e il colore dei caratteri, il colore dello sfondo, ecc. La standardizzazione delle prassi, invece, merita un maggiore approfondimento proprio alla luce di quanto sopra affermato.

Innanzitutto, non esiste una certificazione ufficiale e universalmente riconosciuta che attesti la professionalità degli operatori nel settore, visto che la formazione avviene per la maggiore all'interno delle singole aziende e in qualche raro istituto universitario (tra gli altri *University of Surrey Roehampton*, *University of Wales*, *Copenhagen Business School*, *Universität Autonoma de Barcelona*, *Hogeschol van Antwerpen*, *Haute École de Bruxelles*, *Université du Mons-Hainaut, Università di Bologna*). Per quanto riguarda invece la formazione in media accessibili, il quadro deve essere allargato agli altri paesi anglofoni non europei. Diversamente da quanto accade nei paesi non anglofoni, all'interno della comunità scientifica degli Stati Uniti, dell'Australia e in minor misura del Regno Unito, non si pone il problema della sottotitolazione per udenti, visto che la maggior parte della produzione televisiva e cinematografica avviene in lingua inglese. Ecco quindi che i maggiori centri di ricerca di questi paesi sono volti proprio alla formazione nell'ambito dell'accessibilità ai media. Alcuni esempi oltre al *National Center for Accessible Media* sono *l'Adaptive Technology Resource Centre* della *University of Toronto*, il *Centre for Learning Technologies* della *Ryerson University*, il *Trace Center* della *University of Wisconsin*, il *Technology Assessment Program* della *Gallaudet University*, il *Centre for HCI Design* della *City University London* e il *World Wide Web Consortium*.

96 Cfr. EIA (2002) per la Line 21, EIA (2002) per la North American high-definition television, BBC et al. (1976) per il teletext, ETSI (2002) per la Digital Video Broadcasting e il DVD Forum (2003) per i sottotitoli nei DVD.

In materia di standardizzazione delle prassi in questo settore, sta lavorando il progetto internazionale con base a Toronto *The Open & Closed Project*[97]. Convinti che formare, certificare e avviare professionisti nel settore dell'accessibilità ai media tramite i metodi standardizzati sia vantaggioso oltre che per i garanti, le emittenti e i produttori anche per i professionisti e soprattutto per gli utenti finali, viene proposto il seguente modulo per la produzione di linee guida standard:

- *accessibility*: Make the specification accessible to people with disabilities and others;
- *availability*: Publish and distribute the specification as widely as possible;
- *contribution & collaboration*: Allow contribution and collaboration from all interested parties;
- *decisiveness*: Make decisions quickly based on evidence and fact;
- *representation*: Ensure that involved parties are drawn from a fair cross-section, without imbalance;
- *evidence and fact*: Rely on evidence and fact, even if somewhat in dispute, rather than opinion and feeling;
- *research*: Where evidence and fact are lacking or in dispute, commission custom research;
- *best trumps current*: With a basis of evidence and fact, prefer demonstrably best practices even if at variance with current practice;
- *localization*: Ensure that specifications function well, and are customized for, languages, cultures, audiences, and technologies;
- *testing*: Provide a beta period in which the specification can be torture-tested in the real world;
- *errata and revision*: Make it possible to correct errors in the specification, and develop a program to revise the spec as new technologies arise;
- *training and certification*: Provide a means of training and certifying practitioners.

La pubblicazione dei risultati del progetto saranno sicuramente un grosso passo in avanti verso il processo tanto auspicato della standardizzazione delle prassi e delle tecniche di produzione e proiezione dei sottotitoli per sordi.

3.6 Dalla traduzione alla sottotitolazione per sordi in ottica strategica

La produzione scientifica in materia di traduzione audiovisiva si è concentrata prima sul riconoscimento a pieno titolo della traduzione audiovisiva come disciplina

97 Per maggiori informazioni si veda http://openandclosed.org/

d'interesse accademico e sulla sua subordinazione o meno alla traduzione nel senso convenzionale. Negli ultimi tempi invece si è finalmente focalizzata sulle sue caratteristiche e sull'analisi dei prodotti audiovisivi come testi *per se*. Ecco quindi che si è iniziato a delineare i vincoli che il traduttore audiovisivo deve considerare e ad analizzare le strategie a cui può far ricorso nell'affrontare il suo lavoro. I primi documenti sui sottotitoli per sordi sono stati scritti da professionisti operanti nel settore audiovisivo (Ofcom, NCI, Rai-Televideo) o da associazioni in difesa dei diritti delle persone con problemi di udito (RNID, ENS, FIADDA) all'epoca della loro introduzione in TV. Altre pubblicazioni sono delle relazioni di *case-study* (tra i più importanti Kyle 1996, Gregory e Sancho-Aldridge 1997, Jensema 1999 e Gaell 1999) e affrontano la materia soprattutto dal punto di vista tecnico. Solo poche pubblicazioni hanno invece avuto un oggetto più squisitamente linguistico e traduttivo (Volterra 1981, 1988).

Riflesso di questo processo sono gli incontri della comunità scientifica internazionale, che, da quando il Consiglio d'Europa promosse, nel 1995, il forum sulla comunicazione audiovisiva e sul trasferimento linguistico in occasione del centesimo anniversario della nascita del cinema, hanno iniziato ad affrontare il tema degli studi sulla traduzione audiovisiva. Tuttavia, già McAdam 1985, Baker 1986 e Hindmarsch 1986 erano coscienti dell'esistenza della sottotitolazione per sordi come una delle tante applicazioni della tecnica della sottotitolazione (Luyken *et al.* 1991). Una delle prime attestazioni della sottotitolazione intra-linguistica per sordi riconosciuta come disciplina a pieno titolo all'interno della vasta famiglia della traduzione in generale e della traduzione audiovisiva in particolare la ritroviamo comunque solo qualche anno più tardi: in filigrana in Gottlieb (1992) e in Ivarrson (1992) e in maniera più dettagliata in Gambier (1994). Lo stesso Ivarrson, soltanto nel 1998 definirà la sottotitolazione per sordi come "a subject for a whole book" (Ivarsson e Carroll 1998: 129). Nello stesso anno, Gambier riconosce che la sottotitolazione per sordi possiede delle specificità talmente tanto definite da renderla una tipologia di traduzione *per se* e che merita di essere studiata con maggiore approfondimento. Solo successivamente, sebbene la prima conferenza sull'accessibilità della televisione ai non-udenti si svolse già nel 1971 negli Stati Uniti, si sono moltiplicate le conferenze dedicate anche solo parzialmente a questo tema (tra le più importanti *Scripta manent*, Roma 2001; *Languages and the Media*, Berlino 2002, 2004, 2006 e 2008; *In so many words*, Londra 2004; *TV digital y accesibilidad para personas discapacitadas en un entorno global*

de comunicación, Altea 2004; *Media for All*, Barcellona 2005 e Leiria 2007; *Mu.Tra. euroconferences*, Saarbrücken 2005, Copenhagen 2006 e Vienna 2007; *First International seminar on real-time intralingual subtitling*, Forlì 2006). Infine, tra i contributi più importanti allo studio della sottotitolazione per sordi, non possono non essere menzionati i lavori di De Linde e Kay (1999), di Franco e Araújo (2003) e di Neves (2005), che ha dedicato alla materia la prima tesi di dottorato al mondo.

Nel tentativo di ottenere un quadro di riferimento applicabile allo studio del rispeakeraggio come prodotto, si farà pertanto ricorso agli spunti che derivano sia dagli studi sulla traduzione in generale, sia dagli studi sulla traduzione audiovisiva. Per quanto riguarda i fattori inerenti il processo di redazione e traduzione, Holman e Boase-Beier (1999: 1-17) elencano:

- le caratteristiche linguistiche di ogni lingua (connotazioni di alcuni fonemi, simbolismo sonoro, convenzioni morfo-sintattiche, ordine delle parole, tradizione retorica relativa alla ripetizione, l'uso dei pronomi di prima persona, le forme di cortesia, ecc.). A tal proposito, Gambier (2006) fa notare come questi vincoli siano imposti da ogni lingua di partenza o di arrivo come "chaque matériau (verre, granit, acier, fonte, plastique) impose au sculpteur certaines formes plutôt que d'autres" (2006: 24);
- le convenzioni e le tradizioni dei generi testuali da tradurre e quanto insegnato dalla linguistica in generale e dalla pragmatica in particolare;
- la censura;
- i vincoli imposti dagli editori, committenti, clienti, revisori;
- la tradizione letteraria delle due culture coinvolte (delle lingue e dei testi di partenza e di arrivo);
- la professionalità del traduttore e il suo bagaglio culturale e intellettuale.

Se si osserva più da vicino la traduzione per sottotitoli filmici e cinematografici Gambier (1999) e Bartoll (2004) aggiungono a questa lista:

- i vincoli spazio-temporali (numero caratteri consentito, montaggio delle scene, velocità di eloquio degli attori, ecc.);

- le componenti testuali, paratestuali ed extratestuali che concorrono parallelamente alle competenze linguistiche e cognitive dei fruitori del prodotto finale all'informatività del testo.

Come fa notare Gambier (2006 : 24),

> "[d]ans cet ensemble, il est difficile d'imaginer un travail («original» ou traduit) sans contrôles, sans filtres, sans modèles à suivre et/ou à transgresser - qu'ils soient formulés, dictés par le client ou les récepteurs supposés, par des institutions ou des individus au pouvoir défini".

Ecco quindi che una volta definito il quadro all'interno del quale è vincolato il traduttore per sottotitoli, è necessario ottenere un quadro quanto mai esaustivo delle strategie utilizzate in traduzione nel processo di trasferimento linguistico dal TP a quello di arrivo. In quest'ottica, Gambier suggerisce di ricorrere alla letteratura traduttologica. Volendo considerare solo le ultime riflessioni in materia, Chesterman (1997) identifica una serie di dieci strategie sintattiche (traduzione letterale, prestito/calco, trasposizione, *unit shift*, *group shift*, *constituent shift*, *level shift*, sconvolgimento della coesione, della struttura della frase e del quadro retorico), dieci strategie semantiche (sinonimia, antonimia, iponimia, *converses*, aggiunta/compressione, parafrasi, alterazioni nei livelli di astrazione/concretizzazione, enfasi, stile, uso altri espedienti semantici) e dieci strategie pragmatiche (*cultural filtering*/prestito, esplicitazione/implicitazione, omissione/aggiunta, alterazione del grado di formalità, soggettività e oggettività, della forza illocutoria, della coerenza, traduzione parziale, note/commenti/glosse, riformulazione, uso altri espedienti pragmatici), che ogni traduttore applica nell'espletamento del proprio lavoro.

Con lo stesso obiettivo, Molina e Hurtado (2002 *op. cit. in* Gambier 2006) elencano quindici 'translation techniques' (adattamento, aggiunta, prestito, calco, compensazione, descrizione, alterazione discorsiva, equivalente lessicale o semantico, iponimia/iperonimia, espansione/compressione linguistica, traduzione letterale, modulazione, riduzione, sostituzione, trasposizione).

Cercando di tirare le somme da questa categorizzazione delle strategie (o tecniche) da attuare in traduzione, emerge una distinzione abbastanza evidente tra le strategie psico-cognitive che si concentrano sul processo traduttivo (la decodifica del TP, la valutazione

delle possibilità offerte, la costruzione del TA e la valutazione finale del TA) e quelle più squisitamente linguistiche (tecniche di trasferimento lessico-grammaticale, semantico e pragmatico).

Tra i primi a cercare di standardizzare le tecniche (o strategie) di sottotitolazione, Gottlieb (1992: 166-167) mette insieme gli insegnamenti degli studi sulla traduzione e propone dieci strategie ancora oggi molto valide e utilizzate. Esse sono:

- *expansion*: un elemento culturale viene descritto;
- *paraphrase*: un'espressione idiomatica o idiosincratica viene riformulata;
- *transfer*: traduzione letterale di un'espressione o idiosincrasia;
- *imitation*: nomi propri e altre peculiarità 'intraducibili' vengono riprodotte nella stessa forma dell'originale;
- *transcription*: espressioni non gramaticalizzate nella lingua di partenza sono rese con equivalenti non grammaticalizzati nella lingua di arrivo seguendo la stessa logica che ha portato alla deviazione linguistica iniziale (es: -you must be Igor; -No, it's pronounced Eye-gor. Reso in italiano: -Tu devi essere Igor; -No, si dice Ai-gor) (Perego 2005: 107);
- *dislocation*: quando è il ritmo a essere importante o la ripresa di una determinata parola, si può cambiare la forma di un'espressione dal contenuto non per forza identico;
- *decimation*: la forma del messaggio viene ridotta quantitativamente;
- *deletion*: omissione di espressioni scarsamente rilevanti;
- *resignation*: quando il TP è 'intraducibile' si cerca di sostituirlo con qualcosa di diverso sia nella forma che nel contenuto.

Per giungere al medesimo obiettivo, Lambert e Delabastita (1996) sono invece partiti da un modello tripartito da loro ideato (*competences*, *norms*, *performances*) che considera le tre componenti semiotiche del prodotto audiovisivo (*visual*, *acoustic*, *verbal*). Sulla base di questo modello, sono state individuate cinque macro-strategie, che riprendono la terminologia latina della retorica classica: *repetitio*, *transmutatio*, *adiectio*, *detractio* e *substitutio* (Lambert e Delabastita, 1996: 39-40).

In maniera del tutto simile a Lambert e Delabastita, Gambier cita (2006) anche i tentativi in materia di Ivarsson e Carroll (1998), Lomheim (1995), Kovačič (2000), Schwarz

(2003) e Tveit (2004). Dopo aver constatato la difficoltà da parte di tutti "à appréhender tous les éléments significatifs ensemble, à se détacher aussi du modèle écrit, si prégnant que la tentation est grande de réduire les sous-titres à un problème de comptabilité de mots, avec écart entre ce qui est énoncé et ce qui est écrit", Gambier (2006: 34) riconosce il carattere *target-oriented* della sottotitolazione, identifica alcuni tratti comuni, li raggruppa in macro-strategie concettuali e propone una suddivisione semplificata delle tattiche traduttive che ogni sottotitolatore mette in atto per raggiungere il proprio scopo:

- *réduction*: la caratteristica più tipica in sottotitolazione è la riduzione, generalmente quantitativa del TP. Questo strategia è attuabile tramite l'attuazione delle tecniche di *compression* (le occorrenze ridondanti a livello lessicale, morfo-sintattico o frastico vengono riassunte o compendiate secondo una gerarchia stabilita dal traduttore) e di *élimination* (se la componente video fa eco all'informazione verbale o in presenza di elementi superflui);
- *simplification*: viene operata a livello lessicale e sintattico attraverso parafrasi, paratassi, lessicalizzazione, iperonimia, e simili;
- *expansion*: le strategie di espansione più comuni sono l'esplicitazione, la parafrasi, il prestito diretto e l'equivalenza dinamica (250 dollari diventa 300 euro).

Un ultimo aspetto interessante riguarda la terminologia utilizzata. Gambier parla di strategie, mentre altri ricercatori preferiscono soluzioni meno incentrate sul traduttore come processo (Vinay e Darbelnet 1958) o procedura (Newmark 1988) in quanto, come aveva già fatto notare Lörscher, più che di scelta consapevole sembra più opportuno parlare, anche per la traduzione di testi scritti, di operazione quasi automatica, di meccanismo professionale (1991: 96). Questo è ancor più vero se attuato alla sottotitolazione in tempo reale, in cui i tempi di elaborazione del TP sono molto ristretti (il passaggio dalla percezione del TP alla produzione del TA varia dai tre ai dieci secondi) e sicuramente non concedono spazio a una scelta ponderata e misurata tra tutte le varie opzioni. Resta tuttavia una considerazione fondamentale che fa pendere l'ago della bilancia verso la definizione di Gambier: queste procedure e questi processi non sono norme prestabilite, ma sono il frutto dell'operato del sottotitolatore (o del traduttore) che ha appreso e in un secondo momento semi-automatizzato le tecniche di resa di un TP. Pertanto, quelle che sembrano essere solo in superficie degli automatismi sono in realtà delle proiezioni mentali di un lavoro svolto

intenzionalmente per raggiungere l'obiettivo prefissato. Da questo punto di vista, l'atto del tradurre può essere paragonato all'azione del camminare. Apparentemente, si tratta di un'azione automatica che accomuna quasi tutti gli esseri umani. Tuttavia, in presenza di altre azioni che richiedono uno sforzo supplementare, come bere un bicchiere d'acqua, indicare un punto fisso o mobile o ancora cercare di convincere qualcuno che cammina al proprio fianco dell'esattezza di un'idea o di baciarlo, la stessa azione diventa meno automatica e necessita di sforzi maggiori. Per non parlare della fase di apprendimento, che richiede al bambino o al paziente in riabilitazione uno sforzo continuo e prolungato oltre che tenacia nel superare i numerosi fallimenti.

Ora che si è visto l'insieme di strategie che permettono a un sottotitolatore di raggiungere il proprio obiettivo, è forse necessario introdurre gli ultimi due strumenti di analisi volti allo studio di un testo rispeakerato: l'analisi del genere a cui appartiene il testo da analizzare e la trascrizione e l'analisi multimodali, in grado di permeare la natura semiotica del testo in questione.

3.7 Genre Analysis

Nonostante Allen facesse ancora riferimento all'uso che si fa del genere in botanica (1989: 44) e Stam si chiedesse ancora se la *genre analysis* dovesse essere descrittiva o prescrittiva (2000: 14), essa è riconosciuta come un utile strumento per l'analisi testuale, in quanto offre un quadro teorico di riferimento pur adattandosi alle specificità di ogni singolo testo e alle finalità di ogni ricerca, senza pretesa di essere il frutto di una insindacabile esattezza scientifica valida una volta per tutte. D'altronde la difficoltà di classificare singole manifestazioni come appartenenti a un solo e unico genere è un'esperienza non nuova, perfino in quelle scienze definite esatte. Per quanto riguarda da vicino i media in generale e la televisione in particolare, tale difficoltà è quanto mai presente. Le classificazioni, infatti, mutano nel tempo a seconda dei cambiamenti che avvengono all'interno della società in cui alcuni programmi sono trasmessi (Miller 1984, Freedman e Midway 1994). Inoltre, per molti generi e sotto-generi non esistono definizioni (Fowler 1989: 216, Wales 1989: 206) e quello che per alcuni teorici è un genere per altri potrebbe essere un sotto-genere, se non addirittura una componente (Feuer 1992). Infine, con la diffusione sempre crescente dei programmi internazionali e la produzione di *format* identici in molteplici nazioni, alcuni elementi che compongono un genere sono trasversali ad altri generi (Bordwell 1989).

Swales (1990) si spinge addirittura ad affermare che ogni nuovo testo all'interno di un determinato genere ne modifica la natura fino a creare un nuovo genere o un nuovo sotto-genere. Una prospettiva di questo tipo mette in luce un altro aspetto fondamentale della *genre analysis*, vale a dire il potere degli autori di un dato testo di influire sulle sorti del genere a cui il testo appartiene. Studiando da vicino i generi televisivi, Abercrombie radicalizza questa posizione ribadendo la permeabilità tra i generi e ipotizzando perfino il totale smantellamento del genere in seguito alla caccia all'*audience* tipica della televisione degli ultimi anni (1996: 45).

Tutto questo conduce al problema dell'identificazione delle componenti che si devono considerare per poter definire un genere. In ambito cinematografico, Bordwell (1989) e Stam (2000) hanno ognuno individuato numerosi criteri per definire i sotto-generi filmici come ad esempio il regista, l'attore, la casa produttrice, il periodo storico, il paese, l'ideologia politica, la corrente letteraria, l'argomento trattato, il pubblico di destinazione, la struttura narrativa, ecc. Convenzionalmente, i generi filmici si definiscono in base alla forma (struttura narrativa, stile, corrente letteraria, ecc.) e al contenuto (argomento trattato, ambientazione, ecc.) anche se a tal proposito Stam sostiene che "subject matter is the weakest criterion for generic grouping because it fails to take into account how the subject is treated" (2000: 14). Difficoltà simili si riscontrano anche in ambito televisivo per cui non è possibile stabilire con certezza a quale genere appartengano alcuni programmi. Non esistono più generi precostituiti e la presenza di molteplici componenti all'interno di un unico programma lo rendono irrimediabilmente un ibrido.

Sebbene ogni singolo episodio di un determinato genere possa presentare caratteristiche di altri generi, è tuttavia innegabile che programmi specifici combinano caratteristiche specifiche in maniera del tutto distintiva, tanto da rendere immediatamente individuabile un determinato programma (cfr. Neale 1980). Quest'ultimo sarà contraddistinto da una sequenza di ripetizioni all'interno delle quali emergono differenze più o meno marcate che rendono unico ogni singolo testo del genere a cui appartiene, ma allo stesso tempo ben individuabile come un'occorrenza del genere stesso. Anche in questo caso però lo stesso testo potrebbe essere considerato come appartenente a generi diversi a seconda dei paesi e delle epoche. Ciononostante, alcuni generi televisivi sembrano non soffrire del proliferare e del mutamento a cui sono sottoposti altri generi televisivi, in quanto godono di uno *status* tale da permettere loro di sopravvivere ai cambiamenti della società. È

il caso del notiziario, che in buona parte dei paesi 'occidentali' è costituito da caratteristiche sia formali, sia sostanziali specifiche, tali da renderlo un genere televisivo piuttosto ben definito. Anche in questo caso, però, Norman Fairclough (1992) sostiene che testi giornalistici appartenenti a diversi generi o composti da elementi tipici di generi diversi non sono una rarità. Non esistono più testi che presentino tutte le caratteristiche del genere a cui appartengono. D'altronde, vista l'eterogeneità dei programmi televisivi è quasi impossibile che i generi non si contaminino. E questo vale non solo per i produttori, desiderosi di aumentare gli introiti dei singoli programmi, ma anche per gli spettatori, che fanno sempre più fatica a stabilire con esattezza l'appartenenza di un programma a un dato genere, specialmente se fanno raffronti con il passato (cfr. Abercrombie 1996).

In un contesto come quello summenzionato, in cui attaccare delle etichette è utile soltanto alla catalogazione, ma non all'analisi, forse maggiore successo potrebbe avere l'approccio prototipico, che propone una gerarchizzazione dei testi a seconda della loro maggiore o minore somiglianza con il rappresentante più stereotipico del genere a cui appartiene. Secondo questo approccio, compito del ricercatore è stabilire quali elementi di un determinato genere sono presenti in un testo dato e conseguentemente quale posizione occupa il testo in questione all'interno del *continuum* che dal centro si sposta verso la periferia, come nel caso dei campi semantici. Tuttavia, sebbene questo approccio sia molto più adeguato alla realtà televisiva di quanto non lo sia l'approccio precedentemente delineato, sembra non essere utile ai fini di un'analisi che intende stabilire l'adeguatezza di determinate strategie rispetto ai fini prefissati.

Si viene quindi a imporre l'esigenza di una descrizione dei singoli programmi in termini di parentela con altri programmi (Fowler 1989), piuttosto che in termini di definizione o di prototipicità. Secondo questo approccio, ogni testo viene considerato come un prodotto a se stante, influenzato da altri tipi di testi con i quali condivide alcune caratteristiche e non altre. Sebbene alcuni lo abbiano criticato perché troppo dispersivo e troppo dipendente dalla soggettività del ricercatore, (cfr. Swales 1990) tale approccio offre il vantaggio di liberare lo studio del testo in esame dai vincoli psicologici imposti dalla stereotipizzazione che scaturisce dai due approcci precedenti. A tal proposito, è tuttavia da non sottovalutare il ruolo del pubblico e dell'effetto che ha su di esso il testo e i suoi autori. Senza sfociare in derive ideologiche, è infatti bene tenere a mente che il semplice fatto di essere considerato come appartenente a un genere dato, un testo giornalistico tende a

indossare una determinata veste agli occhi degli spettatori (cfr. Feuer 1992). Ecco quindi che si giunge forse alla posizione più adeguata all'obiettivo della presente tesi, quella di Miller (1984), che sostiene che la definizione del genere deve essere incentrata non tanto sulla forma o sul contenuto, ma sugli strumenti che sono utilizzati dagli autori di un testo per raggiungere l'obiettivo prefissato: il processo, la funzione del processo e del prodotto e infine l'effetto sull'utenza finale.

Che cos'è quindi un genere? Stando a quanto è stato finora delineato, il genere è il rapporto che si viene a instaurare tra uno o più mittenti, i destinatari, il testo (inteso in senso hallidayiano come prodotto di un sistema sociale e semiotico più vasto e pervasivo dei soli mittenti e destinatari dello stesso in un contesto dato), l'intertestualità (cfr. Barthes 1975) e gli strumenti utilizzati sia per produrre, sia per trasmettere il testo in questione. A questo, si aggiunge il codice, in qualche modo dato per acquisito sia dai mittenti, sia dai riceventi, che garantisce il buon esito della comunicazione (cfr. Fowler 1989).

E qual è il ruolo del ricercatore? Se si accetta la posizione attuale, al ricercatore la massima discrezione. Senza assumere atteggiamenti aprioristici, suo compito sarà di notare le somiglianze e le differenze tra i testi in esame e stabilire se possono essere definiti come appartenenti allo stesso genere. Se lo scopo del ricercatore va oltre la definizione di un genere o l'investigazione dell'appartenenza o meno di uno o più testi a un dato genere, allora l'approccio descrittivo permetterà anche il compenetrarsi con un'analisi funzionale all'obiettivo finale. A tal proposito, Tudor (1985) sostiene che soltanto quelle definizioni funzionali ai testi che sono considerati parte del genere in questione possono sopravvivere all'usura del tempo. Se l'obiettivo finale dell'analisi è quindi accertare il raggiungimento dell'obiettivo prefissato da parte di uno o più testi, allora strumenti del ricercatore saranno le categorie di funzione del testo e degli autori, di utenza finale e di strumenti messi in atto per raggiungere l'obiettivo. Un posto centrale è poi occupato dall'unità minima di analisi, indispensabile all'oggettività del ricercatore. A seconda del genere testuale, infatti, si individuerà l'unità minima di analisi che si riterrà più attinente al lavoro da svolgere e si stabilirà la maggiore o minore coerenza con l'obiettivo principale in tutte le fasi e sottofasi del testo in esame.

3.8 L'analisi multimodale

Partendo dal presupposto che "there are many other resources that can be used to create texts in addition to spoken and written word" (2005: 4), Baldry e Thibault sviluppano il concetto di multimodalità, i cui elementi fondanti sono quelli che vengono definiti *semiotic resource system*, cioè "semiotic forms that we can *use* for the purpose of making texts. The forms have particular *functions* in the texts in which they are used" (2005: 18). Questi 'sistemi di risorse semiotiche' sono quindi i materiali tramite i quali viene costruito un testo all'interno del quale svolgono un ruolo fondamentale nella costruzione del significato. Nella creazione di ogni testo, questi sistemi seguono due principi organizzativi essenziali:

- *resource integration principle*: secondo cui un testo non deve essere considerato come la somma "of the different resources used, taken separately" (*ibidem*), ma piuttosto come un sistema, come i diversi modi in cui "the selections from the different *semiotic resource systems* in multimodal texts relate to, and affect each other, in many complex ways across many different levels of organisation" (*ibidem*);
- *meaning-compression principle*: per cui "patterned multimodal combinations of visual and verbal resources on the small, highly compressed scale of any text provide semiotic models of the larger, more complex realities that individuals have to engage with" (*ibidem*).

In altre parole, secondo gli autori, un testo è il risultato della collaborazione di diverse componenti, ognuna delle quali svolge una funzione specifica nel processo di costruzione del significato del testo. Questa funzione ha un senso solo se concepita all'interno di un sistema più grande, che, a sua volta, è la maniera più economica per esprimere un livello più alto di organizzazione testuale a cui prende parte anche il ricevente del testo in questione. In questo contesto, la componente non-verbale non ha quindi più un ruolo di secondo piano, alle dipendenze del testo inteso come espressione della lingua, ma un ruolo di primo piano, di collaborazione con la lingua. Insieme costruiscono il significato del testo.

Per poter quindi cogliere appieno il significato di ogni testo, è necessario attuare una trascrizione del testo che faccia emergere ogni *semiotic resource system* e ne descriva la funzione all'interno del testo. Per raggiungere questo obiettivo, Baldry e Thibault (2005)

propongono la trascrizione fasale, la cui unità minima di analisi è la fase, ovvero quella che Gregory (2002) chiama "a set of copatterned semiotic selections that are codeployed in a consistent way over a given stretch of text". L'analisi fasale segmenta quindi il testo nelle singole unità di base che lo costituiscono (le fasi appunto), le sottofasi e i 'punti di transizione' cioè "when one phase or subphase ends and another begins" (Baldry e Thibault 2005: 47)[98]. All'interno di ogni elemento analizzato saranno descritti i sistemi di risorse semiotiche e il loro contributo al significato generale del testo. In particolare i seguenti aspetti saranno maggiormente da considerare:

- i secondi di inizio e fine di ogni fase;
- la componente video non verbale e in particolare:
  - gli *shot* cioè "visual sequence in which there is no spatial displacement of the camera" (Baldry e Thibault 2005: 187) trascritti sotto forma di *frame*, "a visual transcription of some aspects of the visual track" (*ibidem*);
  - la struttura e l'organizzazione dell'informazione, cioè gli elementi 'rematici' del testo e la maniera in cui questi vengono presentati rispetto a quelli 'tematici', a partire da una serie di opposizioni binarie: destra-sinistra, lontano-vicino, variante-invariante, ignoto-noto, ecc.;
  - la transizione tra i vari *shot* e l'interdipendenza tra gli *shot* da questa determinata (subordinatezza-superordinatezza, continuità-discontinuità, temporalità-logicità, ecc.).
- la maniera in cui le opzioni presenti nel sistema semiotico utilizzato per 'significare' organizzano il rapporto tra il testo e lo spettatore. In particolare:
  - il movimento della telecamera (se si muove – verso destra-sinistra, avanti-indietro, alto-basso, generale-particolare – o è fissa e se si muove perché vuole essere dato un effetto particolare alle immagini o semplicemente perché devono essere ripresi i movimenti dei partecipanti all'evento);
  - la prospettiva dalla quale si assiste alle immagini. Può essere orizzontale (un'inquadratura centrale indica coinvolgimento nell'evento descritto; un'inquadratura obliqua indica distacco) o verticale (un'inquadratura dall'alto verso il basso mette lo spettatore in una posizione di potere sul testo; una

98 In ultima analisi, si potrebbero considerare le fasi e le sotto-fasi come i move e gli step di cui parla Bhatia (2002).

centrale indica, come nel caso precedente, solidarietà, uguaglianza; con un'inquadratura dal basso verso l'alto, infine, lo spettatore è messo in un piano d'inferiorità rispetto a testo);

- la distanza virtuale tra lo spettatore e i personaggi sullo schermo. Siccome vicinanza significa maggiore intimità con le persone che compaiono sullo schermo e lontananza significa distacco, la telecamera potrebbe essere utilizzata in maniera da produrre un certo effetto sullo spettatore. Si distinguono essenzialmente sei gradi di distanza a seconda del piano che viene fatto dei personaggi inquadrati: molto vicino (primissimo piano), vicino (viso e spalle), abbastanza vicino (mezzobusto), abbastanza lontano (tutta la figura), lontano (la figura occupa la metà dello schermo) e molto lontano (la distanza è maggiore;
- la collocazione visiva, cioè la presenza di elementi secondari che non hanno lo status di partecipante, ma servono proprio a caratterizzare una persona contribuendo così al significato generale del testo nella sua multimodalità. Tali elementi possono essere catalogati a seconda della loro deitticità: corpo (tatuaggi, postura più o meno eretta, tratti somatici più o meno marcati, chiari segni di violenza subita, cicatrici, ecc.), abbigliamento (tailleur, scarpe da ginnastica, tuta da operaio, camice, ecc.), ambiente (officina, studio medico, strada, casa, ecc.), ruolo (avvocato, politico, chirurgo, passante, madre della vittima, ecc.), ecc.;
- la prominenza visiva di ogni elemento presente nell'immagine distinta in termini di dimensione, posizione nello schermo, definizione e orientamento nello spazio;
- la direzione dello sguardo dei partecipanti al testo multimodale (ad altri partecipanti, alla telecamera, ecc.) e altri elementi prossemici (pacche sulla spalle, strizzatine d'occhio, smorfia di dubbio, ecc.);

- tutti gli elementi cinestetici presenti nel testo. In particolare saranno considerati:
  - i movimenti secondo la struttura attore-azione-risultato, agente-azione-reazione, ecc. e secondo la loro relazione con il testo verbale;
  - la forza illocutiva dei movimenti, considerata in termini di partecipazione dell'attore nell'azione (entusiasmo, cinismo, imitazione del reale, ecc.);

- la presenza di eventuali movimenti subordinati al movimento principale (causalità, consequenzialità, ecc.).

- la componente audio del testo, verbale e non verbale, extra- e para-linguistica. In particolare:
  - eventuale musica di sottofondo;
  - elementi extra-linguistici che contribuiscono ugualmente alla costruzione del significato (suono di campanella, rumori ambientali, grida, ecc.) a seconda della loro maggiore (*figure*), intermedia (*ground*) o minore (*field*) rilevanza nel testo;
  - la relazione di un elemento sonoro extra-linguistico con un altro elemento linguistico o extra-/para-linguistico;
  - altri elementi che potrebbero caratterizzare la componente sonora del testo (suono di voce femminile o maschile, gutturale o nasale, profonda o acuta, ecc.);
  - elementi para-linguistici come il ritmo dell'eloquio (pause lunghe o brevi, ritmo incalzante o rilassato, ecc.), il timbro della voce, il tono della voce, ecc.;
  - turnazione tra i partecipanti all'evento comunicativo.

La trascrizione multimodale proposta da Baldry e Thibault è sicuramente molto utile agli scopi di un'analisi del significato di un testo audiovisivo. Tuttavia, sembra opportuno aggiungere l'annotazione di altri elementi che contribuiscono non solo alla costruzione del significato generale del testo, ma anche alla definizione dei vincoli posti dal testo al lavoro del sottotitolatore. In particolare sarà interessante annotare:

- il numero di parole al minuto pronunciate nel TP. Si tratta di un elemento significativo che potrebbe contribuire a un'eventuale strategia di riduzione da parte del rispeaker;
- il divario tra la pronuncia del TP e la comparsa del relativo sottotitolo sullo schermo.
- la presenza di didascalie sullo schermo, identificatrici di una località o dei partecipanti all'atto comunicativo;
- la posizione del sottotitolo sullo schermo per determinare se copre o meno altre didascalie sullo schermo o le bocche degli oratori;

3.9 Conclusioni

In questo capitolo sono stati affrontati molteplici aspetti riguardanti la sottotitolazione per sordi di programmi pre-registrati. L'obiettivo finale era duplice: a lungo termine è stato interessante notare quali sono le caratteristiche che deve avere la sottotitolazione per non-udenti intesa come prodotto. Pur essendo concentrata sulla sottotitolazione di programmi pre-registrati, l'analisi contenuta in questo capitolo ha comunque fornito alcune linee guida circa il risultato finale che deve raggiungere ogni sottotitolazione per non-udenti, indipendentemente dalla modalità di produzione e dalla tecnologia impiegata. Nei limiti tecnico-tecnologici imposti dal riconoscimento del parlato, il rispeakeraggio potrà infatti seguire le indicazioni qui contenute per raggiungere il suo scopo d'inclusione sociale; a breve termine, l'obiettivo era di trarre il massimo dell'insegnamento dagli studi sulla traduzione audiovisiva per poter costruire un quadro teorico di riferimento che possa contribuire alla creazione di un modello di analisi testuale per i sottotitoli in diretta prodotti tramite la tecnica del rispeakeraggio.

Per fare ciò è stato necessario attraversare tutte le tappe che hanno portato sia a livello pratico sia a livello teorico alla definizione della sottotitolazione per sordi di programmi pre-registrati in generale e di film in particolare. Dal punto di vista pratico, sono stati ripercorsi i passi compiuti dai primordi fino ai giorni nostri, analizzando le tecniche di produzione e diffusione dei sottotitoli. Dal punto di vista teorico, vista la necessità intrinseca alla sottotitolazione per sordi di essere accessibile all'utenza finale, è stata sottolineata sia l'esigenza di una maggiore concentrazione da parte della comunità scientifica sugli aspetti di ricezione e percezione dei testi audiovisivi, sia la necessità di una standardizzazione delle tecniche e delle prassi di produzione e proiezione dei sottotitoli per sordi, indispensabile allo scambio di idee e al miglioramento della professione. Infine, si sono analizzati i vari modelli di analisi testuale provenienti dai *translation studies* prima e dagli studi sulla traduzione audiovisiva poi che meglio si adattano all'analisi di testi rispeakerati.

In quest'ottica è stato selezionato il modello strategico proposto da Gambier (2006), che ha il vantaggio di essere più flessibile e chiaro degli altri, oltre che di più semplice applicazione. Questo modello, appositamente adattato alle esigenze dell'analisi in questione, costituirà uno dei tre pilastri su ci dovrà poggiare la metodologia di analisi di un qualsiasi

testo rispeakerato. Gli altri due sono: la *genre analysis*, che permetterà, a seconda dei casi, di individuare le unità minime da analizzare in ottica strategica; e l'analisi multimodale, in grado di sviscerare il potenziale semiotico di ogni testo, la cui conoscenza è indispensabile in vista di un raffronto tra un TP e un TA.

Capitolo 4 - Analisi strategica di BBC News

4.1 Introduzione

Grazie ai capitoli precedenti, è stato possibile avere delle indicazioni sugli obiettivi che deve raggiungere un rispeaker nell'espletamento della sua professione, nei limiti professionali e operativi che la contraddistinguono. Oltre ai vincoli imposti dal *software* di riconoscimento del parlato (primo fra tutti l'inevitabile divario temporale tra l'emissione del TP e la ricezione di quello di arrivo), è stato anche possibile comprendere il carico cognitivo a cui è sottoposto un rispeaker nell'espletamento della sua professione (capitolo 2) e individuare le caratteristiche formali che devono avere dei sottotitoli per poter raggiungere l'obiettivo di accessibilità a cui dovrebbero sottostare (capitolo 3). In particolare, è emersa chiaramente l'esigenza di avere informazioni circa la componente audio del TP, cioè a dire non soltanto la componente verbale, ma anche tutti gli strumenti impiegati per la costruzione del TP (pragmatica del linguaggio, colonna sonora ed effetti speciali).

Grazie a queste prime indicazioni e grazie ai contributi derivanti dagli studi sull'interpretazione simultanea e sulla traduzione audiovisiva in generale, il quadro teorico è ora pronto per accogliere l'analisi delle strategie seguite dai rispeaker della BBC. Sarà così possibile derivare delle linee guida che potranno essere di utilità a ogni rispeaker in lingua inglese. Tali linee guida serviranno poi adattate anche per una definizione delle strategie da utilizzare in lingua italiana. A tal fine, sono state inizialmente registrate tra il 4 e il 6 luglio 2005 otto ore di BBC News, un programma in diretta e semidiretta sottotitolato in tempo reale dai rispeaker della BBC[99]. La scelta è caduta in particolare su questo programma che, insieme agli altri due macro-generi per cui è offerta una sottotitolazione tramite rispeakeraggio, cioè a dire le sessioni parlamentari in diretta e le competizioni sportive, sono, secondo il responsabile del servizio di sottotitolazione intralinguistica della BBC "less demanding in terms of accuracy" (Blizzard 2005). Le ragioni sono di politica interna dell'emittente: i programmi in diretta o semidiretta con maggiore *share* (i notiziari di BBC1 e BBC2 ed eventi di grande richiamo come i funerali del Papa, il Live 8 o la notte degli Oscar) sono sottotitolati dagli stenotipisti, in quanto considerati maggiormente preparati e accurati. I programmi sportivi invece sono sottotitolati dai rispeaker in quanto più facili da

99 A partire dal 2006, il servizio di sottotitolazione della BBC è stato smantellato e subappaltato alla società privata RedBee Media, composta per la maggior parte da ex dipendenti della BBC (cfr. Marsh 2006).

sottotitolare perché è possibile attuare una maggiore compressione del TP visto che "you describe the action you can see on the screen, so you do not need to speak all the time" (Marsh 2005). Anche BBC News e le sedute del parlamento sono infine sottotitolate dai rispeaker perché, nonostante siano difficili da sottotitolare, perché molto veloci e privi di una vera e propria componente video in grado di compensare le informazioni provenienti dalla componente audio, sono programmi 'low profile', cioè seguiti da poche persone rispetto ad altri programmi di informazione. Il materiale di riferimento è quindi costituito da otto ore di notiziari di BBC News 24, il canale esclusivamente dedicato all'informazione dell'ultima ora. Per tutti questi programmi, la componente audio verbale del TP è stata trascritta[100] e conseguentemente allineata con la trascrizione dei rispettivi sottotitoli (comprensiva anche delle didascalie relative alla componente para- ed extra-linguistica). Ogni scostamento dalla componente audio verbale del TP (riduzione, riformulazione, ampliamento, ecc.) sarà catalogato sulla base di una tassonomia delle strategie traduttive ispirata da quella compilata da Gambier (2006). Per cercare di capire quali siano state le ragioni dietro un tale riscontro nel TA, si farà, quindi, ricorso sia al modello degli sforzi di Gile (1995), sia alla tassonomia delle strategie operative elaborata da Kohn e Kalina (1996). Quanto al raggiungimento dell'obiettivo finale, la valutazione è lasciata all'interpretazione dei dati provenienti sia dall'analisi strategica, sia da quella del genere e infine dall'analisi multimodale proposta da Baldry e Thibault (2005: 165-249).

Prima di entrare nei dettagli, è forse bene partire dalle norme prescrittive che i rispeaker della BBC sono chiamati a rispettare.

4.2 Le linee guida della Ofcom

In seguito al Broadcasting Act del 1990, le emittenti britanniche sono state obbligate a migliorare la propria produzione di sottotitoli sia in termini quantitativi, cercando di puntare cioè alla sottotitolazione del 100% dei programmi trasmessi entro il 2010, sia qualitativi. Per quanto riguarda questo ultimo punto, l'*Independent Television Commission* (ITC), l'ente britannico deputato al controllo dell'avvenuto rispetto di questi *standard*, in collaborazione con le università di Southampton e Bristol ha condotto nel corso

100 La trascrizione utilizzata è di tipo ortografica, per cui sono stati considerati solo alcuni tratti tipici dell'oralità. In particolare, non sono stati considerati gli eventi non linguistici come le pause piene, le false partenze e gli idiolettismi fonetici, in quanto sistematicamente ignorati dai rispeaker. Sono però state trascritte verbatim tutte le costruzioni sintattiche comprese le false partenze in ragione della loro rilevanza dal punto di vista delle strategie da attuare e dell'esito finale.

degli anni Novanta una ricerca sulla leggibilità dei sottotitoli da parte degli utenti sordi, di tutte le tipologie. Questa ricerca è sfociata, nel 1999, nella pubblicazione dell'*ITC Guidance on Standards for Subtitling*, un manuale che riporta linee guida per ogni aspetto concernente la sottotitolazione per non-udenti di qualsiasi programma televisivo. Attualmente, l'*Office of Communications* (Ofcom) che ha rilevato l'ITC, utilizza tale manuale come strumento indispensabile all'espletamento del proprio lavoro di controllo della qualità della produzione televisiva britannica.

Dal punto di vista grafico, il sottotitolatore deve fare uso della punteggiatura per rendere nel migliore dei modi la sintassi del TA. Inoltre, deve fare uso di colori diversi per sottotitolare ogni elemento non verbale intuibile dalla componente video e utile alla comprensione del TA e per identificare personaggi diversi (in ordine decrescente di importanza e leggibilità: bianco, giallo, ciano e verde) o anche soltanto per notificare l'avvenuto cambiamento di oratore nel caso di programmi in cui l'identità dei personaggi non è così importante come nei film (ad esempio nelle telecronache e nelle sedute parlamentari e in misura minore nei TG). Altri elementi grafici sono la posizione del sottotitolo sullo schermo, che deve occupare preferibilmente la parte bassa, e la sua lunghezza massima (una, due o massimo tre righe di sottotitoli di 32-34 caratteri ciascuna). Per quanto riguarda infine la tempistica e la segmentazione della frase all'interno dei singoli sottotitoli, tutte le indicazioni fornite, nel caso della sottotitolazione in tempo reale, sono vanificate. Nel primo caso si richiede di sincronizzare i sottotitoli con il testo originale e di evitare eccessivi ritardi e sovrapposizioni con immagini relative ad altri enunciati. Questa richiesta è inaccoglibile, visto il fisiologico divario tra il TP e quello di arrivo (cfr. 2.2.3.). Nel secondo caso, si richiede di evitare di produrre sottotitoli difficili da leggere andando a capo compatibilmente con la struttura sintattica del testo dei sottotitoli. Anche in questo caso la linea guida non può essere rispettata vista la modalità di proiezione dei sottotitoli adottata dalla BBC per sottotitolare i programmi i diretta: un misto tra *scrolling* e *roll-up*, che non permette una visualizzazione in blocchi (*pop-on*), in quanto le parole compaiono una alla volta spingendosi l'una l'altra fino alla fine della riga inferiore in un flusso continuo (*scrolling*) per poi salire, in modalità *roll-up*, alla riga superiore, che a sua volta sarà sostituita dalla riga successiva. Sorte simile spetta all'identificazione di oratori che parlano fuori schermo, alla resa di elementi non verbali, degli effetti speciali e delle musiche di sottofondo. In questo caso, ad annullare lo standard qualitativo è la velocità dei tempi di

reazione del rispeaker, nei limiti imposti dal genere in questione. Qualora si trovasse a sottotitolare generi come la telecronaca tennistica, in cui sono previsti applausi, interventi dell'arbitro e disappunto del pubblico, la migliore soluzione è il ricorso alle macro di dettatura. Nei casi in cui un fenomeno è a stento prevedibile, la difficoltà è tale da giustificare un'omissione funzionale alla resa della componente verbale. Un discorso a parte va fatto per la velocità di lettura. Come già accennato, il potenziale del sistema di riconoscimento dl parlato è di 300 parole a minuto, ma gli studi condotti dalla ITC e da alcune università inglesi hanno dimostrato che la velocità massima a cui deve tendere un sottotitolo è di 140 parole al minuto, prorogabili fino a 180 nei casi in cui sono presenti le didascalie esplicative (*add-on*). Nella sottotitolazione in diretta questa operazione è difficilmente attuabile per l'impossibilità cogente di cronometrarsi ed eventualmente di modificare il proprio ritmo.

Per quanto riguarda gli aspetti più specificatamente linguistici, le priorità alla base di qualsiasi lavoro di sottotitolazione intralinguistica, secondo Ofcom (ITC 1999: 4), sono tre:

> - Allow adequate reading time […]
> - Reduce viewers' frustration by:
>   - attempting to match what is actually said, reflecting the spoken word with the same meaning and complexity; without censoring
>   - constructing subtitles which contain all obvious speech and relevant sound effects; […]
> - Without making unnecessary changes to the spoken word, construct subtitles which contain easily-read and commonly-used English sentences in a tidy and sensible format.

Già da questa breve disamina si intuisce la complessità del lavoro del sottotitolatore (di ogni genere di programmi) che deve giostrarsi tra una resa che sia alo stesso tempo fedele alla complessità del TP (2a) e leggibile dal pubblico di destinazione (3). Se questa linea guida si considera nel contesto della sottotitolazione in tempo reale, allora la difficoltà è ancora più evidente. A rendere maggiormente complicato il quadro delle difficoltà sono le linee guida per la sottotitolazione di programmi di informazione, che "should convey the whole meaning of the material" (ITC 1999: 25). Questo non significa che le stese parole debbano essere utilizzate. Per risolvere problemi dovuti alla velocità del TP e un sovraccarico sia per il rispeaker, sia per il *software*, Ofcom introduce il concetto di '*idea unit*', cioè "where a proposition or key information is given" (ITC 1999: 25). Dopodiché, passa

all'enumerazione delle indicazioni minime da rispettare in fase di sottotitolazione. Esse sono:

- Subtitles should contain a reasonable percentage of the words spoken.
- 'Idea units' or key facts should appear as a good percentage of the spoken message […].
- Avoid 'idea units' which are unnecessary or different from the original.[…] (*ibidem*)

Nel caso specifico della sottotitolazione di un notiziario, l'unità concettuale è un concetto abbastanza ampio, non specificato dalle linee guida, che può essere inteso sia come il senso generale di una frase o le singole informazioni in essa contenute. Per evitare di cadere in valutazioni troppo arbitrarie, nell'analisi che segue si è optato per la segnalazione di ogni tipo di intervento avvenuto sul TP. Ogni singola modifica apportata al TP verrà considerata come il frutto di una strategia di riduzione, alterazione o espansione del TP. Nel caso in questione, l'omissione di 'tomorrow' sarebbe annoverata tra le strategie di riduzione. In un secondo momento, si valuteranno l'eventuale aggiunta o perdita di *idea unit* (operazione semantica) e la conseguente efficacia dell'operazione in questione. Sempre nel caso dell'esempio precedente, l'operazione sarebbe considerata adeguata in quanto non ha comportato una perdita di informazioni importanti alla comprensione del TP.

Ritornando alle linee guida della Ofcom, esse entrano nel dettaglio sia della presentazione dei sottotitoli in diretta (compresi quelli precedentemente preparati ma proiettati in diretta), sia della professionalità del sottotitolatore. Di queste, due sono le indicazioni di particolare interesse linguistico per l'analisi proposta:

- Send an apology caption following any serious mistake or a garbled subtitle; and, if possible, repeat the subtitle with the error corrected;
- Do not subtitle over existing video captions where avoidable (in news, this is often unavoidable, in which case a speaker's name can be included in the subtitle if available).

4.3 L'analisi linguistica dei programmi rispeakerati

Alla luce di queste indicazioni, l'analisi linguistica dei programmi sottotitolati dai rispeaker della BBC sarà effettuata considerando una *idea unit*, o unità concettuale, come stringa minima di significato. In particolare, ogni singolo *corpus* è stato suddiviso in unità concettuali. Visto che i sottotitoli non compaiono a blocchi, ma scorrono parola per parola, per

unità concettuale non si è inteso, in questa sede, il singolo sottotitolo, ma ogni enunciato di senso compiuto indipendentemente dalla sua forma (parola-frase, subordinata, incidentale, elemento di una lista, ecc.). A seconda delle dimensioni, è stato possibile riscontrare all'interno di ogni *idea unit*, unità concettuali più piccole dal senso inscindibile dall'enunciato di appartenenza (aggettivi qualificativi in una lista, avverbi di modo, locuzioni, ecc.). Dopo questa segmentazione, le unità concettuali sono state catalogate in due categorie differenti:

- ripetizioni: quelle che sono state ripetute dai rispeaker senza alcuna modifica;
- alterazioni[101]: quelle che hanno subito una trasformazione, anche solo formale.

Un'ulteriore suddivisione è stata effettuata all'interno del secondo gruppo a seconda della tipologia di strategia utilizzata. A tal fine, si è fatto riferimento alla categorizzazione ideata da Gambier, secondo cui tre sono le macrostrategie attuate dai sottotitolatori nel passaggio dal TP al TA: espansione, semplificazione della sintassi e riduzione. Per i nostri scopi, tuttavia, si è resa necessaria una modifica a questa tassonomia. In particolare, si è reso necessario eliminare la macro-categoria della semplificazione della sintassi, che è stata inglobata nelle altre due categorie, a seconda che la semplificazione comportasse una riduzione o un'espansione. All'interno delle due macrocategorie restanti, espansione e riduzione, sono state eseguite due ulteriori suddivisioni: espansione/riduzione non-semantica ed espansione/riduzione semantica. Per espansione/riduzione non-semantica si è inteso ogni modifica al TP che comportasse un mero aumento/decremento del numero di grafemi, senza alcun tipo di intervento sul significato del TP (eliminazioni di tratti dell'oralità, aggiunta di elementi non verbali, dislocazioni sintattiche, ecc.). Espansione/riduzione semantica, di contro, è stata considerata la modifica al TP che comporta una variazione dell'assetto semantico del TP (sinonimia, parafrasi, esplicitazione, ecc.). In una fase successiva, laddove evidente, si è infine operata un'ulteriore suddivisione all'interno di ogni macro-categoria, a seconda della sotto-tipologia di operazione attuata. Un paragrafo a parte è stato poi dedicato agli errori, inevitabili nella sottotitolazione dal vivo. Anch'essi sono stati suddivisi in errori del computer ed errori umani, a seconda che si ritenesse che l'errore comparso sullo schermo fosse dovuto a un'inadempienza dell'operatore (cattiva pronuncia, mancato uso degli ausili a disposizione, ecc.) o semplicemente a un'incapacità del *software* di gestire correttamente l'*input* (sinonimia,

101 Il termine alterazione è inteso qui nel senso che Gambier (1992) attribuisce a ogni forma di riformulazione, cioè di altra cosa, seppur minimamente differente, rispetto al TP.

numeri, dettatura di comandi, ecc.). Degno di nota è infine l'alterazione senza cambiamento del numero di grafemi. Sono stati riscontrati essenzialmente due casi:

- la sinonimia;
- la dislocazione.

Nel primo caso, in presenza di sinonimia verticale, essa è stata attribuita all'espansione o alla riduzione semantiche a seconda che si trattasse di ipo- o iperonimia rispettivamente; nel caso della sinonimia orizzontale, essa è stata considerata come espansione semantica nei pochi casi in cui il rispeaker interviene per correggere il TP, in termini di collocazione, concordanza o altro tipo di errata formulazione del testo dal punto di vista grammaticale o pragmatico. È stata considerata invece come riduzione non-semantica la sinonimia risultante da una erronea o voluta operazione mnemonica da parte del rispeaker. In questo ultimo caso, un'osservazione è forse degna di nota. Benché assimilata a una strategia non-semantica, alla stessa stregua dell'eliminazione dei tratti tipici dell'oralità, dell'aggiunta di elementi non marcati e altri, la sinonimia orizzontale risulta da una parte in un equilibrio semantico tra il TP e il TA, dall'altra in un'operazione mentale molto complessa, paragonata da Gran (1992), ai processi di traduzione interlinguistica.

Nel secondo caso, si tratta, in specifico, dello spostamento all'interno di una frase di lessemi o di sintagmi. Qualora la strategia risulti in una focalizzazione (l'operatore ha voluto mettere in evidenza un elemento della frase rispetto a un altro), la dislocazione è stata considerata come un esempio di espansione semantica. Quando invece è risultata in un semplice spostamento dell'ordine delle parole, probabilmente dettato da una non radicale fedeltà al testo o da una memorizzazione di tipo 'last in, first out'[102], senza alcun effetto sul senso, essa è stata ignorata e collocata nella categoria delle ripetizioni e non delle alterazioni.

Prima di passare all'analisi delle strategie, resta da segnalare che un'unità concettuale contenente diverse istanze di una o più strategie possa essere stata catalogata non come un singolo esempio di espansione/riduzione semantica/non-semantica, ma come tanti esempi di espansione/riduzione semantica/non-semantica in maniera proporzionata al numero delle istanze in questione. Per cui, qualora un'unità concettuale contenesse tre esempi di riduzione

102 Si tratta di una tipologia di memorizzazione utilizzata in interpretazione simultanea sia in ambito professionale, sia formativo, per cui l'ultimo elemento (o unità concettuale) espresso in un enunciato, viene ripetuto per primo onde evitare di memorizzarlo sul breve termine. Contrapposta a questa tipologia, troviamo la tipologia 'last in, last out', che rispetta maggiormente l'organizzazione sintattica del TP, ripetendola in maniera lineare.

non-semantica, tutti e tre gli esempi confluirebbero nella percentuale riferita all'incidenza della riduzione non-semantica sul totale delle operazioni di sottotitolazione effettuate.

4.4 *BBC News*

BBC News 24 è un canale della *British Broadcasting Corporation* volto alla trasmissione di notizie 24 ore su 24 dalle varie postazioni televisive che la BBC possiede in tutti gli angoli della terra. I programmi contenuti in questi canali sono tutti destinati all'informazione e si occupano dei più svariati aspetti. Alcuni sono specializzati in un determinato settore, come *Click* che si occupa di tecnologia, *Teen 24* che sviluppa tematiche di interesse per gli adolescenti e i giovani in generale, *SportsDay* che copre le maggiori notizie sportive del giorno, *OurWorld* che esamina le più importanti notizie provenienti da tutto il mondo, *The Week on Newsnight* che passa in rassegna i migliori film della settimana, *World Business Report* che aggiorna sulle principali notizie provenienti dal mondo degli affari oltre che sulle ultime evoluzioni dei mercati finanziari di Singapore, Francoforte, Londra e New York e infine *Dateline London* che riporta notizie riguardanti il Regno Unito così come sono state trasmesse dai notiziari delle maggiori emittenti mondiali; altri invece sono di informazione generale come *Breakfast* che fornisce informazioni sui fatti di cronaca del giorno precedente e sulle principali notizie di sport, economia, finanza e meteorologia, *HARDtalk* che si occupa delle principali notizie della settimana, *BBC Five O'Clock News Hour* che approfondisce alcuni eventi di cronaca del giorno e infine *BBC News* che, con un *format* di 30 minuti in cui vengono fornite notizie in tempo reale così come escono dalle agenzie o con dei *reportage* in diretta, è il programma maggiormente diffuso durante la giornata, dalle 25 alle 30 volte. Il programma è un notiziario a tutti gli effetti che trasmette notizie di interesse generale in diretta. Alcune notizie sono più importanti delle altre e vengono ribadite nelle edizioni successive con qualche aggiornamento, mentre ad altre ancora è dedicata un'attenzione che si prolunga nel tempo. Nel prossimo paragrafo sarà analizzato *BBC News* come genere e in seguito come TP sottotitolato in diretta.

4.5 Analisi di genere di *BBC News*

BBC News è un notiziario, quindi un esempio stereotipico di genere informativo. Secondo la definizione dell'EBU (*European Broadcasting Union*), un programma informativo in generale ha l'obiettivo primario di informare i telespettatori circa fatti di

cronaca, situazioni, eventi, teorie e previsioni attuali e di fornire loro informazioni sufficienti per avere un'idea chiara di quanto riportato oltre che delle opinioni (1995: 25) auspicabilmente oggettive. Un'altra caratteristica fondamentale del programma di informazione riguarda, sempre secondo l'EBU (*ibidem*), il contenuto, che non può essere sempre attuale ed essere quindi ritrasmesso un anno dopo senza perdere pertinenza.

Per quanto riguarda *BBC News* in particolare, il programma è per lo più in diretta, ma ci sono anche parti in semi-diretta e altre in differita. Questo significa che molti *reportage* che compongono *BBC News* sono proiettati in tempo reale (le immagini fanno riferimento a qualcosa che accade nel momento in cui sono visionate dal pubblico), ma ci sono anche parti, soprattutto quelle introduttive, le cui immagini provengo dallo studio di registrazione, i cui testi sono stati scritti prima e sono letti in diretta. Infine, ci sono parti che sono state registrate prima e sono proiettate in diretta. Questo fa sì che la componente video rispetto a quella audio svolge un ruolo diverso a seconda della parte in onda. Nei casi di diretta in cui il reporter è inquadrato, essa avrà un ruolo secondario rispetto alla componente audio (principalmente verbale). Sempre nel caso dei *reportage*, qualora le immagini mostrino l'evento in oggetto (concerto, fenomeno naturale, competizione sportiva, operazioni di salvataggio, ecc.), la componente video avrà un ruolo decisamente più importante visto che almeno una parte dell'informazione è veicolata proprio da quelle immagini commentate dalla voce del giornalista. Tuttavia, c'è da riconoscere che, nella maggior parte dei casi, l'oggetto della notizia non è mai mostrato per intero (è spesso troppo lungo), ma solo alcuni aspetti sono trasmessi (pezzo più celebre di un concerto, effetti di una catastrofe, goal di una partita di calcio, luoghi dei fatti, partecipanti, ecc.).

Quanto alla strutturazione del telegiornale, esso non segue mai la stessa logica visto che le notizie sono trasmesse appena sono ricevute dall'emittente, sconvolgendo così ogni piano precedentemente ipotizzato. Uno standard è comunque possibile da individuare:

- *immagini di apertura*;
- *titoli*;
- *servizi pre-registrati;*
- *reportage in diretta;*
- *previsioni meteorologiche;*
- *sommario;*

- *immagini di chiusura.*

Immagini di apertura

Un orologio digitale indica l'ora mentre suona il *jingle* del TG.

Titoli

Un presentatore augura il benvenuto alla trasmissione, comunica l'ora e legge i titoli delle principali notizie del telegiornale. La prima sotto-fase, o *step*, è composto da espressioni formulaiche, non lette, che seguono la stessa sequenza a ogni edizione. Malgrado l'assenza di un testo scritto di supporto, non c'è spazio in questo passaggio per l'oralità. Il secondo *step* è costituito dalla lettura del primo titolo dal presentatore o da uno dei due presentatori presenti in studio. In caso di due presentatori, il secondo titolo può essere letto sia dal secondo presentatore con il quale il primo si alternerà nella lettura dei titoli, sia dallo stesso. In questo secondo caso, alla fine della lettura dei titoli, il secondo presentatore inizierà a introdurre il primo servizio giornalistico. Questo *step* è una versione condensata del servizio o dei servizi che seguiranno, che è stata scritta per essere letta in studio. Si tratta pertanto di un testo da un basso tasso di *grammatical intricacy* (complessità grammaticale) e da un alto tasso di *lexical density* (densità lessicale)[103]. I testi sono infatti solitamente caratterizzati da una sola frase principale e da un massimo di due subordinate. Quasi tutti i titoli seguono la struttura sintattica di base (S-V-O) e sono legati dal precedente da paratassi. Ogni lessema che costituisce il titolo svolge quindi un ruolo importante e non può essere omesso, pena una perdita importante del carico informativo del TP. Sfortunatamente per il rispeaker, al raggiungimento di tale obiettivo di completezza si frappone il fattore della compensazione delle immagini, pressoché assente nei titoli, visto che le telecamere inquadrano in maniera più o meno fissa i presentatori. Questo rende le immagini particolarmente statiche oltre che inutili ai fini di un'ottima comprensione del testo da parte dei telespettatori. Per motivi di *marketing* del prodotto, quindi, i giornalisti leggono il testo a una velocità di eloquio particolarmente elevata. Un'ultima difficoltà è rappresentata dalla novità del testo. Il lavoro del rispeaker è quindi, in questa fase, molto delicato in quanto gli viene chiesto di non omettere niente e di seguire l'alto ritmo del TP. C'è però da considerare che né il *software* di riconoscimento del parlato, né il rispeaker

103 Cfr. Halliday 1985.

possono mantenere questo ritmo senza un abbassamento sensibile della qualità del prodotto offerto. Fortunatamente lo sforzo non è da protrarre a lingo nel tempo. Il terzo *step* è costituito da una frase che spiega che i titoli appena letti saranno approfonditi nella fase, o *move*, successivo e da un'espressione di transizione, spesso caratterizzata dalla presa di parola da parte dell'eventuale secondo presentatore o da una frase che lo introduce direttamente o ancora da un ringraziamento funzionale.

Notizie pre-registrate e reportage in diretta

Il *move* introdotto inizia con la ripresa, da parte di un presentatore in studio, del primo titolo, che viene espanso in termini contenutistici. Anche in questo caso, il testo è preparato prima della diretta e letto in tempo reale. Dal punto di vista linguistico, il testo ha le medesime caratteristiche dei titoli, in quanto è grammaticalmente poco complesso e lessicalmente molto denso. Essendo inserito in un contesto introduttivo al servizio che deve essere trasmesso, la velocità di eloquio è inferiore rispetto a quella dei titoli. La transizione da questo breve *step* al successivo può essere garantito in due modi: direttamente da un'espressione formulaica pronunciata dal presentatore che cede la parola al reporter o al servizio (ex: "there is the scene of our sports correspondent, James Munro"); o indirettamente, tramite un evidente calo nella prosodia, indice di fine turno. È così introdotto un *reportage* dal vivo o un servizio pre-registrato. Nel primo caso, il reporter pronuncerà un riassunto introduttivo la cui durata dipende dallo stile personale del giornalista e dalle eventuali domande poste dal presentatore. Linguisticamente, il testo è prodotto oralmente, ma è stato preparato precedentemente. Il reporter può quindi leggerlo direttamente o ricostruirlo sulla base di alcune note. Di conseguenza, sarà grammaticalmente più complesso e lessicalmente meno denso rispetto ai titoli e allo *step* che introduce questo *move*. Essendo di natura 'divulgativa' e basato o meno su un testo che non è stato scritto secondo i canoni del codice scritto, il discorso pronunciato dal reporter è anche più lento e con maggiori elementi dell'oralità che ne rendono più difficile la decodifica, non tanto da parte del presentatore o dei telespettatori, quanto piuttosto da parte dei rispeaker. Alla fine di questa prima parte introduttiva, il presentatore porrà al reporter una serie di domande per delucidare questo o quel passaggio, il quale potrà contare su alcuni appunti non ancora utilizzati o più probabilmente sulla propria conoscenza de fatti. Il testo sarà quindi composto da domande e risposte brevi che si alternano in rapida sequenza. Il *reportage* in diretta può anche essere

costituito da altri generi, quasi sempre pre-registrati. In questo caso, la complessità del testo è notevole con dei cambi di registro e di oratore repentini e imprevisti. Il *move* termina sovente con un'espressione formulaica del reporter che ricorda il suo nome e cognome, la città dalla quale sono trasmesse le immagini e l'emittente per cui lavora (es: "Andrew Hardy, BBC News, Singapore") e con uno scambio di ringraziamenti con il presentatore. Nel secondo caso, è introdotto un servizio pre-registrato. Le caratteristiche del testo saranno alquanto diverse da quelle del *reportage* dal vivo. Innanzitutto, si tratta di un pezzo montato quindi presenterà le caratteristiche di un testo letto (alta densità lessicale e velocità di eloquio, bassa complessità grammaticale, pronuncia chiara e assenza di tratti dell'oralità. Un'altra caratteristica importante del servizio preregistrato è la natura composita delle immagini che accompagna o descrive. Raramente un servizio sarà solo composto, come in molti *reportage* dal vivo, dal giornalista che parla di fronte alla telecamera. Piuttosto, sarà intervallato da interviste, conferenze stampa, immagini in diretta o di repertorio, ecc. le immagini svolgono quindi un ruolo compensativo importante, identificando l'oratore e mostrando e spesso disambiguando ciò di cui si parla. In questo caso il servizio non termina con uno scambio di ringraziamenti, ma con la fine del servizio e con la telecamera che torna a inquadrare il presentatore di turno. Il *move* si conclude ufficialmente con il presentatore che volge lo sguardo dallo schermo da dove sono provenute le immagini del servizio o del *reportage* alla telecamera centrale. Questo è un segnale che il presentatore passerà a un altro servizio o *reportage* o a un altro *move*.

Previsioni meteorologiche

Le previsioni del tempo sono introdotte dal presentatore con una espressione formulaica (ex: "Now, let's have a weather forecast, with Jo Farrow"). L'inquadratura passa al volto o al mezzo busto o busto intero del giornalista che si occupa di meteorologia. Quest'ultimo ringrazia il presentatore e subito inizia a spiegare le previsioni. Il testo è preparato precedentemente e letto in diretta in sintonia con le immagini che passano alla cartina del Regno Unito, per poi stringere su una regione in particolare. Il giornalista prima spiega le previsioni per ogni città e poi riassume brevemente la situazione nella regione (ex: 'So, wet weather in the north-east'). Le immagini passano quindi alla regione successiva e così fino alla fine delle previsioni meteorologiche. In questo *step*, le immagini svolgono un ruolo preponderante. Le parole sono infatti solo una descrizione di quanto compare sullo

schermo (piccole icone che rappresentano tempo soleggiato, parzialmente o totalmente nuvoloso, piovoso, ventoso, nebbioso, grandioso, tempestoso e mare calmo o più o meno agitato posizionate sulla cartina del Regno Unito e indicate dal giornalista man mano che produce il TP). Ecco quindi che il testo è particolarmente rapido, con una forte presenza di coordinate e un'alta densità lessicale. Il *move* in questione può sia terminare con un'espressione formulaica da parte del giornalista che ha appena finito di parlare e che quindi cede la parola al presentatore, sia con il presentatore che riprende la parola ringraziando il giornalista delle previsioni.

Sommario

Il *sommario* è un interessante *move* enunciato dai presentatori e solitamente posizionato dopo le previsioni meteorologiche. Viene introdotto solitamente da un'espressione formulaica (ex: "Some news summary first") e la sua struttura è simile a quella dei titoli. Esso è infatti una lista delle notizie che sono state approfondite nel corso del giornale e dei giornali precedenti sotto forma di brevi titoli (comunque più lunghi rispetto ai titoli). Si tratta di testo scritto per essere letto, caratterizzato da bassa complessità grammaticale, da alte densità lessicale e velocità di eloquio e dall'assenza di compensazione da parte delle immagini. Se il sommario può essere paragonato ai titoli dal punto di vista della struttura, la sua funzione è tuttavia piuttosto diversa. Il sommario non vuole infatti essere un'introduzione alle notizie, ma uno stato dell'arte delle notizie del giorno. Se nell'edizione che si sta per concludere sono state date alcune notizie, nel sommario saranno riassunte le medesime oltre che quelle più importanti che sono state date nel corso delle edizioni precedenti. L'ultimo *step* è caratterizzato da una transizione garantita da un'espressione formulaica pronunciata dai presentatori che ricorda ai telespettatori che le notizie, i titoli e il sommario dell'edizione che sta per terminare sono disponibili sull'archivio interattivo della BBC. Dopo questo *step* possono seguire le notizie dell'ultimo minuto o l'ultimo *move*.

Immagini di chiusura

Non sono sempre proiettate, in quanto notizie dell'ultimo minuto potrebbero impedirne l'emissione a favore del jingle di apertura del programma successivo.

4.6 Analisi strategica di *BBC News*

La componente verbale di un'edizione di *BBC News* svolge un'importanza fondamentale, in quanto il programma ha uno scopo informativo. È proprio per questa ragione, che il RNID (*Royal National Institute for Deaf and Hard-of-hearing People*) ha chiesto agli *Access Services*[104] della BBC di sottotitolare questo programma *verbatim* garantendo così ai telespettatori non udenti piena accessibilità[105]. Tuttavia, la sua importanza varia a seconda del ruolo che la componente visiva svolge in ogni singolo fase e sotto-fase. D'altronde, come si è visto nei capitoli iniziali, molti ricercatori sono d'accordo nel sostenere la posizione opposta, quella cioè della riformulazione in nome di una maggiore leggibilità e udibilità del sottotitolo. Ofcom, l'autorità britannica preposta alla vigilanza in materia di competizione e rispetto delle regole da parte del comparto dell'industria della comunicazione[106], sembra trovare un compromesso, proponendo di "reduce the amount of text by reducing the reading speed and removing unnecessary words and sentences" (ITC 1999: 27). Ciò che è importante, è infatti la resa di quello che Ofcom definisce "the whole meaning" (*ibidem*). Per fare questo, "subtitles should contain a reasonable percentage of the words spoken" (*ibidem*). Inoltre, quelle che definisce 'idea units', cioè le unità concettuali, dovrebbero "appear as a good percentage of the original" (*ibidem*). E infine, "'idea units' which are unnecessary or different from the original" (*ibidem*) possono essere omesse.

Grazie all'analisi approfondita di otto edizioni di BBC News in generale e delle singole fasi e sotto-fasi in particolare, è stato possibile evidenziare le strategie messe in atto dai rispeaker e riscontrare il grado di aderenza dei rispeaker alle indicazioni del RNID e di Ofcom.

4.6.1 Metodologia

La metodologia adottata in questo passaggio della ricerca si è basata sul concetto di 'idea unit' proposto da Ofcom già brevemente introdotto. Nonostante quest'ultimo definisca

104 All'epoca in cui il programma analizzato è stato trasmesso, i sottotitoli erano ancora forniti dagli Access Services della BBC. Oggi, i sottotitoli trasmessi dai programmi della BBC sono forniti da RedBee Media, una società privata che ha assunto il personale degli Access Services della BBC e che ha acquistato lo stesso software.

105 La prima giornata di studio internazionale sulla sottotitolazione intralinguistica in tempo reale, svoltasi a Forlì il 17 novembre 2006, ha approfondito un dibattito che da qualche tempo sta interessando l'ambito accademico, ovvero il significato di accessibilità da parte di un qualsiasi utente a un prodotto audiovisivo. Nel caso dei sordi preverbali, le due teorie contrapposte sostengono da una parte la resa verbatim del TP (cfr. Mereghetti 2006), dall'altra una riformulazione principalmente sintattica dello stesso (cfr. Eugeni 2007).

106 Ofcom è la principale autorità nel settore dell'accessibilità alla TV del Regno Unito. Regola la politica della BBC nel settore della sottotitolazione in pre-registrato, della sottotitolazione in tempo reale, dell'audio descrizione e dell'interpretazione simultanea. Le sue linee guida sono un punto di riferimento per tutte le emittenti britanniche.

l'unità concettuale come la porzione di testo "where a proposition or key information is given" (1999: 27), in pratica l'identificazione scientifica di una *idea unit* è una questione aperta. Si dovrebbero far combaciare i confini dell'unità concettuale con quelli di un'unità grammaticale, come lessema, predicato o frase? O piuttosto con quelli di unità testuale e quindi, nel caso specifico, un sottotitolo o un gruppo di sottotitoli che coprono un intero periodo? Consideriamo l'esempio seguente:

> TP[107]: London have to focus on the key 45' presentation tomorrow to all the IOC members
>
> TA: London have to focus on the key 45 minutes presentation (…) to (…) the IOC[108]

È possibile affermare senza timore di essere smentiti che le *idea unit* del TP siano state trasferite nel TA? Se prendiamo in considerazione il significato globale delle due frasi, cioè l'obbligo da parte di Londra di concentrarsi sulla presentazione al Comitato Olimpico Internazionale, sembra non ci siano dubbi sull'esattezza della resa. Lo stesso dicasi per la resa di 'all the IOC members' con 'IOC'. IOC è l'acronimo per Comitato Olimpico Internazionale, l'organismo chiamato a prendere le principali decisioni in materia di Giochi Olimpici. Come è ovvio, è composto da un certo numero di membri che, insieme, agiscono in nome dell'organismo a cui appartengono. Visto che una presentazione non può che essere fatta di fronte a delle persone, generalizzare, così come è stato fatto nel TA dell'esempio, non comporta alcuna modifica al senso del sintagma in questione. Lo spettatore che ha seguito la notizia fin dall'inizio, inoltre, inferisce facilmente l'informazione veicolata anche dalla semplice parola 'presentation', visto che la presentazione di fronte al COI non è un'informazione nuova.

---

107 TP sarà di seguito utilizzato per indicare il testo pronunciato dall'oratore sullo schermo e TA per indicare i sottotitoli. Vista l'impossibilità di accedere al TM, si è preferito non considerare questa versione onde evitare inutili previsioni. Quando si avrà una certa sicurezza circa la versione del TM, vi si farà accenno in fase di discussione dei dati. Il carattere sottotitolato sarà utilizzato, nel TP, per indicare le porzioni di testo che sono state omesse o semplicemente alterate e, nel TA, per indicare la rispettiva resa.

108 Tutti gli esempi all'interno del presente lavoro sono tratti dalla trascrizione delle sedici edizioni di BBC news e dei relativi sottotitoli componenti il corpus analizzato. La trascrizione di una edizione di BBC News e dei relativi sottotitoli è stata riprodotta in allegato.

Ma che dire dell'omissione di 'tomorrow'? Si tratta, anche in questo caso, di un'informazione nota ai più, ma con uno suo specifico carico semantico. Che cosa succede quando la si omette? Se, da una parte, lo spettatore attento e informato ringrazia, dall'altra, il linguista sottolinea l'inevitabile e irreparabile perdita. Indipendentemente dalla natura della perdita, al ricercatore resta il dubbio sulla sua catalogazione: deve essere trattata come l'omissione di un'unità concettuale o piuttosto come l'effetto di una strategia di riduzione di un'unità concettuale più grande e troppo complessa dal punto di vista informativo? Alla luce di queste considerazioni, nel seguente lavoro è stata adottata la seguente distinzione:

- micro-unità: tutte le componenti portatrici di significato all'interno di una frase (sintagmi, avverbi, incisi, ecc.) e che contribuiscono al significato generale di un'unità più vasta e completa;
- macro-unità: ogni tipo di frase che fornisce un'informazione di senso compiuto, indipendentemente dal fatto che sia principale o secondaria, coordinata o subordinata.

Fatta questa indispensabile distinzione, il secondo passo è stata la trascrizione del TP e la sua conseguente segmentazione in macro-unità concettuali. Il TA è stato poi allineato con il TP e infine sono stati messi a confronto i due testi. La prima tappa dell'analisi è stato l'isolamento nel TP delle macro-unità non rese nel TA. Delle macro-unità rimanenti sono state esaminate le strategie traduttive: alcune di loro sono state ripetute fedelmente nel TA, mentre altre sono state alterate nella forma a livello di micro-unità o di morfema grammaticale[109]. Queste ultime sono state infine analizzate mediante una speciale tassonomia largamente ispirata alla tassonomia di Gambier (2006) per l'analisi delle strategie traduttive utilizzate dai sottotitolatori (interlinguistici) professionisti di testi pre-registrati (cfr. cap. precedente). Le macro-strategie prese in considerazione sono:

- espansione: una macro-unità concettuale è stata spiegata o disambiguato attraverso l'uso di più caratteri rispetto al TP;
- riduzione: una macro-unità ha subito un'omissione parziale oppure una riformulazione totale o parziale (compressione);

109 In questo testo si farà un uso improprio, ma utile, della distinzione fatta in linguistica tra morfema grammaticale e morfema lessicale. Il primo identificherà tutte quelle parole che hanno una mera funzione grammaticale o testuale (articoli, preposizioni, connettori, intercalari, ecc.); il secondo tutte quelle parole portatrici di un valore semantico importante (nomi, verbi, avverbi, aggettivi, ecc.).

- errori: una micro-unità è stata alterata da un errore umano o del software di riconoscimento del parlato.

Da sottolineare, è la natura quantitativa della tassonomia. Come si vede, infatti, il criterio per l'attribuzione di una strategia alla categoria dell'espansione o della riduzione non è altro che, rispettivamente, l'aumento o la diminuzione del numero di caratteri; nell'ultimo caso, l'alterazione fisica ingiustificata dei caratteri. Si tratta quindi di un discrimine molto concreto che non concede alcuno spazio alla speculazione. Tuttavia, all'interno di queste categorie è innegabile la presenza di due sottocategorie abbastanza evidenti. Ignorarle avrebbe rappresentato un'imperdonabile negligenza.

Ecco quindi che un'ulteriore suddivisione è stata presa in considerazione, tra le strategie semantiche, quelle cioè che hanno comportato un'aggiunta o una riduzione di significato rispetto al TP; e quelle non-semantiche, la cui applicazione ha condotto a un mero cambiamento del numero di caratteri del TA, senza peraltro intaccare il significato generale o quello delle singole micro-unità. Quanto agli errori, va detto, prima di tutto, che pur comportando un aumento o una diminuzione del numero dei caratteri del TP, essi non possono essere inseriti in una delle due categorie precedenti, perché sono delle azioni non volute e spesso indipendenti dalle intenzioni del sottotitolatore. Eppure, gli errori sono stati classificati come una forma di alterazione del TP, in quanto i telespettatori ricevono qualcosa di diverso rispetto a quanto hanno ricevuto acusticamente i normoudenti. L'unica distinzione che va fatta è quindi tra quelli che possono essere imputabili a una *défaillance* del rispeaker e quelli che invece possono essere considerati errori del software. Non si pone invece la questione dell'alterazione semantica, in quanto inutile ai fini del presente lavoro.

4.6.2 Analisi strategica generale

Il testo analizzato è composto di 1101 macro-unità concettuali. 208 sono state completamente omesse (18.9% delle macro-unità). 555 sono state ripetute e fedelmente trascritte (50.4% delle macro-unità). Le rimanenti 338 macro-unità sono state alterate (30.7% delle macro-unità). In particolare, di queste 338 macro-unità, il 7.1% è il risultato di strategie di espansione, l'89.1% di riduzione, il 3.4% di errori del software o del rispeaker e lo 0.4% di correzione cosciente da parte dell'operatore. Queste ultime strategie non sono

state catalogate a parte, ma come forme di espansione o compressione semantica a seconda dell'esito quantitativo dell'operazione.

Omissioni

Le macro-unità che sono state omesse sono per lo più delle frasi contenute in passaggi più rapidi della media (59.8% delle macro-unità omesse), ma anche frasi enunciate in un momento in cui i sottotitoli hanno un *décalage* superiore alla media (12.8% delle macro-unità omesse) o espressioni formulaiche (27.4% delle macro-unità omesse). Per quanto riguarda le macro-unità perché contenute in passaggi più rapidi della media, bisogna notare che la velocità media di un'edizione di mezz'ora di *BBC News* è di 183 parole al minuto. Qualora uno *step* fosse sensibilmente più rapido della media, il numero totale di unità omesse aumenterebbe. Quello che colpisce è che quasi mai le unità omesse sono veramente le unità con il minore carico semantico. Non sono per forza meno importanti di altre che sono state tralasciate. L'unica spiegazione per la loro omissione sembra quindi essere l'accidentale presenza in uno *step* più rapido della media. Nell'esempio che segue, il reporter si trova a Gleneagles, un paese scozzese dove dovrà tenersi un G8 e intervista diversi passanti chiedendo loro opinioni circa le misure di sicurezza che sono state adottate dalla polizia locale. Uno degli intervistati dice:

> TP: It seems too much for what is going on. No-one at Gleneagles is gonna know what's happening. <u>All these big marches and things I don't think it is useful at all.</u>
>
> TA: It seems too much for what is going on. No-one at Gleneagles will know what is going on <u>(…)</u>

In questo passaggio la velocità media è di 191 parole al minuto. Un altro aspetto che potrebbe influire sulla scelta del rispeaker di omettere la frase in questione è la velocità di presa di parola degli oratori, otto secondi in media. Visto che il *décalage* medio tra il TP e i rispettivi sottotitoli è di 4.3 secondi, è facile intuire come il rispeaker sia stato obbligato a ridurre il TP. Tuttavia, la sua scelta è caduta su un aspro commento dell'intervistato. Perché proprio questa frase e non una delle altre due? Perché il rispeaker ha optato per la sua eliminazione e non per una sintesi? La risposta a questi quesiti sembra ovvia: l'omissione è

più rapida e più efficace in termini spaziali e temporali. Quanto alla scelta dell'omissione dell'ultima frase in luogo di una delle altre due, le ragioni possibili sono almeno due: 1) il rispeaker si rende conto che il TP è composto da due opinioni personali simili tra loro ("It seems too much for what is going on" e "All these big marches and things I don't think it is useful at all") e da un commento generico ("No-one up at Gleneagles is gonna know what's happening"). Il rispeaker decide quindi di omettere la seconda delle due opinioni perché ridondante e inutile alla comprensione generale del testo; 2) nel momento in cui il rispeaker sta terminando di ripetere la seconda macro-unità, il giornalista inizia a commentare le interviste le interviste. Il rispeaker decide quindi di omettere l'ultima unità in maniera da non sovraccaricare la propria memoria a breve termine o il carico semantico del TA. Probabile è anche una combinazione delle due ipotesi.

Un altro esempio di unità concettuale omessa perché contenuta in un passaggio più rapido della media è il seguente:

> TP: I feel very proud to be a British athlete and a Paralympian. There is not another country in the world with so much attention to Paralympian athletes as the UK both in terms of financial support, and of media coverage.
>
> TA: I feel very proud to be a British athlete and a Paralympian. (…) Both in terms of financial support, and of media coverage.

Questo esempio è tratto dalla conferenza stampa della squadra britannica incaricata di promuovere Londra come città ospitante i Giochi Olimpici del 2012. La conferenza è appena iniziata e i membri della squadra si stanno presentando e stanno rivolgendo delle parole di benvenuto ai giornalisti. La presa di parola è molto rapida (15.2 secondi ognuno). Quando uno dei membri, Dame Tanni Grey Thompson, prende la parola, parla a 169 parole al minuto, una velocità che, pur essendo inferiore rispetto alla media, è più rapida rispetto alla media dello *step* in questione (148 parole al minuto). Il rispeaker è disorientato da questo repentino cambiamento e quindi decide di omettere alcune porzioni del TP. Anche in questo caso, la decisione sull'unità da omettere sembra arbitraria. Lo spettatore che legge tali sottotitoli ha l'impressione che qualcosa manchi rispetto al TP anche se un certo nesso logico può essere inferito. In ottica contrastiva c'è poi da sottolineare la mancata resa di un

concetto importante: il Regno Unito è il paese che dedica più attenzione agli atleti disabili rispetto agli altri paesi del mondo.

Un altro caso di omissione di macro-unità concettuale è il secondo, vale a dire il caso di unità eliminate perché il rispeaker è in ritardo nella produzione di sottotitoli, malgrado la velocità di eloquio del passaggio in cui l'unità in questione è stata omessa non sia superiore alla media. Le ragioni di questo tipo di omissioni sono essenzialmente due: 1) il testo è particolarmente denso (il numero di morfemi lessicali per ogni macro-unità concettuale è superiore alla media, che si attesta attorno a 4.7). Nel caso dei titoli, il numero di morfemi lessicali per macro-unità sale a 6.1 e la velocità media supera di poco la media generale (186 parole al minuto). Nella fase produttiva del rispeakeraggio, il carico cognitivo diventa sempre più difficile da gestire e il divario tra il TP e il TA aumenta inevitabilmente. Il seguente esempio dimostra chiaramente che nessun'altra motivazione è possibile:

> TP: Liverpool have received a £32 offer from Chelsea for the star, which they say they will turn down. West Ham have been busy on the transfer of Cardiff defender Danny Gabbidon, one of three players to join West Ham today. And the Lions have beaten Auckland, but only just.
>
> TA: Liv(…) were received a £32 offer (…) for the star, which they say they will turn down. (…) (…) And the Lions have beaten Auckland, but only just.

Mentre la compressione è un *escamotage* per ridurre quantomeno lo spazio che andrà a occupare il riferimento alla squadra di calcio della città di Liverpool[110] e la prima omissione ('from Chelsea') non è così rilevante ai fini della comprensione del TP, benché il dato omesso sia informativamente di peso, la seconda e la terza omissione sono un chiaro esempio di quanto detto precedentemente: la macro-unità omessa corrisponde a un'informazione che sarà recuperata dal telespettatore solo nel *move* delle notizie o dei *reportage*. Nel *move* in questione, l'unità è stata omessa perché le informazioni iniziavano a sovraccaricare la memoria a breve termine del rispeaker. Una della informazioni elencate è

110 In realtà, è possibile che siano stati utilizzati degli accorgimenti tecnici, come housestyle, short form o macro di dettatura, per produrre un TA dato (in questo caso la parola Liv, che è il termine con cui ci si riferisce familiarmente alla squadra di calcio Liverpool FC), tramite un testo di mezzo decisamente più lungo (ad esempio, rispettivamente: 'liverpool', 'squadra 3', 'liverpoolmacro'). In tutti questi casi, il guadagno rispetto al TP sarebbe solo in termini di velocità di lettura del TA, quindi di spazio, non di tempo per produrlo.

stata quindi eliminata *in toto*; 2) l'ingresso video svolge un ruolo decisivo e quindi la componente verbale non è così rilevante ai fini della comprensione del testo. È sicuramente preminente la multimodalità del testo in generale e la sincronia con le immagini in particolare. Da questo deriva la ovvia necessità per il rispeaker di ridurre il più possibile il divario tra la produzione del TP e del TA. Un caso emblematico è rappresentato dalle previsioni meteorologiche, in cui il testo descrive le immagini, già sufficientemente iconiche, che si susseguono in rapida successione. In questi casi (così come nelle telecronache sportive), l'omissione non comporta una perdita totale di significato. Ecco quindi che nei sottotitoli intere macro-unità sono cancellate con l'unico obiettivo di non ritardare ulteriormente la comparsa dei sottotitoli rispetto al TP e di non far apparire quindi delle informazioni riferite a una regione sotto la mappa e le relative icone di un'altra regione. Un chiaro esempio di quanto appena affermato è il seguente:

> TP: In the late morning we'll begin to see the clouds increasing across northern England as this wet weather moves in, moving across Cumbria by lunchtime
>
> TA: (…) We'll begin to see the clouds increasing across northern England (…)(…)

Mentre per la prima omissione, l'immagine compensa l'informazione verbalmente perduta (un orologio digitale posizionato in alto a destra dello schermo indica le ore della tarda mattinata), la seconda e terza omissione sono solo parzialmente compensate da un'icona che mostra una nuvola piena di pioggia posizionata sopra le regioni settentrionali della Gran Bretagna. Inoltre, coloro che preferiscono seguire i sottotitoli senza aver prima dato uno sguardo alle immagini si perderanno una parte importante delle informazioni, specialmente se interessati dalle previsioni.

La terza tipologia di macro-unità omesse concerne le espressioni formulaiche. In questo caso, si potrebbe parlare di una vera e propria strategia, in quanto l'operazione in questione è applicata da tutti i rispeaker della BBC in tutti i contesti in cui un'espressione formulaica considerata ridondante è presente. Il caso più eclatante riguarda la transizione tra la fine del *reportage* dal vivo e lo *step* successivo. Il reporter infatti spesso conclude con una frase senza verbo in cui viene indicato prima il suo nome e cognome, l'emittente per la quale lavora e infine la città dalla quale sta parlando. Nel testo analizzato sono state

riscontrate 57 espressioni formulaiche, tutte sistematicamente omesse, anche qualora il passaggio in cui si trovavano fosse sensibilmente più lento rispetto alla media. La ragione di questo tipo di omissione risiede nella natura semiotica del testo. In particolare, per un determinato lasso di tempo dall'inizio del *reportage*, un sottopancia in sovrimpressione fornisce lo stesso tipo di informazioni. Sottotitolare l'espressione in questione risulterebbe in un'inutile ridondanza. Donde l'inevitabile omissione da parte dei rispeaker.

Un ultimo caso abbastanza comune di omissioni di macro-unità concettuali, trasversale alle prime due tipologie analizzate (omissioni in un passaggio rapido e omissioni in un momento in cui i sottotitoli sono in ritardo), è quello delle fusioni, vale a dire quelle operazioni di 'taglio e cucito' per cui una proposizione o porzione di proposizione è eliminata e la proposizione (o porzione di proposizione) precedente è giustapposta o unita a quella successiva tramite elementi di coesione, come nel caso seguente:

> TP: ...a female pilot who wanted to work part-time to look after her baby daughter. She went to court for indirect sexual discrimination against the airline after her request to cut her hours by half was turned down.
>
> TA: ...a woman who wanted (…) (…) to cut her hours by half (...).

In quei casi in cui la fusione non crea problemi di coesione e di coerenza, anche se interi pezzi di informazione sono omessi, la strategia non può che essere ben accetta, in quanto il beneficio in termini di riduzione del divario tra TP e TA è notevole. Nel caso appena citato, la coesione è salva, ma salta subito all'occhio la mancanza di informazioni essenziali a definire l'identità professionale della donna, la natura professionale delle ore e la volontà, altrimenti apparentemente frivola, della donna di dimezzare il proprio orario di lavoro.

Ripetizioni

Più della metà delle macro-unità concettuali (il 50.4% del totale delle macro-unità) è stato fedelmente ripetuto e ortograficamente trascritto, elementi dell'oralità compresi. Questo dato dimostra chiaramente come la preoccupazione maggiore dei rispeaker della BBC sia quella di ripetere il TP il più fedelmente possibile. Questo per due ragioni

essenziali: le associazioni in difesa dei telespettatori sordi chiedono una resa possibilmente *verbatim* nei sottotitoli, quindi l'uso della ripetizione come stretegia traduttiva soddisfa appieno le loro richieste; per ottenere il massimo risultato (la resa *verbatim* del TP) con il minimo sforzo mentale e cognitivo, il rispeaker può limitarsi all'indispensabile: ripetere fedelmente il TP, prestando attenzione alla pronuncia, introducendo la punteggiatura e cambiando i colori a seconda degli oratori. Tuttavia, questo è possibile soltanto se, nel TP, la velocità d'eloquio, la densità lessicale, la presa di parola, il carico semantico e la compensazione video sono abbastanza basse da permettere quanto richiesto dal minimo indispensabile. Tuttavia, un tale concorso di circostanze è alquanto aleatorio. Inoltre, nemmeno una simile situazione ideale garantisce una resa *verbatim* del TP per un periodo superiore alle cinque macro-unità concettuali. Questo probabilmente perché, così come succede in interpretazione simultanea, mantenere lo stesso ritmo del TP è un'attività mentalmente dispendiosa se protratta sul medio-lungo periodo. Per amor del vero, c'è infine un'attenuante da considerare derivante dal concetto stesso di verbatim e di trascrizione ortografica: per quanto un giornalista si sforzi di parlare in maniera chiara e pulita, il suo eloquio sarà inevitabilmente caratterizzato da fenomeni tipici dell'oralità, come false partenze, autocorrezioni, ripetizioni e contrazioni. Se a operazioni di pulizia di questi tratti nella produzione del testo di mezzo aggiungiamo tutte le piccole operazioni di espansione e riduzione quantitativa delle unità concettuali tramite l'aggiunta o l'eliminazione di morfemi meramente grammaticali, risulta chiaro come, in realtà, il numero di macro-unità ripetute in condizioni esterne favorevoli possa superare quota cinque. Inoltre, malgrado tutti gli sforzi possibili nella produzione del testo di mezzo da parte dei rispeaker, i software di riconoscimento del parlato non sono ancora tecnicamente in grado, da una parte, di gestire le inevitabili imperfezioni nella pronuncia da parte dei rispeaker (una sillaba non perfettamente pronunciata, due parole non perfettamente distinte l'una dall'altra, un respiro più affannoso del normale, ecc.), dall'altra, di disambiguare grazie al contesto e di fare distinzione tra un monosillabo e uno foneticamente simile. Questi due aspetti negativi della tecnologia del riconoscimento automatico del parlato si traducono nell'introduzione di errori nella fase di produzione del TA, che, nel corpus in esame, è del 3.4% delle macro-unità alterate, ovvero l'1.15% del corpus analizzato.

Espansioni

Le espansioni e le riduzioni contano per meno di un terzo delle strategie. Come già detto, finché le condizioni lo permettono, i rispeaker tendono a ripetere il TP nella maniera più fedele possibile e a omettere soltanto quelle proposizioni considerate come ridondanti quali le espressioni formulaiche, i ringraziamenti, eccetera. Quando le circostanze cambiano (perché il rispeaker è in ritardo rispetto al TP principalmente a causa di un aumento repentino della velocità di eloquio media o della densità lessicale rispetto alla media o ancora perché le immagini impongono una riduzione del divario tra il TP e il TA), allora il rispeaker, nel rispetto di un certo ritmo di produzione dei sottotitoli (circa 180 parole al minuto), è spinto a omettere intere proposizioni, spesso le ultime che compongono uno *step* o quelle che si trovano all'interno di periodi particolarmente complessi dal punto di vista sintattico.

In questo quadro, è facile intuire come non vi sia grande spazio per le espansioni (7.1% delle strategie di alterazione), che risultano essere pertanto la strategia meno impiegata (cfr. Gambier 2006). La ragione principale di questo fenomeno è la difficoltà, intrinseca al processo di rispekeraggio, di aggiungere informazioni o di spiegare concetti particolarmente condensati in tempi così ristretti. Tuttavia, l'espansione di macro-unità è talvolta indispensabile per recuperare delle informazioni precedentemente omesse o per esplicitare un concetto ritenuto poco chiaro. A queste strategie, che costituiscono un terzo delle strategie di espansione, si aggiungono quelle strategie di espansione grammaticale volte a una resa maggiormente adatta allo scritto di forme spesso abbreviate (ad esempio: it's → it is, gonna → going to, ecc.). Oltre all'obiettiva difficoltà di aggiungere testo in condizioni così vincolanti dal punto di vista temporale, un'altra ragione per quest'ultimo tipo di espansione riguarda la difficoltà, nella resa, di ripetere esattamente ogni parola del TP così come è stata prodotta dall'oratore. Il rispeaker, infatti, dopo aver ascoltato le prime parole di una frase, capisce il senso generale della stessa e inizia a ripetere le parole in base al ruolo grammaticale e semantico che attribuisce *hic et nunc* a ciascuna di esse e non meramente al suono che esse producono. Di conseguenza, così come accade in maniera più evidente nell'interpretazione simultanea (per ragioni non solo cognitive, ma anche legate alla natura sintattica della lingua di arrivo), il rispeaker introduce, automaticamente e quasi inconsapevolmente, nella produzione del testo di mezzo, sinonimi morfo-sintattici ('we will' per 'we'll', 'going to' per 'gonna', 'will not' per 'won't', 'that' al posto di una giustapposizione, 'thank you very much' per 'thanks' o 'thank you', ecc.), elementi di

coesione ('and', 'but', 'well', ecc.), dislocazione delle parole o dei sintagmi (tematizzazioni, rematizzazioni, ecc.) e altri morfemi grammaticali che comportano un aumento del numero di caratteri nel TA rispetto al TP, ma non una vera e propria alterazione semantica.

Per quanto riguarda la loro rappresentatività nel totale delle strategie di espansione, l'espansione semantica è perlopiù costituita da operazioni a livello lessicale o sintagmatico, mai drastico. La ragione, ancora una volta, sta nell'impossibilità di aggiungere frasi, a meno che queste ultime non siano state omesse interamente in una fase precedente e ripetute in un secondo momento. La strategia più utilizzata è l'esplicitazione (18%), un termine iperonimico che comprende micro-strategie come la disambiguazione di un acronimo ('IOC' diventa 'the International Olympic Committee'), il completamento di un nome ('Seb' e 'Steve Redgrave' diventano rispettivamente 'Sebastian Coe' e 'Sir Steven Redgrave'), l'attenuazione di una generalizzazione ('a third of the members' diventa 'about a third of the members'), la spiegazione di un concetto evidentemente ritenuto poco chiaro ('spent time on buses' diventa [*sic*] 'spent time sitting in buses'). Un'altra importante strategia è la sinonimia lessicale, che rappresenta il 14% delle espansioni ('are not allowed' per 'cannot', 'request' per 'ask', 'comments' per 'remarks', 'huge amount' per 'galaxy', ecc.). Degno di nota è anche il caso del rispeaker che corregge il TP (2.5% delle espansioni). Benché sembra sia "strictly forbidden" (cfr. Marsh 2005), può succedere che un rispeaker corregga automaticamente un oratore, perché più semplice dal punto di vista cognitivo ('changes is' è stato reso con 'changes are') o perché, nel caso di parole inventate o pronunciate male, il software di riconoscimento del parlato non le riconoscerebbe e quindi non le trascriverebbe correttamente. Infine, degno di nota, è l'unico caso di espansione a livello frastico:

TP: Gunmen have settled a ferocious battle with police

TA: Reports say that people have started to fight with the police

Questo esempio è tratto dal sommario, un *move* abbastanza semplice da rispeakerare visto che il rispeaker ha già sentito e compreso il contenuto di quanto sarà detto. In questo caso, il rispeaker non solo introduce un elemento di *hedging*[111] non presente

111 Hedging è una strategia di presa di distanza da quanto si dice, tramite cui l'oratore non si assume la piena responsabilità della veridicità di quanto asserito (cfr. tra i primi Lakoff 1972 e Halliday e Hasan 1985).

nell'originale, ma ha anche semplificato lessicalmente il TP senza che lo stesso sia impoverito semanticamente. Si tratta di un caso isolato, ma è una dimostrazione del fatto che la riformulazione non è soltanto possibile, ma offre anche grandi possibilità in termini di accessibilità al TP. Restano da dimostrare l'effettiva applicabilità di tali strategie sul lungo periodo oltre che l'efficacia sul pubblico di destinazione.

Quanto all'espansione non semantica, la sinonimia morfo-sintattica costituisce il 36% delle espansioni, l'introduzione di elementi di coesione il 26% e infine le dislocazioni il 2%. Quest'ultimo dato è determinato essenzialmente da due fattori: 1) le dislocazioni spesso comportano una modifica del senso globale della frase in cui si trovano e quindi vengono calcolate anche nella categoria espansione semantica; 2) vista l'inutilità di dislocare senza ottenere una maggiore chiarezza del TA, la presenza delle dislocazioni in questa categoria è probabilmente dettata dalla sola summenzionata ragione dell'automatismo del rispeaker nel rendere grammaticalmente e semanticamente, ma non foneticamente, ciascuna parola o ciascun sintagma del TP.

Riduzioni

La riduzione è la strategia più utilizzata così come viene riconosciuto dalla maggior parte degli studiosi in materia (cfr. Chaume 2004, Bruti e Perego 2005, Gambier 2006). Come già accennato precedentemente, nel corpus analizzato corrisponde all'89.1% delle strategie di alterazione utilizzate e si scompone in due micro-strategie, l'omissione e la compressione. La prima, omonima di una delle tre macrostrategie (omissione, ripetizione, alterazione), consiste nell'eliminazione di una o più micro-unità all'interno di una macro-unità concettuale. Come nel caso delle espansioni, l'omissione di micro-unità può essere semanticamente più o meno importante. Ecco quindi che anche questa micro-strategia è ulteriormente suddivisa in omissione semantica e omissione non semantica. L'omissione semantica consiste nella perdita di lessemi contenutisticamente importanti e non inferibili dal resto del contesto. L'omissione non semantica è invece la cancellazione quasi automatica di tratti dell'oralità ('you see', 'it's, it's, it's very interesting', 'do you think that much protesters, many protesters will come to court', ecc.), di morfemi grammaticali inutili ('that' → ø, 'we will' → 'we'll', 'thank you' → 'thanks', ecc.) e di parole o micro-unità ridondanti dal punto di vista multimodale.

La seconda, la compressione, è un'operazione che consiste nel ridurre il numero di caratteri del TP senza peraltro omettere alcunché o ricorrere a strategie di 'taglio e cucito'. Come nel caso precedente, anche qui la distinzione tra compressione semantica e non semantica è degna di nota. La compressione semantica è composta da micro-strategie come la riformulazione, la sinonimia verticale e orizzontale e la focalizzazione. Queste sono anche le micro-strategie che compongono la compressione non semantica. Il discrimine tra l'una è l'altra categoria è lo scarto di significato che viene a crearsi nel caso della compressione semantica, assente invece nella compressione non semantica.

Per quanto riguarda l'incidenza nel corpus analizzato, l'omissione semantica rappresenta il 5.2% delle strategie di riduzione. Nella maggior parte dei casi (69% delle omissioni semantiche), essa implica la cancellazione di parti di una micro-unità concettuale e quindi una generalizzazione della macro-unità concettuale di cui è parte, come nell'esempio che segue:

> TP: Scientific studies have found links to an increased risk of asthma and allergies, which laboratory tests have linked to allergies such as asthma and even certain types of cancer.
>
> TA: (…) Studies have found links to an increased risk of asthma and allergies, which (…) tests have linked to allergies such as asthma and even certain types of cancer.

Tuttavia, molti sono anche i casi di micro-unità interamente cancellate (31% dei casi). La strategia spesso implica una perdita totale di informazioni non compensate dalla componente video (78% delle micro-unità semanticamente omesse). Per amor di precisione, la micro-unità eliminata è spesso posizionata alla fine della proposizione o addirittura dello *step* in cui è compresa. Questo potrebbe significare che il rispeaker ripete il TP parola per parola finché la macro-unità pronunciata dall'oratore del TP non finisce. A questo punto, se la macro-unità è seguita da una pausa che permette la fine della resa da parte del rispeaker, allora, la macro-unità sarà completamente trasferita nel TA, se, invece, la macro-unità è posizionata alla fine di uno *step* o è seguita da un'altra macro-unità concettuale particolarmente densa dal punto di vista del contenuto o ancora da delle immagini che

esigono una riduzione repentina del divario tra il TP e il TA, il rispeaker si vedrà obbligato a tagliare per non sovraccaricare la propria memoria a breve termine:

> TP: The Government has been an enthusiastic supporter of Internet for business and education
>
> TA: The Government has been an enthusiastic supporter of Internet for business (…)

Nel 22% dei casi di omissione semantica di micro-unità, la micro-unità in questione può essere facilmente inferita dal contesto multimodale (audio e/o video, verbale e/o non verbale). Si tratta spesso di informazione ridondante, o che il rispeaker percepisce come tale, e che è omessa dal rispeaker, a profitto della sua memoria a breve termine e del carico cognitivo dei telespettatori. Prendiamo il caso seguente:

> TP: The controversy started when an opposition website published this photograph, alleging the bearded man was Mr Ahmedinejad
>
> TA: The controversy started when an opposition website published this photograph, alleging the (...) man was Mr Mahmoud Ahmedinejad

Nell'esempio, l'intervistato mostra la fotografia di tre uomini, due dei quali con una lunga barba. Uno di questi due uomini, il presidente iraniano Ahmadinejad, è cerchiato di rosso. Il rispeaker ha qui optato per l'omissione della specificazione 'bearded', che peraltro è totalmente inutile visto che nella foto ci sono due uomini barbuti, in quanto il riferimento al presidente iraniano nella foto è stato totalmente compensato dalla componente video. In questo caso, nonostante la strategia in questione implichi semanticamente una perdita, quest'ultima è compensata dalla componente video e quindi la perdita non è più rilevante ai fini della ricezione del prodotto audiovisivo. È questo uno di quei casi che Rundle (2008: 107) definisce di collaborazione tra le due componenti diamesiche del prodotto audiovisivo. Sempre secondo Rundle, il ruolo del sottotitolatore è di cogliere questi momenti di collaborazione in anticipo e di fornire nei sottotitoli "an interpretative key which is added to

the original dialogue (and all the other communicative channels of the film) in a form that allows us to absorb both at once and use one to understand the other" (*ibidem*).

Quanto all'omissione non semantica, essa rappresenta il 46% delle strategie di riduzione ed è la strategia di riduzione maggiormente utilizzata. Le ragioni sono alquanto evidenti: se il TP è orale-orale (non scritto per essere letto ad alta voce, ma preparato solo sotto forma di note, mentalmente o non preparato affatto[112]), tutti i tratti dell'oralità (in particolar modo ripetizioni, false partenze, esitazioni, autocorrezioni, parole pronunciate male e intercalari) sono facilmente omettibili da parte del rispeaker (35.2% dei casi di omissione non semantica) a vantaggio suo (se è macchinoso e artificiale dettare a una macchina un testo orale che dovrà essere letto, con tanto di punteggiatura, sarebbe ancor più macchinoso e artificiale riprodurre il testo con tutti i tratti dell'oralità producendo così l'evidente paradosso di un testo (tra)scritto[113]), del pubblico di destinazione (che non dovrà decifrare una trascrizione, ma avrà la possibilità di ricevere un testo già formalmente digerito) e del meccanismo di trascrizione da parte del software (che commetterà meno errori nel riconoscere un testo coeso piuttosto che un testo foneticamente e sintatticamente non rispondente alle sue impostazioni); in secondo luogo, poiché il processo traduttivo è di natura orale, ma mira alla produzione di un testo scritto, tutti gli elementi non semanticamente e non grammaticalmente rilevanti che vanno a sconvolgere l'ordine lineare del discorso (fatismi, appellativi, deittici, morfemi grammaticali, ecc.) sono automaticamente omessi (47.8% delle omissioni non semantiche), a vantaggio del carico cognitivo del rispeaker.

Se consideriamo gli esempi che seguono, sarà facile comprendere come l'omissione degli elementi appena citati dia vita a un testo più semplice da produrre per il rispeaker, leggibile da parte del pubblico e comunque completo dal punto di vista informativo:

TP: Thank you, James, for...

TA: Thank you (...) for...

112 Cfr. Cortelazzo 1985.
113 Ibidem.

In questo primo esempio, rispeakerare il TP *verbatim* significherebbe dettare al software 'THANK YOU COMMA JAMES COMMA FOR...'. L'omissione dell'appellativo si traduce in un guadagno sensibile in termini di tempo e sforzo e in una perdita pragmaticamente irrilevante dal punto di vista comunicazionale, visto che le immagini compensano totalmente l'omissione in questione. Consideriamo il seguente esempio:

TP: He said <u>that</u> it is quite clear that...

TA: He said (…) it is quite clear that...

L'omissione di morfemi grammaticali semanticamente inutili e grammaticalmente ridondanti è una strategia frequente (29.2% delle omissioni non semantiche di micro-unità), così come la loro aggiunta (23.3% delle espansioni non semantiche) e la loro ripetizione. Risulta quindi evidente che la natura effimera e discorsivamente molto radicata di questi elementi renda di difficile applicazione la loro sistematica eliminazione. La ragione è strettamente collegata a quanto prima espresso. Parlare e rispeakerare sono due attività orali e transeunti (cfr. Gottlieb 2005: 16) collegate l'ultima alla prima mediante la memoria a breve termine. E visto che il TP non è scolpito in maniera indelebile nella mente del rispeaker, vari fattori (come una bassa qualità di trasmissione del segnale, un alto grado di complessità grammaticale o di densità lessicale, un elevato numero di turni o di tratti dell'oralità, velocità di eloquio diverse, o ancora stress e stanchezza del rispeaker) contribuiscono a rendere il TP un'entità foneticamente instabile nella mente del rispeaker, che avrà delle difficoltà nel rendere fedelmente ogni singolo morfema grammaticale.

L'ultimo caso di omissione non semantica è costituita dall'eliminazione di lessemi, locuzioni o interi sintagmi. La maggior parte degli esempi riscontrati è composto di lessemi contenuti in micro-unità semanticamente non rilevanti e/o pragmaticamente ridondanti, come nel seguente:

TP: we determined <u>as a council</u> three priorities

TA: we determined <u>(...)</u> three priorities

Questo esempio è tratto da una conferenza stampa. L'oratrice è consigliere comunale, è stata presentata come tale e sin dall'inizio del suo discorso ha parlato dell'azione del consiglio comunale in questione alternando i termini 'we', 'the council' e 'we as a council'. Sottolineare per l'ennesima volta che il termine 'we' deve essere inteso come l'insieme dei consiglieri non è un'unità concettuale rilevante dal punto di vista dell'informazione e della coesione e quindi l'omissione della specificazione 'as a council' è stata considerata come semanticamente insignificante. Questa è la stessa ragione per cui il 51.9% delle omissioni lessicali è stato catalogato come omissione non semantica.

Degno di nota è anche il caso delle fusioni. Come nelle fusioni di macro-unità, questa strategia è trasversale alle strategie appena menzionate. Non implicano mai una perdita di informazioni importanti, ma rendono sempre la grammatica meno intricata e di conseguenza i sottotitoli più semplici da leggere e quindi accessibili, come nell'esempio:

TP: there where some locals from Edinburgh who were involved

TA: there where some locals from Edinburgh (…) involved

In questo caso, così come in molti altri casi, la relativa è omessa e l'informazione in essa contenuta è stata attaccata alla principale.

Per quanto riguarda la compressione, anch'essa si scompone in compressione semantica e compressione non semantica. Contrariamente a quanto accade per le omissioni di microstrategie, la compressione semantica è molto utilizzata e rappresenta il 30% delle strategie di riduzione. La compressione semantica si divide ulteriormente in due microstrategie principali: la compressione di singoli lessemi, come la sinonimia orizzontale e verticale (61.5% delle strategie di compressione semantica), e la compressione di intere locuzioni o frasi (38.5% delle strategie di compressione semantica). Quanto alla compressione di lessemi in generale e alla sinonimia in particolare, quest'ultima produce nella maggior parte dei casi delle generalizzazioni della micro-unità di cui fa parte, e conseguentemente della macro- unità concettuale in questione, come nel seguente caso:

TP: in favour of a female pilot...

TA: in favour of a woman...

Qui, il rispeaker sta sottotitolando una notizia e spiega che una donna pilota è stata licenziata dalla compagnia area per cui lavorava perché aveva chiesto di diminuire il proprio monte ore settimanale per avere più tempo per badare la figlia che aveva appena avuto. Nel contesto informativo, la specificazione che la donna fosse un pilota non è così rilevante come il fatto che lavorava per una ben nota compagnia aerea e che da essa è stata licenziata per il motivo appena menzionato. Tuttavia, la resa del termine 'female pilot' con l'iperonimo 'woman' comporta una perdita di informazioni non compensata dal contesto (che restringe comunque il campo agli operatori nel settore dell'aviazione) o da informazioni precedenti o successive. Questo è vero non solo per quei casi di iperonimia in cui ad essere generalizzato è un nome, ma anche per i verbi il cui effetto si estende a tutta la micro-unità concettuale in cui si trova, come nel seguente caso:

TP: so they realise the needs that are actually wanted

TA: so they know what [sic] the athletes.

Qui, il rispeaker ha messo in atto una vera e propria strategia di riformulazione di un testo orale-orale. Questo esempio è tratto da una conferenza stampa. Con uno stile un po' macchinoso, il reporter sta spiegando che la giuria del COI è composta da atleti o ex atleti, quindi sa quali sono le esigenze degli atleti. Il rispeaker capisce il senso dell'unità concettuale e cerca di renderla stilisticamente più accettabile, senza però ottenere il risultato sperato. Intuitivamente, il rispeaker ha detto o 'so they know what the athletes want' o 'so they know what they need'. In entrambi i casi, la comunicazione con il software di riconoscimento del parlato non è passata. Nel primo, il software confonde il termine 'want' per il comando vocale 'stop' o 'point' che portano entrambi all'introduzione del punto. Nel secondo, il software confonde 'they need' con 'the athletes'. Indipendentemente dall'errore di riconoscimento del parlato è possibile notare come il rispeaker abbia sostituito 'realise' con 'know' ottenendo una semplificazione dell'unità concettuale e una diminuzione oltre del numeri di caratteri anche del senso dell'unità stessa, in quanto viene eliminata l'idea del processo cognitivo dietro la conoscenza delle necessità degli atleti. La ragione che sta alla

base di questa strategia non è ben chiara. Si tratta senza dubbio di una strategia, che, nelle intenzioni, sarebbe utile alla leggibilità del TA, in quanto il TP è quantitativamente ridotto con una semplificazione del concetto espresso e senza un'effettiva perdita del senso generale dell'unità concettuale. Eppure, non si contano molti casi simili. Quindi, non si tratta di una strategia comune, ma di un caso isolato dettato più probabilmente da una non perfetta memorizzazione della forma dell'unità concettuale, che non da un'effettiva volontà di ridurre il TP.

Per quanto riguarda i casi di iponimia, essi sono quanto mai rari e la ragione appare alquanto semplice: generalmente, se il giornalista ha dei dettagli, li espone e il rispeaker non ha certamente maggiori dettagli di quanto ne disponga il giornalista. In quei rari casi in cui questo accade, il rispeaker farà comunque fatica a utilizzarli al posto di quanto espresso nel TP col rischio di dover rendere il concetto di più difficile accesso e di dover modificare anche la sintassi. I pochi casi di iponimia riscontrati sono o facilmente intuibili dal contesto (e quindi il rispeaker anticipa l'informazione fornita dal giornalista) o il termine utilizzato dal rispeaker nella produzione del testo di mezzo è presente nell'unità concettuale immediatamente successiva a quella che deve essere resa dal rispeaker (quest'ultimo accetta quindi il suggerimento in fase di resa dell'idea precedente e lo elimina dalla micro-unità in cui è presente o elimina l'intera micro-unità in cui è presente), come nel seguente esempio:

TP: the <u>case</u> will be heard at the employment <u>appeal</u> tribunal

TA: the <u>appeal</u> will be heard at the employment (…) tribunal

Apparentemente, il rispeaker sostituisce il termine 'case' con il termine 'appeal' ottenendo così un'espansione quantitativa della micro-unità in cui esso è contenuto. *De facto*, la macro-unità in cui questa strategia è stata attuata è composta da meno caratteri rispetto alla medesima nel TP. Il rispeaker raccoglie quindi il suggerimento offerto dal rema e sostituisce 'case' con appeal. Una seconda possibilità, molto più probabile, è che il rispeaker sente 'appeal' nel momento in cui sta iniziando a dettare la macro-unità in questione e la usa inconsciamente per iniziare a rendere la macro-unità memorizzata. In seguito, si rende conto che una ripetizione della parola 'appeal' sarebbe poco probabile e

che una sua eliminazione non comporterebbe alcun effetto negativo nella coesione e nella coerenza del TA e quindi omette 'appeal'.

Quanto alla sinonimia orizzontale, la sua incidenza sul totale delle strategie di compressione semantica è decisamente basa. Una ragione può essere ricercata nella relativa inutilità di una tale operazione: ridurre di qualche carattere un termine senza ottenere un reale guadagno in termini di leggibilità risulterebbe semplicemente in uno sforzo cognitivo da parte del rispeaker non richiesto. I casi riscontrati comportano una vera e propria riduzione quantitativa del TP ('local people' tradotto con 'locals', 'attempts' con 'try', ecc.).

Un ulteriore interessante caso di compressione semantica è il seguente:

TP: And in a separate instant, at least four people have been killed

TA: And in a separate attack, several people have been killed

In questo caso, il rispeaker non ha riformulato il TP per ridurlo (la riduzione è limitata a un carattere), ma perché ha preferito essere più chiaro del giornalista. Il risultato pone però una questione di etica professionale: è corretto dire più di quanto dica il giornalista? Quest'ultimo si assume la responsabilità di quel che dice in un contesto di fiducia con i telespettatori e con le persone che contribuiscono a vario titolo alla realizzazione del telegiornale, mentre il rispeaker è solo una delle anonime pedine che contribuiscono alla realizzazione di sottotitoli che sono offerti come servizio dall'emittente. E se la politica di quest'ultima fosse il racconto neutro delle notizie provenienti da ogni angolo del mondo, la sostituzione di 'instant' con 'attack' non comporterebbe forse un'ingiustificata presa di posizione da parte del responsabile della produzione dei sottotitoli? 'Attack' rimanda chiaramente a una deliberata azione violenta di sconvolgimento dello *status quo* per motivi di varia natura (religiosa, politica, etnica, ecc.), mentre 'instant' ha solo una valenza temporale.

Un'ultima strategia di compressione semantica, la compressione frastica, si compone essenzialmente si riformulazioni, sintesi, lessicalizzazioni e dislocazioni. Consideriamo gli esempi che seguono:

TP: his support has been absolutely critical

TA: he's given huge support

E ancora:

TP: that's why London has gone creating an athletes commission asking the athletes: what do you want out of a London bid?

TA: that's why the bid (…) has asked athletes what they want out of a London bid

Queste due strategie rispettivamente di riformulazione e di sintesi del TP rappresentano una percentuale decisamente bassa del totale delle strategie di riduzione (0.9%), ma dimostrano come una sistematica riformulazione o sintesi del TP sia, laddove necessario, non soltanto utile alla leggibilità del TA, ma anche possibile e soprattutto senza perdite importanti nel trasferimento linguistico (nel primo caso, è stato il *focus*, mentre nel secondo è stata persa la micro-unità riguardante la commissione, che comunque svolge un ruolo secondario nella macro-unità in cui è inserita.

Per quanto riguarda la compressione non semantica, anch'essa può essere scomposta in compressione di singoli lessemi e compressione di locuzioni e frasi, nonostante la definizione di compressione di frasi sembri, nei termini, una contraddizione. Quanta alla compressione non semantica di singoli lessemi, la sinonimia è la strategia più utilizzata. L'iperonimia conta per 15.4% del totale delle strategie di compressione non semantica. Il ricorso a una tale strategie sembra essere dettato da un ragionamento apparente da parte del rispeaker come nel caso seguente:

TP: if they concentrate on the smoking and the alcohol-related cancers

TA: if they concentrate on the smoking and drinking cancer

Come succede anche in altri casi, una parola composta ('alcohol-related') è stata qui sostituita da una sinonimo ('drinking') morfologicamente più simile al primo elemento della lista ('smoking'), oltre che più vantaggioso rispetto al relativo termine nel TP.

Come riscontrato nelle compressioni semantiche, l'iponimia è molto meno frequente (1.3% del totale delle strategie di compressione non semantica) e la ragione sembra essere anche qui la medesima. Inoltre, comprimere senza ottenere un risultato apprezzabile in termini semantici, sembra essere veramente inutile. Ecco quindi, che solo in quei casi in cui un iponimo comporta un evidente guadagno in termini di caratteri risparmiati, il rispeaker ricorre a tale strategie, come si evince dall'esempio che segue:

TP: He has not been the asset that some would have liked the French bid to be

TA: It has not been the asset that some would have liked him to be

Questo esempio è tratto da un *reportage* in diretta durante il quale il giornalista spiega come il presidente francese Jacques Chirac si sia comportato in una maniera politicamente scorretta durante l'incontro con il COI. Mentre 'the French bid' deresponsabilizza in maniera evidente il comportamento del presidente attribuendo la delusione delle aspettative di qualcuno all'intera squadra francese, il rispeaker punta il dito direttamente sul presidente francese ottenendo così una maggiore comprensibilità del TA, uno snellimento della sua struttura grammaticale e una riduzione quantitativa dei caratteri utilizzati. Ancora una volta, però, questa strategia pone il problema deontologico del rispeaker che, come l'interprete di simultanea, dovrebbe non interpretare il TP al posto del proprio pubblico, ma lasciare al pubblico la libertà di decifrare che cosa si nasconde dietro le parole, soprattutto in un contesto intralinguistico.

L'ultima strategia che compone il gruppo delle strategie di compressione lessicale non semantica è la sinonimia orizzontale, con un'incidenza del 7.4% sul totale delle strategie non semantiche di riduzione. Essa comprende operazioni come la sinonimia lessicale ('going on' per 'happening', 'today' per 'in the day', 'locals' per 'local people', 'some' per 'a series of', ecc.), sinonimia morfosintattica ('he's going to be meeting' tradotto con 'he'll meet', 'over the course of' con 'on', 'Buckingham' con 'Buck', 'United Nations' con 'UN', ecc.) e l'anafora ('her' al posto di 'Homolka', 'they' al posto di 'London team', ecc.).

Quanto alla compressione non semantica di locuzioni e frasi, ritroviamo ancora una volta la sinonimia orizzontale e l'anafora oltre alle altre strategie menzionate nel caso della

compressione semantica, come la riformulazione, la sintesi, la dislocazione e la lessicalizzazione. Come nel caso precedente, queste strategie sono poco utilizzate dai rispeaker della BBC rispetto alle altre strategie di riduzione non semantica (1.5%). Tuttavia, qualora la loro attuazione risultasse in un guadagno evidente di caratteri, queste strategie sono applicate con risultati più che apprezzabili in termini di spazio, tempo di lettura e coesione, come nei casi qui di seguito riportati:

TP: David Lomas is reported to be in a stable condition in hospital

TA: David Lomas is said to be stable in hospital

TP: This is a 45 minute presentation which is going to be crucial, isn't it? A 45 minute presentation and…

TA: This is a 45 minute presentation which is going to be crucial, isn't it? Yes, and…

TP: We're going to hear from Steve Redgrave, we're going to hear from Matthew…

TA: We're going to hear from Steve Redgrave and Matthew…

TP: Our policing plan for policing this event has been…

TA: Our response has been…

Questi esempi mostrano come una riformulazione o sostituzione sistematica di tutti quei passaggi semanticamente e non semanticamente ridondanti risulti essere non solo possibile, ma anche molto apprezzabile in termini editoriali.

Errori

Come è stato già detto, sia gli errori umani, sia quelli dovuti a un malfunzionamento del *software* (in totale costituiscono il 3.4% di tutte le alterazioni riscontrate) non sono stati considerati come una strategia di espansione o di riduzione in quanto non sono delle operazioni volontarie. Tuttavia, il risultato di questi errori è ben visibile dal pubblico da casa e talvolta possono anche minare la comprensione del TP. Come regola generale, i fattori che influenzano il grado di gravità di questo o quell'errore sono sia obiettivi e misurabili, sia soggettivi e assolutamente impossibili da misurare. Nel primo gruppo di fattori ritroviamo il grado di differenza rispetto al termine desiderato ('the athletes' per 'they need' è obiettivamente di più difficile interpretazione rispetto a 'two much' per 'too much'), il grado di 'visibilità' dell'errore (è più facile individuare e provare a interpretare un errore evidente come 'I'm not going to dis parish anybody' traducente 'I'm not going to disparage anybody', piuttosto che un errore plausibile sia nella forma, sia nel contenuto come 'China has fallen in love with the Netherlands' traducente 'China has fallen in love with the Net') e il grado di disambiguità offerto dal contesto semiotico e semantico (è più automatico inferire il reale testo del primo esempio qui di seguito riportato piuttosto che quello dell'esempio successivo, in cui il contesto non riesce a spiegare la natura del sostegno del pubblico britannico:

> TP: So you've got cancer of the lung, the pharynx, the larynx, cancer of the mouth and the lip.
>
> TA: So you've got cancer of the lung, far inches, the larynx, cancer of the mouth and the lip.
>
> TP: There isn't unanimous support for this bid from the British public
>
> TA: There isn't nam support (…) from the British public

Quanto alle varianti soggettive, possiamo elencare l'attenzione del telespettatore, la sua familiarità con l'argomento trattato e la sua prontezza di riflessi nell'individuare e nel correggere gli errori. Proprio per il peso che questi tre fattori hanno nel grado di gravità

dell'errore, gli esempi che saranno qui di seguito elencati non saranno valutati in base al loro maggiore o minore grado di gravità, ma in base alla loro natura specifica.

Prima di iniziare l'analisi degli errori riscontrati nel corpus, è necessario fare un distinguo essenziale tra errori imputabili all'oratore (alta velocità di eloquio, rapide e numerose prese di parola, termini tecnici inaspettati, pronuncia idiolettale, ecc.), al canale (bassa qualità della trasmissione del segnale in ingresso, rumore ambientale, ecc.) e al rispeaker (stress, stanchezza, uso scorretto degli strumenti a sua disposizione, ecc.). Siccome risulta impossibile individuare eventuali errori imputabili al canale, il numero di categorie in cui collocare gli errori sono stati limitati a due: errori del software ed errori del rispeaker[114]. Un'ultima nota da premettere è la natura dell'errore (sia esso imputabile al software o al rispeaker), che, nel rispeakeraggio, non risulta mai in una non-parola (contrariamente a quanto accade con la stenotipia), ma in un morfema o in un gruppo di morfemi che possono essere più o meno o per niente plausibili.

Consideriamo ora l'esempio seguente:

TP: To add to that, the politics, the rumours, the conversations in corners and…

TA: To add to that, the politics, the Romans, the conversations in corners and…

In questo esempio, il rispeaker non ha probabilmente pronunciato male il testo di mezzo. Semplicemente, il software, che notoriamente non effettua un'analisi semantica a livello frastico, ha riconosciuto l'input orale sulla base dei tre fonemi che maggiormente contraddistinguono la pronuncia della parola 'rumours', vale a dire /r/, /m/ e /s/. In base a un calcolo probabilistico, il software, indeciso tra 'Romans', 'rumours' e altre quattro plausibili soluzioni, ha selezionato quello con la più alta percentuale di ricorrenza nel vocabolario di base. Il risultato è un evidente errore, visto che 'the Romans' è chiaramente fuori contesto. Lo spettatore individua quindi facilmente l'errore, ma ha l'arduo compito di scoprire il termine nascosto dietro l'errore. La soluzione giungerà alla mente del telespettatore

114 Per fugare dubbi circa la scientificità di questa categorizzazione sul nascere, è forse necessario premettere che l'attribuzione alla categoria 'errori del rispeaker' è stata effettuata grazie all'esperienza professionale acquisita nel corso degli anni di ricerca in materia riportati in questo lavoro. La natura di questa categorizzazione è quindi di tipo probabilistico e intuitivo, ma con un buon grado di approssimazione, vista la ricorrenza di molti meccanismi che portano all'errore.

immediatamente, grazie al campo semantico introdotto da 'the conversations in corners', o dopo alcuni istanti, quando il giornalista si riferirà nuovamente ai 'rumours':

TP: Matthew referred to rumours and counter rumours

TA: Matthew referred to the Moors and counter rumours

Ancora una volta, il software non ha riconosciuto correttamente il testo di mezzo, confondendolo ancora una volta con una popolazione storica ('the Moors'), ma il resto del TA (così come il TP) permette una disambiguazione abbastanza immediata: da 'counter rumours' sarà facile risalire a 'rumours' come primo elemento della lista.

Un caso interessante di parole non riconosciute dal software per una lacuna nella programmazione riguarda i monosillabi, che sono spesso confusi con altri monosillabi. In questo caso, gli elementi fonetici che confondono il software impedendogli un corretto riconoscimento del termine desiderato sono sia il suono consonantico in posizione di testa, sia quello vocalico in posizione sia di testa, sia di coda. Nel pronunciare 'oh', ad esempio, il rispeaker è inconsciamente molta più attenzione al fonema vocalico che non alla coda della sillaba. Nel corpus di riferimento, il risultato del processo di riconoscimento di questo termine è spesso 'off'. La stessa analisi vale per 'to' e 'true', 'bid' e 'bed', 'you're' e 'you', ecc.

Un'estremizzazione di questo tipo di errori è il caso degli omonimi ('to', 'two' e 'too'; 'there', 'they're' e 'their'; 'fast and' e 'fasten'; ecc.), che ricorrono costantemente nei casi di errori del software riscontrati.

Un ultimo tipo di errori del software è caratterizzato dalla confusione che fa il software tra una parola dettata e comando vocale e viceversa, come nei due casi qui di seguito riportati:

TP: on my left, David Hemery,

TA: on my left, David Hemmery come up

TP: The point I've been trying to make

TA: The. We're trying to make

Per quanto riguarda gli errori umani, la trattazione necessita maggiore attenzione in quanto le possibili sviste e dimenticanze sono numerose. Il primo caso di errore umano è attribuibile a un calo dell'attenzione del rispeaker:

TP: Everybody is saying it's too close to call and nobody is prepared to predict the result

TA: Everybody is saying it's too close to call and anybody is prepared to predict the result

Questo errore non può essere attribuito al software perché 'nobody' e 'anybody' sono troppo distanti foneticamente da poter essere scambiati l'uno per l'altro, a meno che il rispeaker non abbia prodotto una pausa piena prima di pronunciare 'nobody'. Quanto a plausibili spiegazioni di quanto è accaduto in cabina, il rispeaker ha voluto cambiare la frase da affermativa a negativa, dimenticando però di modificare anche il verbo. Un'altra possibile spiegazione sta nell'eventualità che il rispeaker abbia effettivamente modificato il verbo senza però che il software abbia riconosciuto correttamente 'isn't'. Un'ulteriore ma psico-linguisticamente poco plausibile spiegazione di quanto avvenuto nel processo di produzione del testo di mezzo è la seguente: il rispeaker non comprende bene il TP e lo ripete trascurando la correttezza grammaticale del TA.

Un altro esempio di errore umano ricade nella categoria riguardante la memoria a breve termine. Nell'esempio che segue è semplice rendersi conto che il rispeaker abbia completato l'unità concettuale in questione senza peraltro ricordarsi il dato numerico:

TP: You can see more police here in any five minutes than I've seen in the past ten years

TA: You can see more police here in any five minutes than I've seen in the past five years

Si tratta di fenomeno che si verifica con una certa frequenza anche in interpretazione simultanea, in quanto il dato non è mai contenutistico, ma sempre aleatorio e quindi non può che agganciarsi alla sola memoria ecoica. Sfortunatamente, in questo contesto, la memoria ha ecoica ha funzionato troppo bene. Visto che 'five' era stato già pronunciato dall'intervistato e ripetuto dal rispeaker stesso, esso ha influenzato la memoria ecoica del rispeaker, che non ha potuto fare altro che usarlo nella produzione del TA in luogo del già dimenticato 'ten'.

Un ultimo esempio di errore umano riguarda gli errori di ortografia. In precedenza è stato affermato che il software di riconoscimento del parlato non produce mai una non-parola. Può però produrre un termine in luogo di un altro. In molti casi di omofonia tra nomi comuni, al rispeaker non resta che sperare che il software riesca a 'inferire' dal contesto. In altri casi, però, il rispeaker dispone di diversi espedienti per evitare questo tipo di errori. Il più semplice di questi strumenti è l'uso dei comandi vocali. Se un termine esige una formattazione diversa da quella base, probabilisticamente più accreditata, come nei casi riscontrati di 'standstill' (e non come è stato scritto 'stand still'), 'set-up' (e non 'set up') e 'State' (e non 'state'), allora il rispeaker non potrà dettare semplicemente la forma base, ma dovrà dettare al software la formattazione desiderata (rispettivamente 'ONE WORD STANDSTILL', 'SET HYPHEN UP' e 'CAPITAL STATE').

Qualora l'emittente per cui i sottotitoli debbono essere prodotti avesse una politica editoriale diversa da quella che è stata adottata nella produzione del vocabolario di base e quindi avesse bisogno che alcune parole più comuni siano scritte in un modo piuttosto che in un altro (nel caso di BBC Parliament, 'President' e non 'president', nel caso si BBC News 'UN' e non 'U. N.' e nel caso di BBS Sports 'offside' e non 'off-side'), allora il rispeaker può preventivamente fare ricorso agli *house-style*. In particolare, dovrà fare, per ogni programma, una lista delle parole che vorrebbe essere scritte in modo differente da come il software solitamente le scrive; selezionare, e quindi attivare, la lista di *house-style* giusta prima del sottotitolazione di un dato programma; e assicurarsi che, in diretta, il software trascriva la parola così come compare nella lista di *house-style* selezionata. Nel corpus analizzato, non sono stati riscontrati errori attribuibili a u mancato uso di un *house-style*.

Gli *house-style* possono essere utilizzati solo per quelle parole che non hanno omofoni, altrimenti l'unica versione che uscirebbe dal processo di elaborazione del testo di

mezzo da parte del software sarebbe quella segnalata nella lista di *house-style* selezionata. Per quelle parole, soprattutto i nomi propri, che hanno un omofono abbastanza ricorrente (come 'Butt' e 'but' e 'Cruz' e 'cruise'), una soluzione possibile potrebbe essere il ricorso alle macro di dettatura. In particolare, il rispeaker deve fare una lista delle parole per cui vorrebbe creare una macro e attribuire, a ognuna di essa, una 'word-macro', cioè un'etichetta che, se pronunciata come una parola unica, viene riconosciuta dal software come una macro e resa nella forma desiderata. Per convenzione, questa etichetta si compone della parola in questione e del suffisso 'macro' ('Buttmacro', 'Cruzmacro', ecc.). Nel corpus analizzato, l'unico caso di uso scorretto di macro di dettatura è stata la resa di 'Beckham' con 'back ham' durante una conferenza stampa in cui il celebre calciatore era notoriamente uno degli oratori.

4.6.3 Analisi strategica delle fasi e sotto-fasi

Nel paragrafo precedente è stato possibile osservare l'occorrenza media di ogni singola strategia nell'intero corpus. Per cercare di comprendere meglio le ragioni dell'applicazione di una strategia in un determinato passaggio e la professionalità dei rispeaker della BBC a seconda delle fasi, o *move*, sarà ora approntata un'analisi strategica per ognuna di esse. Nelle figure 14 e 15, sono riportate in maniera schematica rispettivamente le macro-strategie per ogni fase e le strategie di alterazione per ogni fase.

	Non-rese	Rese	
FASI	**Omissioni**	**Ripetizioni**	**Alterazioni**
Titoli	18.2	48.2	33.6
Servizi	16.3	50.4	33.3
Reportage	21	45.1	33.9
Meteo	25.3	32.6	42.1
Sommario	14.5	57.2	28.3
Media	18.9	47.4	33.7

Figura 14: Incidenza delle strategie per ogni fase di *BBC News*.

FASI	**Errori**			**Espansioni**			**Riduzioni**						
	umani	**software**	**Tot**	**Sem**	**NS**	**Tot**	**Omissioni**			**Compressioni**			**Tot**
							Sem	**NS**	**Tot**	**Sem**	**NS**	**Tot**	
Titoli	40	60	3	75	25	8.7	13.7	86.3	44.4	65	35	55.6	88.3
Servizi	43.2	56.8	1.9	31.6	68.4	10.4	18.8	81.2	37.1	58.6	41.4	62.9	87.7
Reportage	68.3	31.7	7.5	30.1	69.9	1.1	4.6	95.4	37.6	43.9	56.1	62.4	91.4

Meteo	**71.9**	**28.1**	**1.3**	**26.5**	**73.5**	**4.3**	**22.1**	**77.9**	**53.2**	**59.9**	**40.1**	**46.8**	**94.4**
Sommario	**41.1**	**58.9**	**1.2**	**68.7**	**31.3**	**8.8**	**44.6**	**55.4**	**30.3**	**46.7**	**53.3**	**69.7**	**90**
Media	**62.2**	**37.8**	**3.4**	**35**	**65**	**7.1**	**13.8**	**86.2**	**37.8**	**50.5**	**49.5**	**62.2**	**89.5**

Figura 15: Distribuzione delle strategie di alterazione per ogni fase di *BBC News*.

Titoli

I titoli sono caratterizzati da un tasso di complessità grammaticale inferiore (quasi tutte le proposizioni seguono una struttura sintattica lineare) e da una densità lessicale e da una velocità di eloquio superiori rispetto alla media. Da un'attenta analisi strategica delle strategie utilizzate nella sottotitolazione in diretta dei titoli è possibile osservare che l'omissione di macro-unità è leggermente inferiore rispetto alla media (18.2% rispetto a una media di 18.9%). Una possibile ragione sta nell'assenza, in questo passaggio, di espressioni formulaiche e di unità ridondanti. Omettere un'unità concettuale in questo *move* significa eliminare un'intera informazione o una parte importante dell'informazione veicolata da ogni titolo. Eppure, il tasso di incidenza delle omissioni non è di molto inferiore alla media. Risulta abbastanza ovvio pensare che la velocità di eloquio, la 'novità' delle informazioni e la densità concettuale che caratterizzano questo *move* controbilancino l'assenza di unità omettibili. Quanto alla ripetizione, essa è leggermente superiore alla media (48.2% contro una media del 47.4%) compensando così i dati riguardanti l'omissione di macro-unità concettuali. Una delle prime cause di questo atteggiamento nei confronti del TP deriva indubbiamente dalla 'novità' delle informazioni in esso contenute, che impediscono al rispeaker di anticipare i concetti o riformularli in un tempo talmente breve da richiedere immediata reazione, in ragione della densità lessicale e della varietà degli argomenti, che rischiano di sovraccaricare la propria memoria a breve termine. Infine, proprio in virtù dell'effetto compensatore delle strategie di ripetizione nei confronti delle strategie di omissione delle macro-unità concettuali, l'incidenza delle strategie di alterazione risulta essere molto vicina alla media (33.6% contro 33.7%).

A guardare bene i dati circa la distribuzione interna delle singole sotto-strategie, però, è possibile notare come l'espansione sia sensibilmente più presente (8.7% contro una media di 7.1%), la riduzione meno presente (88.3% contro una media di 89.5%) e gli errori siano invece leggermente inferiori rispetto alla media (3% contro 3.4%). Mentre il dato

riguardante gli errori non sorprende[115], è interessante notare come l'espansione sia più presente della riduzione rispetto alla media in un passaggio che ha tutte le caratteristiche per non lasciare spazio a un tale fenomeno. Una possibile spiegazione di questo sta nel peso di ogni singola strategia sul totale del *move* in questione. In particolare, una strategia di espansione sul totale delle strategie attuate sull'intero corpus sposta la media di meno di un millesimo, mentre una strategia di espansione sul totale delle strategie attuate sul solo *move* dei titoli, che conta il minor numero di unità concettuali rispetto al totale delle unità concettuali, ha un impatto di molto maggiore. Lo stesso vale per spiegare la distribuzione interna delle micro-strategie di espansione, che vede il 75% delle espansioni di natura semantica e il 25% di natura non semantica contro, rispettivamente, il 35% e il 65% della media. In questo caso, un altro fattore di indubbia influenza sul lavoro del rispeaker è la mancanza di tempo (conseguente a un'alta velocità di eloquio e un'alta densità lessicale) nel trasformare i verbi dalla loro forma contratta alla forma standard. Come nel caso precedente e come già affermato, in mancanza di un alto tasso di espansioni non semantiche, le espansioni semantiche hanno un'incidenza maggiore sul totale delle strategie di espansione dei titoli. Degno di nota è il posizionamento della maggior parte delle espansioni, vale a dire la fine delle proposizioni. Questo potrebbe significare che il rispeaker approfitta delle brevi pause tra un titolo e l'altro per migliorare sia la forma sia il contenuto delle unità concettuali nel TA

Per quanto riguarda le strategie di riduzione, una distinzione importante va fatta tra le omissioni di micro-unità e le compressioni. Infatti, la media delle riduzioni non differisce di molto dalla media del corpus (88.3% contro 89.5%). Tuttavia, mentre la distribuzione interna delle omissioni riflette pressoché interamente la media (le omissioni semantiche rappresentano il 13.7% del totale contro il 13.8% della media e le omissioni non semantiche rappresentano l'86.3% contro l'86.2% della media), le compressioni semantiche e non semantiche hanno una distribuzione interna simile a quella delle espansioni, cioè a dire praticamente opposta alla media del corpus (rispettivamente 65% e 35% contro 50.5% e 49.5 %). Le ragioni di questo fenomeno sono plausibilmente le stesse delle espansioni: visto

115 Maggiore è la densità lessicale e minore è la complessità grammaticale, meno monosillabi saranno presenti nelle macro-unità che compongono il move in questione. Essendo i monosillabi una delle maggiori fonti di errore di riconoscimento per colpa sia del rispeaker (che potrebbe pronunciarli attaccati alle parole limitrofe), sia, soprattutto, del software (che notoriamente non riesce a fare molta distinzione tra i monosillabi), una logica conseguenza della loro minore incidenza è la minore presenza di errori.

che non c'è tempo per ridurre i morfemi grammaticali, l'incidenza delle riduzioni non semantiche è, rispetto alla media, sensibilmente inferiore.

Concludendo con gli errori, anch'essi sono in linea con le altre due modalità di alterazione delle macro-unità concettuali dei titoli, in quanto l'incidenza sul totale delle strategie di alterazione è simile alla media, ma la distribuzione interna si discosta di molto dalla media: il 40% degli errori è imputabile al rispeaker contro il 62.2% della media, mentre il restante 60% è imputabile al software contro una media del 37.8%. Si tratta di un dato strano e parzialmente contrario alle aspettative. Le caratteristiche testuali dei titoli sono infatti dei fattori che generalmente aumentano lo stress del rispeaker (alta velocità di eloquio, alta densità lessicale, novità delle informazioni, ecc.). Tuttavia, il fatto che il rispeaker sia all'inizio del suo turno potrebbe giocare a suo favore, in quanto il rispeaker sarà meno stanco e quindi meno sensibile allo stress. Un altro motivo che potrebbe giustificare questo ribaltamento dei dati, strettamente collegato a quello appena menzionato, è l'alto tasso di errori imputabili al software. Essendo all'inizio del suo turno, il software lavora 'a freddo'[116], impiega più tempo nell'elaborazione dei dati in ingresso e produce quindi un numero di errori comprensibilmente superiore rispetto alla media.

Servizi pre-registrati

In questa analisi, in virtù delle loro differenze di fondo, servizi pre-registrati e *reportage* dal vivo sono considerati separatamente anche se cronologicamente non possono esserlo, in quanto, normalmente, sono sottofasi dello stesso *move* oppure incassati l'uno nell'altro. Come i titoli, i servizi pre-registrati sono testi pre-strutturati scritti per essere letti, quindi caratterizzati da una bassa complessità grammaticale e da un'alta densità lessicale. Contrariamente ai titoli, però, la velocità di eloquio è percettibilmente inferiore a quella dei titoli a causa del ruolo importante che svolgono le immagini, assenti nei titoli. In ragione dell'impatto massiccio che hanno le strategie utilizzate per sottotitolare i servizi pre-registrati sulla media delle strategie (circa la metà delle unità concettuali che compongono il corpus analizzato è composto da servizi pre-registrati), la loro incidenza si accosta maggiormente alla media di quanto non accada nel caso di *move* quantitativamente inferiori.

116 La calibrazione del microfono, che solitamente va effettuata prima del proprio turno di lavoro, non risolve totalmente i problemi di adattamento del software alla voce del rispeaker. Inoltre, proprio perché parte della memoria è impegnata nell'adattamento alle caratteristiche della voce del rispeaker, il software è più sensibile a eventuali sovraccarichi di input.

Gli scostamenti dalla media sono quindi da considerare diversamente rispetto a quanto è stato appena fatto con le strategie riguardante i titoli.

Tuttavia, alcune differenze possono essere identificate. L'omissione di macro-unità è del 2.6% inferiore rispetto alla media. Il motivo è probabilmente una velocità di eloquio inferiore alla media. Questo è confermato dai dati sulla ripetizione di macro-unità, che controbilanciano il dato delle omissioni con un'incidenza superiore alla media del 3%. La ragione che sta alla base di una tale strategia è dettata dalla possibilità di ripetere il TP parola per parola. Quanto alle strategie di alterazione, esse corrispondono più o meno alla media: 33.3%, contro il 33.7% del totale delle strategie di alterazione sul corpus di riferimento. Anche la distribuzione interna di questa macrostrategia conferma la relativa semplicità di rispeakerare questo *move* rispetto alla difficoltà media. Una prima prova di quanto affermato sta nelle espansioni, che costituiscono il 10.4% delle strategie di alterazione, cioè a dire il 3.3% in più della media. Questo dimostra che con maggiore tempo a disposizione (dettato principalmente dal numero maggiore di pause tra un oratore di una parte di un servizio e l'altra e dalla maggiore interazione tra il testo e le immagini), il rispeaker si può permettere di disambiguare parti del discorso lessicalmente o sintatticamente poco chiare. Tuttavia, la distribuzione interna di questa strategia sembra affermare il contrario: l'espansione non-semantica è superiore alla media (68.4% contro 65%). Una prima spiegazione di questo fenomeno sta nella natura linguistica del *move* in questione. Essendo un testo preparato in anticipo, l'autore ha avuto modo di pensare sia alla forma sia al contenuto del TP. Di conseguenza, il margine di manovra che avrà il rispeaker sarà sul piano morfo-sintattico, non tanto su quello lessico-semantico. Un altro motivo della bassa incidenza di espansioni semantica è da ricercarsi nell'alta densità lessicale delle macro-unità concettuali contenute in questo *move*, superiore alla media. Questo rende impossibile un'ulteriore espansione semantica del testo. Questo non è in contrasto con quanto appena affermato circa la maggiore disponibilità di testo. Se è vero infatti che la velocità di eloquio è più bassa rispetto alla media e che ci sono più pause nel testo è anche vero che tra le pause ci sono più informazioni da elaborare.

Per quanto riguarda le riduzioni, vale il medesimo criterio alla base delle espansioni. Da una parte, esse hanno un'incidenza inferiore alla media (87.7% contro 89.5%), ma, dall'altra, la compressione ha un'incidenza di molto superiore alle altre due strategie di riduzione: omissione di micro-unità ed errori. La causa non è tanto un'effettiva

mancanza di strategie di omissione di micro-unità concettuali o l'eventuale assenza di errori, quanto la combinazione dei due fattori summenzionati: sufficiente tempo per elaborare il TP e un'alta densità lessicale del TP. In queste condizioni, il rispeaker può ridurre le unità concettuali in caso di sovraccarico della propria memoria a breve termine riducendo quantitativamente il testo, tramite sintesi e riformulazioni. Anche con il tempo necessario a una buona compressione puramente quantitativa, risulta quindi molto probabile anche l'accidentale eliminazione di lessemi semanticamente rilevanti. E infatti la compressione semantica ha un'incidenza superiore alla media, 58.6% contro 50.5%, determinando così una differenza dell'8.1% tra l'incidenza delle compressioni non semantiche in questo *move* e la media delle compressioni non semantiche. La ragione sta nell'altro elemento che distingue un testo preparato in anticipo da un testo improvvisato oralmente: la complessità grammaticale. Essendo il *move* in questione meno intricato dal punto di vista grammaticale rispetto alla media, l'intervento del rispeaker si concentrerà sulla resa alternativa di alcune micro-unità o sulla loro eliminazione. Tuttavia, i dati riguardanti questo ultimo aspetto dicono chiaramente che l'omissione di micro-unità ha un'incidenza dello 0.7% inferiore rispetto alla media sul totale delle strategie. Come nel caso delle omissioni di macro-unità concettuali, il tempo svolge qui un ruolo favorevole a strategie di minor impatto semantico sul TA. Se è possibile una riformulazione, perché omettere? Laddove una omissione è necessaria, visto anche il ruolo dell'alta densità lessicale, essa avrà una natura semantica (superiore del 5% rispetto alla media). Come già detto, una elevata densità lessicale causa un sovraccarico di memoria al rispeaker, anche in condizioni temporali favorevoli. Una delle strategie più semplici risulta quindi essere l'omissione di un aggettivo o di un sostantivo da una lista o un inciso piuttosto che un morfema grammaticale. Quest'ultima operazione è infatti cognitivamente più complessa rispetto alla prima, perché mentre l'omissione di un elemento da una lista necessita semplicemente di una lieve sospensione della memoria in attesa dell'elemento successivo o della fine del sintagma di cui è parte, l'omissione grammaticale comporta la sospensione temporanea della memoria e il necessario sforzo di adattamento della grammatica del TA, salvo rari casi (tra i più comuni l'omissione di 'that' in una relativa e quella di un pronome soggetto in una serie di coordinate con lo stesso soggetto).

Per quanto riguarda gli errori, essi hanno un'incidenza dell'1.9%, sensibilmente inferiore rispetto alla media del 3.4%. La ragione sta, ancora una volta, nel relativamente

semplice compito che hanno sia il rispeaker, sia il software nel sottotitolare questi passaggi. Il primo, che in molti casi avrà ricevuto anticipatamente il servizio da sottotitolare, sarà agevolato sia dalla maggiore lentezza del *move*, sia dalla conoscenza pregressa del TP. Il secondo sarà già stato istruito sulle eventuali parole di difficile riconoscimento nella fase preparatoria (tramite macro, *house-style* o aumenti della probabilità statistica di un dato termine) e limiterà la produzione di errori di riconoscimento ai soli mono- e bi-sillabi. La distribuzione interna riguardante gli errori umani (43.2% contro una media del 62.2%) e quelli dovuti a un malfunzionamento del software (56.8% contro una media del 37.8%) è una dimostrazione di quanto appena affermato.

Reportage in diretta

Questo *move* è indubbiamente quello che presenta maggiori sfide al rispeaker per molteplici aspetti. Innanzitutto, esso costituisce circa la metà dell'intera edizione del giornale. Da punto di vista linguistico, esso è caratterizzato da un'alta complessità grammaticale, da una generalmente elevata densità lessicale, con un'alta presenza di nomi propri e termini tecnici o contestuali non sempre noti al rispeaker e al software, da numerosi turni di parola (specialmente nelle interazioni con il giornalista presente in studio), da un'elevata presenza di tratti dell'oralità (false partenze, auto-riformulazioni, ridondanze e pause piene) e da una qualità della trasmissione del suono raramente paragonabile a quella degli altri *move* (soprattutto nei collegamenti da un luogo rumoroso o all'aperto e nei collegamenti telefonici). Quanto alla velocità di eloquio, essa dipenderà da numerosi fattori contestuali (percepita difficoltà dell'argomento, tempo a disposizione per il collegamento, eventuale interazione con lo studio, ecc.) e soggettivi (padronanza dell'argomento trattato, grado di pre-strutturazione del discorso, capacità retorica, velocità media di eloquio, ecc.). I *reportage* dal vivo possono sia essere dei *move* a sé stanti, sia essere contenuti in un altro servizio, sia contenere altri servizi pre-registrati (montati o semplicemente registrati e proiettati in un secondo momento senza interventi editoriali precedenti) o dal vivo (conferenze stampa, interviste a testimoni o a esperti, tele-cronache, ecc.). Date queste condizioni, è chiaro comprendere come il rispeaker sia obbligato a fare ricorso, sempre come *extrema ratio*, all'omissione in maniera sensibilmente superiore rispetto alla media (21% contro 18.9%) e alle ripetizioni fedeli in maniera inferiore rispetto alla media (45.1% contro 47.4%).

Quanto alle alterazioni, la natura del TP sembra non avere alcuna influenza sulla loro incidenza rispetto alla media (33.9% contro 33.7%). Interessante, è da notare la distribuzione interna delle strategie di alterazione rispetto alla media. Come prevedibile, le espansioni hanno un tasso di incidenza addirittura inferiore a quello degli errori (1.1%). Una delle possibili ragioni è la percentuale di informazioni non note rispetto a quelle già note al rispeaker e al software. In queste condizioni, il rispeaker non si impegna in un'aggiunta di informazioni o anche semplicemente di caratteri, perché non sa che cosa succederà immediatamente dopo e deve mantenere al minimo il divario tra il TP e il TA. Tuttavia, in maniera quasi automatica trasforma quelle peculiarità tipiche di un testo orale in inglese scritto standard. Ne è una prova l'incidenza delle strategie di espansione non semantica, superiori del 4.9% rispetto alla media. Si tratta essenzialmente di trasformazione di verbi dalla forma contratta a quella standard (we'll → we will, she's → she is, ecc.) e di aggiunta del pronome soggetto in una serie di coordinate con il medesimo soggetto (Homolka made a deal with the prosecutors, (…) testified against her husband and (…) got off → Homolka made a deal with the prosecutors, she testified against her husband and she got off). In questi casi, il rispeaker non ha difficoltà a completare la forma verbale o ad aggiungere il pronome laddove necessario, perché la forma del TP è quella che corrisponde alla versione standard in inglese scritto. Si tratta quindi di un'operazione quasi automatica, che non comporta ulteriori sovraccarichi cognitivi. Quanto alla espansione semantica, essa rappresenta il 30.1% del totale delle strategie di espansione, contro una media del 35%. Questo dato è una chiara dimostrazione di quanto sopra accennato: in caso di informazioni totalmente sconosciute al rispeaker sia nella forma sia nel contenuto, la sua tendenza a completare il TP con delle informazioni note o con delle rese semanticamente più chiare è limitata allo stretto necessario. Completa il quadro delle motivazioni che giustificano una diminuzione, rispetto alla media, delle strategie di espansione semantica, l'alta incidenza delle strategie non semantiche dovute a un alto tasso di auto-riformulazioni, pause piene, intercalari e ripetizioni ridondanti.

Contrariamente a quanto ci si potrebbe aspettare, le riduzioni non controbilanciano completamente la bassa incidenza delle strategie di espansione sul totale delle alterazioni. Esse infatti rappresentano il 91.4%, cioè l'1.9% in più rispetto alla media. Una possibile ragione è l'alta incidenza degli errori, che sono più del doppio rispetto alla media (7.5% contro 3.4%). Un altro dato di difficile interpretazione è la quasi totale uguaglianza rispetto

alla media delle strategie di compressione e di omissione di micro-unità concettuali: rispettivamente 62.4% e 37.6% contro una media del 62.2% e del 37.8%. Come avviene nel caso dei servizi pre-registrati, il rispeaker tende di più a riformulare e a riassumere piuttosto che a omettere. Questo perché i *reportage* dal vivo sono spesso caratterizzati da una densità lessicale che, pur essendo elevata, è inferiore alla media. Di conseguenza, ogni unità concettuale è espressa con un numero più elevato di parole rispetto a testi scritti per essere letti. Il rispeaker farà quindi relativamente poca fatica a dire il testo in altre parole e non sarà pertano necessario omettere concetti, seppur accessori.

Un dato interessante riguarda il rapporto tra le strategie semantiche e quelle non semantiche. Le strategie di compressione non semantica costituiscono il 56.1% delle strategie di compressione rispetto a una media del 49.5% e quelle di omissione non semantica di micro-unità concettuali il 95.4% rispetto a una media dell'86.2% la ragione principale è la medesima dell'incidenza delle espansioni non semantiche: vista l'alta presenza di lessemi grammaticali, il rispeaker riesce a comprimere od omettere quelle micro-unità più legate all'aspetto grammaticale del testo e pertanto più ridondanti, senza peraltro perdere micro-unità semanticamente rilevanti. Per quanto riguarda le omissioni di micro-unità, esse intervengono sui tratti più squisitamente orali come le false partenze, le auto-riformulazioni e le esitazioni. Essendo superiori rispetto alla media, l'omissione di questi ultimi tratti fa aumentare percettibilmente l'incidenza delle omissioni non semantiche rispetto a quelle che hanno un impatto sostanziale sulla natura del processo traduttivo e quindi sul TA.

Gli errori, come già accennato, sono più del doppio rispetto alla media. Il motivo di questo fenomeno risiede nel tasso di novità delle informazioni, sia dal punto di vista della forma, sia dal punto di vista del contenuto, e nel conseguente stress derivante sia dall'obiettiva difficoltà di ripetere il TP senza poterne prendere sufficientemente le distanze da avere un'idea del contenuto (per non aumentare troppo il *décalage* con il TA), sia dall'obbligo di ridurlo o espanderlo laddove le circostanze lo richiedano e o lo permettano. Questa spiegazione dello stress ha un riscontro immediato nel tasso riguardante gli errori umani (68.3%), superiore del 6.1% rispetto alla media (62.2%), che dipende, a sua volta, da un altro fattore di stress: l'impossibilità da parte del rispeaker di introdurre parole nuove e macro di dettatura nel software. Uno strumento potrebbe essere utilizzato per superare questa *impasse*: l'uso della tastiera per scrivere la parola in questione. Tuttavia, la fonetica

dell'inglese non garantisce il buon esito di una tale operazione, vista la non binomica corrispondenza con l'ortografia. Inoltre, il rispeaker potrebbe non rendersi immediatamente conto della novità di un termine per il software.

Previsioni meteorologiche

Se i *reportage* presentano la difficoltà di essere prevalentemente dei testi pensati e prodotti oralmente, le varie componenti diamesiche che li compongono non sono talmente dipendenti le une dalle altre da non concedere una certa libertà al rispeaker nel sottotitolarli. Nei casi in cui questo si verifica maggiormente (conferenze stampa, interviste e telecronache), la complessità semiotica del TP non è così vincolante da impedire una buona gestione del *décalage* tra il TP e il TA. Le previsioni meteo, invece, presentano in un unico *move* tutte le difficoltà degli altri *move*: il giornalista parla a una velocità di eloquio molto sostenuta, con un elevato tasso di densità lessicale e una bassa complessità grammaticale. Inoltre, la componente video svolge un ruolo talmente tanto importante che il rispeaker, deve non solo mantenere lo stesso passo per non perdere informazioni importanti, ma anche ridurre il *décalage* al minimo onde evitare di proiettare quelle che produce sotto immagini che non hanno niente a che fare con il TP. Di fronte a queste difficoltà, i rispeaker della BBC omettono macro-unità concettuale per una percentuale di molto superiore alla media (25.3% contro il 18.9%). Le unità maggiormente omesse sono contenute negli *step* di transizione che pongono la sintesi delle previsioni che il giornalista ha appena esposto per una data regione o Stato. Non è raro però notare la presenza di altre omissioni di unità concettuali che non offrono altra ragion d'essere che la casualità. Meglio, il rispeaker non riesce a tenere il passo del TP e quindi, per non sovraccaricare troppo la propria memoria a breve termine o per porre rimedio a una memoria a breve termine già sovraccaricata, tralascia qualcosa riprendendo la sua resa dall'ultima unità concettuale memorizzata o inserendo un punto e riprendendo la sua resa da quella ancora a venire. Paradossalmente, però, queste omissioni non possono essere qualificate come semanticamente rilevanti, perché le immagini compensano *in toto* il senso del TP, che più di accompagnare le immagini, le descrive, svolgendo così un ruolo del tutto ancillare alla componente audio. In seguito alla forte incidenza delle omissioni di macro-unità e, come si vedrà in seguito, delle alterazioni, le ripetizioni rappresentano il tasso più basso all'interno dell'intero corpus, il 32.6%, contro una media superiore del 14.8% (47.4%).

Per quanto riguarda le alterazioni, anch'esse presentano il tasso più alto di tutti gli altri *move*, 42.1%, cioè a dire l'8.4% in più rispetto alla media (33.7%). All'interno di questa macro-strategia, le espansioni rappresentano soltanto il 4.3%, un dato che è inferiore del 2.8% rispetto alla media. La ragione di questo dato risulta chiara nella misura in cui il TP presenta molte più difficoltà e molte più incognite rispetto a ogni singolo altro *move*. Per cui non c'è né tempo né spazio per poter aggiungere informazioni al TP o modificarne la grammatica. Inoltre, c'è da considerare l'alta incidenza delle strategie di riduzione (sia compressione, sia omissione di micro-strategie) che fa abbassare la rappresentatività di ogni altro dato. Un dato interessante riguarda il raffronto con i dati riguardanti i *reportage* dal vivo, che, nella forma, sono molto simili alle previsioni meteorologiche. Le espansioni nei *reportage* in diretta costituiscono solo l'1.1% del totale delle strategie di alterazione e quindi il dato riguardante le previsioni meteo si vestono di un nuovo significato, contraddittorio rispetto a quanto è stato appena affermato riguardo a quest'ultimo *move*. In realtà, c'è da considerare un aspetto statistico che potrebbe non essere ininfluente: il campione di riferimento è, nel caso delle previsioni meteo, quantitativamente di molto inferiore rispetto a quello dei *reportage* in diretta. Ogni caso di espansione nelle previsioni meteorologiche ha quindi una maggiore incidenza sul campione di riferimento rispetto a quanto invece accade nei *reportage*, che costituiscono circa la metà del corpus. Inoltre, c'è da ricordare che l'aspetto maggiormente interessante è l'esigua incidenza di questi dati e che, spesso, quelle che risultano essere delle strategie di espansione sono *de facto* il risultato di operazioni quasi automatiche e spesso inconsce da parte del rispeaker. Per concludere, il dato riguardante la distribuzione interna delle strategie di espansione potrebbe completare il quadro delle motivazioni di questo dato. Le espansioni semantiche, infatti, rappresentano il 26.5% mentre quelle non semantiche il 73.5%, rispettivamente il dato più basso e quello più alto nella loro categoria. Questo significa che la maggior parte delle strategie di espansione è composta proprio da quelle operazioni che sono state definite come automatizzate e quasi inconsce, come la trasformazione di forme contratte in forme standard, l'introduzione di congiunzioni tra due coordinate o l'uso di connettori più lunghi rispetto a quelli utilizzato nel TP. Questo riduce l'intervento deliberato del rispeaker a poche espansioni semantiche, che però hanno un'incidenza maggiore rispetto a quelle contenute in *move* quantitativamente più significativi, come i servizi pre-registrati e i reportage i diretta.

Per quanto riguarda le riduzioni e la distribuzione interna a questa categoria di strategie, i dati che sono stati ricavati riflettono la tendenza generale del *move* in questione. Le strategie di riduzione corrispondono infatti al 94.4% del totale delle strategie di alterazione del TP, il dato più elevato nella sua categoria. Questo gioca a favore delle riduzioni semantiche che, come intuibile, hanno un'incidenza superiore alla media. Per quanto riguarda le omissioni di micro-unità concettuali, che rappresentano il 53.2% del totale delle strategie di riduzione (il 15.4% in più rispetto alla media), le omissioni semantiche rappresentano il 22.1% delle omissioni di micro-unità mentre quelle non-semantiche solo il 77.9%, cioè l'8.3% in meno rispetto alla media e ben il 17.5% in meno rispetto a quanto accade nei *reportage* in diretta. Questa apparente contraddizione è presto spiegata dalla natura linguistica del TP, che, contrariamente alla natura orale-orale dei *reportage* dal vivo, sono testi pre-strutturati scritti per essere letti. Di conseguenza, i tratti dell'oralità sono, in questo *move*, di gran lunga inferiori rispetto ai *reportage*. Un ultimo dato che spiega quanto appena affermato proviene dall'alta incidenza delle omissioni semantiche, seconde, per quantità, solo alle strategie nella resa del sommario.

Quanto alle compressioni, esse rappresentano il dato più basso della loro categoria, il 46.8%, cioè a dire il 15.4% in meno rispetto alla media. Pur rimanendo un dato importante, è facile comprendere come le compressioni non svolgano un ruolo preponderante nelle operazioni di riduzione del TP nelle previsioni meteo. Riformulare o riassumere in condizioni semiotiche particolarmente vincolanti, come è il caso del *move* in questione, risulta essere molto complicato, soprattutto perché il linguaggio delle previsioni meteo è specialistico, nel senso che fa uso di un lessico molto preciso e particolarmente ristretto, pressoché insostituibile con sinonimi più brevi, peraltro assenti o poco utilizzati. La distribuzione interna alle compressioni conferma questa intuizione: le compressioni non semantiche costituiscono il 40.1% del totale delle compressioni, ossia il 9.4% in meno rispetto alla media. Nella maggior parte dei casi riscontrati, il rispeaker interviene poco a livello grammaticale, ma si concentra soprattutto sulle unità concettuali (micro o macro che siano) determinando così una bassa incidenza delle compressioni non semantiche. La maggiore preoccupazione dei rispeaker sembra essere infatti una resa con lessemi di origine germanica di lessemi di origine latina, con un evidente guadagno in termini di caratteri e di comprensibilità del TA da parte degli spettatori (questa è forse una scelta delibertata dei rispeaker, in quanto ricorrente).

Quanto agli errori, i dati sono in controtendenza rispetto a quanto ci si possa aspettare: essi costituiscono infatti solo l'1.3% delle strategie di alterazione, meno della metà della media (3.4%) e quasi un sesto dei dati riguardanti i *reportage* in diretta. I fattori determinanti un tale dato sono già stati esplorati: la relativa bassa frequenza di operazioni di riformulazione e di sintesi comporta un minore stress per il rispeaker e, conseguentemente, una bassa incidenza degli errori da esso provocati; inoltre, la terminologia appartiene a un settore abbastanza circoscritto rispetto a quello generale degli altri *move* e questo fa diminuire le probabilità di errore del software. A questo si aggiunge la compensazione delle immagini, che potrebbero addirittura facilitare, soltanto in questi casi, il rispeaker nella resa del senso del TP. Ecco quindi che si spiega anche la distribuzione interna degli errori, determinati maggiormente dal software, 71.9%, ben il 34.1% in più rispetto alla media (che come negli altri *move* fa fatica a riconoscere lessemi grammaticali e monosillabi), che non dal rispeaker (28.1%).

Sommario

Il sommario è un caso di *move* relativamente facile da sottotitolare. I dati relativi sembrano confermare questa affermazione. L'omissione di macro-unità registrano il tasso più basso della loro categoria, 14.5%, le ripetizioni il dato più elevato della categoria, 57.2% e infine le alterazioni il più basso della categoria, 28.3%. Oltre a questi primi dati, altre caratteristiche intrinseche suffragano l'ipotesi della relativa facilità del *move* in questione. Per quanto riguarda la forma, infatti, il sommario assomiglia in parte ai titoli (poche proposizioni per notizia, elevata densità lessicale, bassa complessità grammaticale e assenza di compensazione dalla componente video del testo) e in parte ai servizi pre-registrati (predominanza di informazioni conosciute, assenza di termini e concetti ignoti e una relativamente bassa velocità di eloquio). Quest'ultimo accostamento è confermato dai dati relativi alla distribuzione delle tre macro-categorie considerate. Se quelle del sommario sono le più e le meno elevate della rispettiva categoria, quelle dei servizi preregistrati sono quelle che più si avvicinano a questi dati (le omissioni sono il 16.3% del totale delle strategie usate per sottotitolare i servizi pre-registrati, contro 14.5% del sommario, le ripetizioni il 50.4% contro il 57.2% del sommario e infine le alterazioni il 33.3% contro il 28.3%). Ma è rispetto alla media che i dati riguardanti le strategie utilizzate per sottotitolare il sommario acquisiscono la loro vera identità. Le omissioni di macro-unità, infatti, sono il 4.4% in meno

rispetto alla media, le ripetizioni il 9.8% in più rispetto alla media e le alterazioni il 5.4% in meno.

Quanto ai dati riguardanti la distribuzione interna delle strategie di alterazione, esse sembrano invece confermare il primo raffronto summenzionato, quello con i titoli. Le espansioni infatti hanno più o meno la stessa incidenza (l'8.8% nel caso del sommario e l'8.7% nel caso dei titoli). Lo stesso dicasi per la distribuzione interna delle espansioni, che, in entrambi i casi, si discostano di molto dalla media. L'espansione semantica rappresenta infatti il 68.7% nei sommari e il 75% nei titoli, contro una media del 35%, mentre l'espansione non semantica rappresenta il 31.3% nei sommari e il 25% nei titoli contro una media del 65%. Contrariamente a quanto accade con i titoli, in questo caso, la ragione non sta soltanto in una carenza di tempo da parte del rispeaker che gli impedisce di compiere operazioni semanticamente non rilevanti, ma anche e soprattutto nella natura delle unità concettuali del TP, che sono già note al rispeaker. Questi ha infatti già sottotitolato i titoli e le singole notizie. Ha quindi bene in mente la terminologia utilizzata per rendere i concetti che saranno espressi nel sommario e in molti casi anche una predeterminata struttura sintattica. Nella fase di sottotitolazione sceglierà quindi il termine e la struttura sintattica più appropriati o quelli che gli vengono prima in mente senza dover sovraccaricare troppo la propria memoria a breve termine.

Per quanto riguarda le riduzioni, esse rappresentano il 90% del totale delle strategie di alterazione, un dato che si accosta molto ai dati relativi sia ai titoli (88.3%), sia ai servizi pre-registrati (87.7%). La somiglianza tra le diverse strategie dei diversi *move* in questione si ferma qui. Le omissioni di micro-unità rappresentano infatti il 30.3% del totale delle strategie di riduzione, cioè il 7.5% in meno rispetto alla media, il 6.8% in meno rispetto ai servizi pre-registrati e ben il 14.1% in meno rispetto ai titoli. Questo accade principalmente per il motivo appena menzionato: il rispeaker conosce già il contenuto delle notizie e non ha quindi bisogno di molto sforzo per memorizzare il TP e per produrre il TA. Di conseguenza, non avrà bisogno di omettere seppur minime porzioni di unità concettuali visto che il suo bagaglio di conoscenze gli permette di svolgere il proprio lavoro senza troppo dispendio di energie. Questo aspetto ha però un effetto collaterale: il rispeaker è alla fine del proprio turno di sottotitolazione e quindi la sua attenzione e la sua reazione sono progressivamente diminuite nel corso del rispeakeraggio. Se quindi il testo da sottotitolare è relativamente più semplice rispetto ai *move* precedenti, anch'esse sono inferiori rispetto all'attenzione e alla

reazione a sua disposizione nei primi *move*. Inoltre, il TP può variare rispetto alle conoscenze del rispeaker o, più in generale, presentare delle difficoltà impreviste. Ecco quindi che un eccessivo rilassamento dell'attenzione e della reazione necessarie al rispeakeraggio in generale da parte del rispeaker porta a un peggioramento della resa. Una dimostrazione di questo sta nella natura delle omissioni di micro-unità. Paradossalmente, infatti, il rispeaker che si rilassa eccessivamente rispetto a quanto richiede il rispeakeraggio in quel passaggio e che quindi incappa in ostacoli imprevisti si trova obbligato a omettere interi sintagmi per evitare di rimanere indietro. Solo così si spiegano i dati riguardanti la distribuzione interna delle omissioni di micro-unità: quelle semantiche registrano infatti il dato più elevato della propria categoria (55.4%) e quindi quelle non semantiche il dato più basso (44.6%).

Quanto alle compressioni, esse rappresentano il 69.7% delle strategie di riduzione, il 7.5% in più rispetto alla media, il 6.8% in più rispetto ai servizi pre-registrati e ben il 14.1% in più rispetto ai titoli. Nessuna spiegazione a questo fenomeno sembra essere veramente soddisfacente. Di certo, vale lo stesso motivo adottato per spiegare la bassa incidenza delle omissioni di micro-unità, ossia un più elevato tasso di conoscenza del TP e conseguentemente una più rapida elaborazione del TA, che porta a privilegiare la struttura sintattica e la terminologia già memorizzate piuttosto che quelle eventualmente nuove. Questa motivazione sembra spiegare, in parte anche la distribuzione interna delle strategie di compressione. Quelle semantiche rappresentano il 46.7% e quelle non semantiche il 53.3%. Come le strategie di omissione di micro-unità, le strategie di compressione sono quantitativamente molto distanti da quelle sia dei titoli (rispettivamente il 65% e il 35%), sia dei servizi pre-registrati (rispettivamente il 58.6% e il 41.4%). Quello che sorprende è che questi dati riflettono invece quelli riguardanti i *reportage* in diretta (rispettivamente il 43.9% e il 56.1%). Tuttavia, la ragione non può essere la medesima dei reportage, vale a dire una complessità grammaticale più elevata rispetto alla media. Come è già stato detto, il sommario è caratterizzato da forse la più bassa complessità grammaticale dei *move* considerati. La ragione potrebbe piuttosto essere l'eliminazione di interi sintagmi che, come nel caso delle omissioni di micro-unità, comporta un'equa omissione sia dei morfemi grammaticali, sia di quelli lessicali.

Per quanto riguarda gli errori, la loro incidenza è comprensibilmente la più bassa della categoria, 1.2%, appena lo 0.1% in meno rispetto alle previsioni meteorologiche, per

motivi che sono da ricercare nell'assenza di parole inaspettate nel TP. Tuttavia, gli errori da parte del software rappresentano un dato secondo solo a quello delle previsioni meteo (58.9%), simile a quello dei titoli. La ragione è da trovare nel fatto che il sommario è generalmente l'ultimo *move* che un rispeaker deve sottotitolare. Sia la voce del rispeaker, sia la memoria del software sono sovraccaricate e alcuni errori, che in altri *move* non sarebbero prodotti, sono inevitabilmente proiettati sullo schermo.

4.7 Conclusioni

Dall'analisi contrastiva tra il TP e il TA di un corpus di otto ore di BBC News emerge un quadro complesso delle strategie utilizzate dai rispeaker della BBC per sottotitolare in tempo reale questo programma. Di conseguenza, anche l'interpretazione dei dati estrapolati non risulta essere semplice. Tuttavia, alcune linee guida chiare possono essere ricavate. Innanzitutto, i rispeaker della BBC propendono per una resa il più possibile *verbatim* del TP, malgrado gli ostacoli che si frappongono al buon esito di questo obiettivo. Ecco quindi che finché il rispeaker non soffre dello stress ambientale in cui si trova a lavorare e le condizioni testuali del programma lo consentono (velocità di eloquio non superiore alle 180 parole al minuto, buon equilibrio tra densità lessicale e complessità grammaticale, buon equilibrio tra informazioni nuove e informazioni note, forma grammaticalmente corretta e coesa e pragmaticamente coerente), la politica editoriale della BBC è di trascrivere ortograficamente il TP (47.4% dei casi). A questa tendenza si sottraggono alcuni *step* del testo, che vengono sistematicamente omessi in quanto considerati evidentemente come semioticamente irrilevanti (le immagini compensano *in toto* l'informazione fornita verbalmente).

Al modificarsi delle caratteristiche testuali del TP, i rispeaker sono però obbligati a scostarsi dalla ripetizione fedele del TP e adottano delle strategie che consentono loro di aggirare o risolvere il problema contingente pur mantenendo un'apprezzabile qualità del TA. In linea generale, essi riescono sempre a rendere le idee del TP, ma può succedere che in quei passaggi più rapidi e lessicalmente più densi degli altri qualche unità concettuale sia omessa a discapito della coerenza e della coesione del TA. Questo avviene principalmente quando il *décalage* tra il TP e il TA deve essere ridotto per motivi dipendenti dalla resa del rispeaker in quel passaggio o dalla semiotica o dalla sola velocità di eloquio del TP. Per quanto riguarda tutti gli altri casi (33.7% in media), le strategie adottate dipendono

fortemente dal tipo di variazione dall'*optimum* delineato e quindi dal genere testuale che il rispeaker si trova a sottotitolare. In *BBC News*, si trovano a convivere numerosi generi testuali come le interviste, le conferenze stampa, i *reportage* in diretta, i servizi pre-registrati, la lettura dei titoli, del sommario e delle previsioni meteorologiche, frammenti di telecronache, ecc. Ognuno di questi generi presenta delle caratteristiche testuali ben definite in base a criteri precisi.

La prima distinzione va fatta tra testo orale-scritto (scritto per essere letto) e testo orale-orale (preparato e prodotto senza affidarsi alla lettura). A queste due macro-categorie appartengono tutti i generi testuali presenti in *BBC News* con gradi diversi a seconda delle caratteristiche di ciascuno, che sono una combinazione delle summenzionate condizioni testuali. Innanzitutto, quando il testo è lessicalmente più denso della media, così come quando è più veloce della media e quindi al di sopra delle possibilità sia del software, sia del rispeaker, la tendenza di quest'ultimo è di comprimere (62.2% dei casi di alterazione del TP) od omettere (37.8%) qualche micro-unità. Nel primo caso, il rispeaker farà ricorso a sintesi o riformulazioni che andranno a ridurre in egual misura morfemi grammaticali (49.5% dei casi di compressione) e morfemi lessicali (50.5%). Nel secondo caso, il rispeaker cercherà di limitare al minimo le omissioni di informazioni rilevanti, andando a eliminare morfemi grammaticali (86.2% dei casi di omissione di micro-unità) piuttosto che morfemi lessicali (13.8%). Quando poi il testo è più grammaticalmente complesso rispetto alla media, il rispeaker tende a sfoltirlo tramite un uso intelligente non tanto dell'omissione di quei morfemi puramente accessori (come nel caso del pronome relativo oggetto 'that' per introdurre una relativa) o delle espansioni grammaticali di alcuni termini in seguito all'omissione di un morfema grammaticale, comunque utilizzate, ma soprattutto della coordinazione e della lessicalizzazione. Quando invece il testo presenta più informazioni nuove di quante il rispeaker, e conseguentemente il software, non sia già a conoscenza, lo stress del rispeaker aumenterà e così la sua tendenza a ripetere il TP il più fedelmente possibile, anche qualora vi sia un alto tasso di oralità. Qualora sia necessario ridurre, si nota un atteggiamento prudente di predilezione dell'omissione di micro-unità rispetto alla sintesi e alla sinonimia. Quando infine il testo è maggiormente orale-orale e presenta quindi numerosi tratti dell'oralità, il rispeaker tende ad adattarlo alla lettura, eliminando tutti quei tratti non utili alla lettura (ripetizioni, false partenze, auto-riformulazioni, intercalari, ecc.) ed espandendo le forme lessicali contratte nella loro forma standard.

Da questa sintesi delle strategie dei rispeaker della BBC emerge un aspetto interessante che potrebbe contribuire enormemente al dibattito attorno la politica editoriale da adottare quando si parla di sottotitolazione per non-udenti, sia in diretta (come in questo caso), sia in pre-registrato. Voci discordanti tendono infatti a opporre due definizioni diverse di accessibilità al TP: da una parte, viene invocata la resa *verbatim*, quindi una trascrizione ortografica, del TP in nome delle pari opportunità tra udenti e non-udenti e dall'altra, una riformulazione del TP in nome della leggibilità del TA, diamesicamente diverso. Mentre la prima posizione comporta *a posteriori* l'obiettiva difficoltà per il pubblico non udente di leggere un sottotitolo alla stessa velocità in cui il pubblico udente ascolta il relativo testo, preparato appositamente per essere da esso ricevuto, la seconda fa sollevare dubbi circa l'effettiva 'inferiorità' delle competenze del lettore dei sottotitoli rispetto al tele-ascoltatore (che si traduce purtroppo, nelle rivendicazioni delle associazioni in difesa dei non-udenti[117], in 'inferiorità' delle persone non-udenti rispetto alle persone udenti)[118] e accuse di compassione e pietismo nei confronti dei fautori[119]. Quale può essere quindi il contributo dell'analisi appena riportata a questo apparentemente inconciliabile scontro? Dai dati emerge chiaramente la volontà dei rispeaker della BBC di ripetere il TP il più fedelmente possibile. Questo però non è sempre possibile ed essi sono quindi obbligati ad attingere a una serie di strategie volte a far fronte alle difficoltà contingenti e che si traducono in un adattamento del TP al mezzo usato per veicolare il TA. Quest'operazione comporta l'alterazione della forma del TP, ma non il contenuto, che è alternativamente veicolato o da una forma diversa (spesso quantitativamente ridotta) o dalle altre componenti semiotiche del TP che rimangono inalterate nel TA (componenti video verbali e non verbali). Così facendo, tutte le unità concettuali del TP sono veicolate nel TA, pur essendo le singole parole talvolta sacrificate in nome di una maggiore leggibilità del TA. Alla luce di questo spostamento dell'attenzione dalle singole parole alle unità concettuali, una riscrittura accettabile delle due posizioni summenzionate potrebbe essere la seguente:

- sottotitolazione *verbatim*, vale a dire la resa (soprattutto per ripetizione, ma anche per omissione, esplicitazione, sintesi, riformulazione, ecc.) di tutte le unità concettuali del TP nel TA;

117 Cfr. Mereghetti 2006.
118 Cfr. D'Ydewalle 1987: 321.
119 Cfr. Donaldson 2004.

- sottotitolazione *non verbatim*, ossia l'omissione, l'esplicitazione, la sintesi, la riformulazione, ecc. e laddove possibile la ripetizione delle unità concettuali del TP nel TA, sempre avendo come criterio guida la leggibilità dei sottotitoli.

Come è evidente, le due posizioni si sono avvicinate sensibilmente e la loro differenza sembra essere molto meno netta. Da qui, si potrebbe tentare un passo ulteriore verso la formulazione di standard nazionali e internazionali in grado di garantire una maggiore facilità di esportazione dei sottotitoli prodotti da un'emittente all'altra, in una filosofia dello scambio purtroppo non ancora presente.

Capitolo 5 - Per una piena accessibilità del TG ai sordi segnanti italiani

5.1 Introduzione

Una volta descritto il rispeakeraggio, abbozzato un quadro teorico all'interno del quale poter inserire gli studi sul rispeakeraggio e applicato un modello di analisi allo studio delle migliori prassi in materia, è giunto ora il momento di applicare le linee guida derivanti dai risultati appena ottenuti al contesto italiano. Si tratta di un corollario necessario a uno studio che non si vuole solo teorico e astratto, ma anche di utilità accademica e professionale, con enormi potenzialità sociali.

Nel caso specifico, si è optato per un lavoro mirato e pratico, che coinvolgesse tutti gli attori necessari all'applicazione del rispeakeraggio: comunità scientifica, rispeaker, comunità sorda ed emittente televisiva. Con questo spirito di collaborazione è nato il progetto SALES (Sottotitolazione Simultanea per l'Apprendimento Linguistico, l'Emancipazione e la Sicurezza dei Sordi), portato avanti in collaborazione con il VOICE project della Commissione Europea[120], il Subtitle Project dell'università di Bologna[121], il dottorato in Lingua Inglese per Scopi Speciali dell'università di Napoli Federico II, l'Ente Nazionale Sordi (ENS) della regione Emilia Romagna e la televisione di Stato della Repubblica di San Marino (RTV). L'obiettivo del progetto è di ridurre l'emarginazione sociale delle persone audiolese tramite il rispeakeraggio dei prodotti televisivi dell'emittente sammarinese in particolare e delle emittenti di lingua italiana in generale.

Per raggiungere quest'obiettivo, si è prima cercato di analizzare il potenziale bacino di utenza dell'emittente pubblica sammarinese tramite una serie di questionari e test di natura linguistica. In secondo luogo, visto che è maturata la scelta di restringere il pubblico di destinazione ai sordi segnanti e in seguito a un'intuizione sorta nel corso dei test, si è passati all'analisi delle interpretazioni in Lingua dei Segni Italiana (LIS) di diversi telegiornali nazionali, per cercare di individuare delle strategie utili al rispeakeraggio in italiano. Da queste analisi, che hanno arricchito le linee guida illustrate nei capitoli precedenti, è stata stilata una dettagliata lista di strategie per ogni categoria linguistica (lessico, sintassi, semantica e pragmatica) la cui validità è stata testata in un esperimento unico nel suo genere, la sottotitolazione in diretta tramite rispeakeraggio del secondo

120 Cfr. http://voice.jrc.it
121 Cfr. www.subtitleproject.net

confronto televisivo, nel 2006, tra due candidati Premier alle elezioni per la formazione del nuovo governo nazionale.

5.2 Il bacino di utenza

Tenere in considerazione il pubblico di destinazione è uno degli aspetti che maggiormente influenza la produzione di sottotitoli, soprattutto se progettati per scopi speciali. Secondo Nord (2000: 195),

> (t)idea of the addressee the author has in mind is a very important […] criterion guiding the writer's stylistic or linguistic decisions. If a text is to be functional for a certain person or group of persons, it has to be tailored to their needs and expectations. An "elastic" text intended to fit all receivers and all sorts of purposes is bound to be equally unfit for any of them, and a specific purpose is best achieved by a text specifically designed for this occasion.

Alla luce delle competenze del rispeaker descritte nei capitoli precedenti, risulta chiaro che il tempo di permanenza dei sottotitoli sullo schermo e la qualità e quantità di un'eventuale riformulazione da parte del rispeaker sono i fattori maggiormente influenzati dalle competenze linguistiche dell'utenza finale. Ecco quindi che un'introduzione alla sordità in generale e al potenziale bacino di utenza della RTV in particolare sembra opportuno.

La sordità è un concetto molto complesso e che spesso viene confuso con la semplice carenza di udito. In realtà i fattori sono numerosi e molto diversificati gli uni dagli altri. Dal punto di vista prettamente medico, la sordità si definisce in base a fattori come:

- la quantità calcolata in decibel di udito perduto (20-40 = lieve, 40-70 = media, 70-90 = grave, >90 = profonda);
- la qualità della ricezione del residuo di udito;
- la capacità da parte del soggetto di distinguere i suoni percepiti da altri suoni affini;
- l'età in cui il soggetto ha perso l'udito: prima, durante o dopo l'apprendimento linguistico (sordità pre-/ peri-/ post-linguale)[122]. Poche persone sorde post-linguali tendono a rapportarsi e a identificarsi con gli altri sordi.

122 In termini anagrafici, si considera pre-linguale l'età che va da 0 a 12 mesi e peri-linguale da 1 a 12 anni.

Inoltre, principalmente all'interno della comunità sorda pre- e peri-linguale grave o profonda, molti sono i fattori più specificamente culturali che possono incidere considerevolmente sulle competenze linguistiche di ciascuna di queste persone. Essi sono:

- il metodo linguistico con cui sono state educate (segnante o oralista);
- la lingua madre o semplicemente quella predominante (Lingua dei Segni o lingua orale[123]/labiolettura o bilinguismo);
- la comunità linguistico-culturale in cui si identificano (comunità dei Sordi[124] o comunità degli udenti).

In termini demografici, la percentuale di persone medicalmente considerate audiolese in Europa varia dal 5 al 15% dell'intera popolazione[125]. In questo dato, rilevante è l'incidenza delle persone anziane presbiacusiche, che perdono cioè l'udito progressivamente e per ragioni dovute all'invecchiamento delle cellule dell'orecchio. Se si valutano invece solo coloro che hanno sviluppato una sordità severa o profonda in età pre-linguale o peri-linguale, il dato scende drasticamente allo 0,1% della popolazione. In Italia, di questo 0,1%, solo una parte si dichiara membro della comunità sorda segnante, mentre la maggior parte delle persone al di sotto dei trent'anni preferisce alla Lingua dei Segni Italiana, il metodo oralista[126]. Dal punto di vista linguistico, questa distinzione si traduce in due sottocategorie di sordi:

- i sordi segnanti, che parlano cioè la LIS come loro lingua madre e l'italiano come prima lingua straniera;
- i sordi oralisti, che leggono, scrivono, parlano e, grazie alla labiolettura o ad ausili tecnici[127], capiscono l'italiano parlato.

123 Il termine 'orale' è qui impiegato in contrapposizione concettuale al termine 'segnico', riferendosi in generale a tutte le lingue che dispongono del canale fono-acustico e, laddove disponibile, di quello grafico.
124 Il termine Sordo indica l'appartenenza alla comunità dei sordi segnanti, che oltre alla lingua condivide anche una cultura specifica con gli altri membri della comunità, profondamente diversa da quella che, per contrasto, è definita la comunità degli udenti. Dal punto di vista legislativo, in Italia sono considerate sorde le persone che hanno perduto almeno 70 dB di udito prima dei 12 anni di età. Cfr. www.ens.it
125 Il censimento delle persone sorde dipende dal riscontro dei singoli medici, che, oltre a utilizzare strumenti e criteri diversi, non possono obbligare i propri pazienti a sottoporsi a un esame audiometrico. Non esiste pertanto un dato esatto circa la popolazione sorda, ma solo delle deduzioni statistiche.
126 Da qualche tempo, si sta cercando di diffondere il metodo bilingue, ma ancora molte sono le reticenze da parte delle associazioni in difesa dei diritti dei sordi alla diffusione di questo metodo. Cfr. http://www.webxtutti.it/cond_sordi.htm (ultimo accesso 28/11/2008).
127 Cfr. www.fiadda.it

Da questa breve disamina, risulta chiaro il significato delle parole di Nord circa un eventuale tentativo di soddisfare le esigenze di un bacino di utenza appartenente a una gamma di profili linguistici così vasta. Pur conscio dell'impossibilità di soddisfare le esigenze di tutti i telespettatori sordi, compito più semplice invece sembra essere delineare un profilo medio all'interno di una determinata sottocategoria linguistica. Sulla base di una ricerca che testi le competenze linguistiche di un campione il più possibile rappresentativo della sottocategoria in esame, sarà poi possibile fissare alcune linee guida per la produzione di sottotitoli accessibili alla sottocategoria linguistica di sordi scelta. Tuttavia, bisogna tenere in considerazione che optare per una categoria o per l'altra implica pesanti ripercussioni sul lavoro del sottotitolatore. C'è infatti una sensibile differenza tra sottotitolare per persone che non riescono a sentire nessun suono del TP e che conoscono l'italiano come una lingua straniera e sottotitolare per un pubblico la cui lingua madre è l'italiano e che percepisce la quasi totalità delle frequenze del parlato. Nel primo caso, la sottotitolazione è un ausilio dal forte impatto cognitivo e 'traduttivo', nel secondo caso, invece, il sottotitolo è un semplice strumento per controllare che la comprensione del testo orale sia stata accurata. Visto che l'obiettivo del progetto SALES è di rendere accessibili i programmi televisivi in diretta al maggior numero possibile di persone, la scelta della categoria su cui tarare i sottotitoli è caduta sulla comunità dei sordi segnanti pre-linguali. Le ipotesi alla base di questa scelta sono state le seguenti:

- dal punto di vista linguistico, persone per cui l'italiano è una lingua straniera e che non riescono a percepire nessuna frequenza del parlato da prima che iniziasse il loro apprendimento della lingua materna orale si pongono al polo negativo del continuum rappresentato dalle sottocategorie dei sordi in materia di comprensione di un testo scritto (il sottotitolo);
- nonostante Nord (2000) dica chiaramente che un testo "elastico", che abbia l'obiettivo di soddisfare le esigenze di molteplici tipologie di pubblico, non soddisfa le esigenze di nessuna tipologia, tarare la sottotitolazione di un programma televisivo su una categoria di persone che si pone al polo negativo del continuum in materia di comprensione di un sottotitolo potrebbe essere di utilità a tutte le altre sottocategorie, quanto meno in termini di comprensione del testo in questione.

5.3 La ricerca

Come si è visto, per poter soddisfare le esigenze e le aspettative di una determinata tipologia di pubblico attraverso la produzione di sottotitoli mirati è necessario conoscere le competenze linguistiche di questa categoria. Il gruppo di ricerca incaricato di stabilire il profilo linguistico medio della sottocategoria dei sordi segnanti pre-linguali è stato costituito da un ricercatore in materia di sottotitoli per sordi, un'interprete di lingua dei segni socio dell'ENS, un educatore di un istituto superiore per sordi, un'insegnate di sostegno delle scuole elementari e sei presidenti dei circoli ENS[128] delle province coperte dal segnale della RTV.

Per quanto riguarda la costituzione del *focus group*, invece, una prima difficoltà è immediatamente emersa nel gestire l'eterogeneità dei volontari. La maggior parte di questi era sorda segnante pre-linguale, ma alcuni segnanti non sapevano affermare con certezza quando avevano perduto l'udito. Altri segnanti ancora avevano perduto l'udito in età peri-linguale. Una cospicua quantità di sordi inoltre non era segnante, ma aveva comunque perduto l'udito in età pre-linguale. È stato dunque deciso di accettare, in un primo momento, ogni volontario, salvo poi scartare i suoi risultati o collocarli insieme a quelli di uno o più *control group*, qualora fossero sensibilmente discordanti dai risultati del resto del gruppo.

Dal punto di vista metodologico, inizialmente è stato analizzato in dettaglio il profilo sociale di ognuno e poi ne sono state testate le competenze linguistiche, sia in termini di comprensione dei testi che sono stati loro sottoposti, sia di velocità di lettura degli stessi. Visto il numero elevato dei volontari (197), è stato deciso di suddividere il *focus group* in piccoli gruppi, per lo più in base alla zona geografica di appartenenza, in modo da limitare gli spostamenti di ognuno. Il tutto è stato video-registrato con il consenso dei partecipanti in modo da poter monitorare in seguito le singole reazioni. Dai risultati è stato possibile delineare il profilo linguistico dell'utente tipo e conseguentemente ottenere le linee guida per un rispeakeraggio mirato.

5.3.1 Il profilo sociale

Come già anticipato, il primo passo dopo la costituzione del *focus group* è stato il tentativo di delineare il profilo sociale medio dell'utente sordo segnante. Per ottenere questo

128 L'ENS è l'unica associazione nazionale in difesa dei sordi segnanti. Altre associazioni nazionali come la FIADDA sono invece propense alla promozione dell'oralismo come unica modalità di riabilitazione del soggetto sordo. Per lo scopo della presente ricerca, si è ritenuto più opportuno coinvolgere solo l'ENS. Le sedi provinciali dell'ENS coinvolte sono quelle di Rimini, Forlì-Cesena, Ravenna, Ferrara, Bologna e Modena. Del focus group hanno fatto parte anche un gruppo di sordi della Repubblica di San Marino che non sono però riuniti in associazione.

risultato, si è inizialmente spiegato al *focus group* il fine ultimo della ricerca in corso. Motivati dall'idea di contribuire attivamente e in maniera molto influente al risultato finale, i sordi che hanno preso parte alla ricerca hanno risposto alle domande[129] di un primo questionario scritto, nonostante la loro generale predilezione per l'uso della lingua dei segni. Dal questionario sono emersi molti dati personali interessanti che hanno contribuito a un primo orientamento della ricerca[130] e in particolare:

- età: l'età media del *focus group* è di 40,1 anni. 38 persone (19,29%) hanno tra i 20 e i 30 anni; 61 (30,96%) tra i 30 e i 40 anni; 58 (29,44%) tra i 40 e i 50; 26 (13,2%) tra i 50 e i 60; 13 (6,6%) tra i 60 e i 70. Completa il gruppo una signora di 93 anni (0,51%);
- anni di sordità: la maggior parte del *focus group* (178, corrispondente al 90,36%) dichiara di aver perso l'udito in età pre-linguale o peri-linguale o non ricorda esattamente. Non ricorda comunque di aver mai udito o parlato. Solo pochi ricordano di aver parlato, di cui 11 (5,58% del totale) dichiarano di aver perduto l'udito durante il loro secondo anno di vita, uno (0,51% del totale) di averlo perduto durante il terzo anno di vita, due (1,02% del totale) durante il quarto, quattro (2,05% del totale) durante il quinto e uno (0,51% del totale) durante il settimo anno di vita;
- grado di istruzione: nel 1880, il cosiddetto Consiglio di Milano stabilì che, da quel momento in poi, nelle scuole speciali per sordi il metodo da seguire per l'insegnamento avrebbe dovuto essere quello oralista e non più la lingua dei segni. Tuttavia, soprattutto durante il regime fascista, molta importanza continuava ad essere accordata alla LIS, che era usata in tutti i contesti para- ed extra-scolastici. Per tutto il corso degli anni Settanta, infine, si è compiuto il passaggio degli studenti sordi dalle scuole speciali alle scuole 'normali'. Questi eventi hanno influito notevolmente sull'istruzione dei sordi perché, viste le maggiori difficoltà di un sordo rispetto a un normo-udente ad affrontare il percorso scolastico ed

129 Per agevolare le risposte da parte di molti segnanti, le domande erano prevalentemente chiuse o a risposta multipla. Solo una, circa le opinioni personali sul servizio sottotitoli offerto dalla televisione italiana pubblica e privata, era aperta. Oltre ad avere un'idea generale circa la loro visione delle cose in materia, quest'ultima domanda aveva anche l'obiettivo di aggiornare eventualmente i dati ottenuti dal CNR in occasione dell'inizio dei lavori del servizio sottotitoli della RAI. Cfr. Volterra 1986.

130 Oltre ai risultati presentati, il questionario comprendeva anche domande che sono state meno rilevanti ai fini della ricerca come il luogo di nascita, il domicilio, il genere, l'occupazione, il numero di sordi all'interno del nucleo familiare, le abitudini in materia di lettura in generale e di uso dei sottotitoli per sordi in particolare.

eventualmente universitario[131], la maggior parte dei sordi componenti il *focus group* si è limitata agli anni di scuola dell'obbligo imposti e in particolare: otto anni per le persone tra i 20 e i 40 anni circa, cinque per quelli tra i 40 e i 70 anni circa e tre anni di scuola dell'obbligo per la signora 93enne. Tuttavia, alcuni sordi, segnanti e oralisti, hanno continuato anche oltre la scuola dell'obbligo tanto che il 15,48% (13 persone) delle 84 persone tra i 40 e i 60 anni hanno frequentato anche le scuole medie inferiori (otto anni di istruzione) e il 26,26% (26 persone) delle 99 persone al di sotto dei 40 anni ha frequentato anche le medie superiori e due persone sono studenti universitari;

- lingua materna: gli eventi sopra descritti circa l'evoluzione della scuola dell'obbligo per le persone sorde hanno inoltre creato una spaccatura profonda tra i sordi pre- e peri-linguali. In termini molto generali, visto che la frequentazione dei corsi si limitava agli anni della scuola dell'obbligo, mentre il resto della vita veniva spesso trascorso all'interno degli istituti o comunque all'interno della comunità sorda dove la LIS era il metodo di comunicazione più naturale, le persone sorde con più di quarant'anni sono per lo più segnanti. I più giovani invece sono *grosso modo* suddivisi in due macro-categorie: sordi segnanti nati in famiglie di sordi e che hanno l'italiano come prima lingua straniera[132]; sordi oralisti che hanno l'italiano come lingua materna e che conoscono la LIS solo come prima lingua straniera o la ignorano completamente. Quanto al *focus group*, 118 (59,9%) persone sono sorde segnanti e sono nate da almeno un genitore sordo. 53 (26,9%) persone sono ugualmente madrelingua LIS, ma sono nati da genitori udenti. Infine, 26 sordi (13,2%) sono nati da genitori udenti e hanno l'italiano come lingua materna. Quanto al bilinguismo, 151 (76,65%) hanno dichiarato di utilizzare le due lingue indifferentemente, di cui 109 (72,19% del parziale) sono madrelingua LIS e i restanti 42 (27,81% del parziale) hanno acquisito la LIS solo in un secondo momento. Questi ultimi dichiarano in ogni caso di avere maggiori competenze in italiano che in LIS.

131 Cfr. Pirelli 2006.

132 Secondo alcuni insegnanti del CNR, dopo aver acquisito una buona padronanza della lingua inglese, alcuni sordi italiani, sia segnanti sia oralisti preferiscono comunicare (via posta elettronica, chat o SMS) in inglese piuttosto che in italiano. Una delle probabili ragioni è la relativa maggiore semplicità morfo-sintattica della lingua inglese rispetto all'italiano. Inoltre, per quanto riguarda in particolare i segnanti, c'è una sostanziale somiglianza strutturale tra le lingue dei segni e la lingua inglese.

5.3.2 Le competenze linguistiche

Dopo aver identificato e discusso il profilo sociale del *focus group* all'interno del gruppo di ricerca, è stata anche esaminata la modalità con cui sarebbero stati condotti i test successivi volti alla comprensione delle reali esigenze linguistiche del *focus group* stesso.

Metodologia

Per essere certi che i test risultassero utili a tale scopo, è stato adottato un doppio approccio: grazie a un approccio quantitativo, si sono potuti ottenere dati precisi circa la velocità di lettura dei singoli volontari e dati indicativi circa la percezione della loro comprensione del testo audiovisivo appena letto. Grazie a un approccio qualitativo[133], è stato possibile ottenere dati certi circa l'effettiva comprensione da parte dei singoli volontari di un testo audiovisivo sottotitolato.

Partendo dal presupposto che gli utenti di un prodotto audiovisivo sono *active producers of meaning* (de Certeau 1990) e che decodificano il testo a seconda del contesto socio-culturale in cui avviene la visione e della maniera in cui la vivono, è risultato immediatamente necessario avere un'idea più precisa dell'uso che l'utente avrebbe fatto del sottotitolo.

Per raggiungere tale scopo, è stato necessario individuare la tipologia di testo da utilizzare come materiale dell'esperimento, selezionarlo e infine sottotitolarlo compatibilmente con gli obiettivi della ricerca. Visto che il progetto SALES è nato con l'intento di trovare la migliore soluzione per sottotitolare intra-linguisticamente i notiziari della RTV, la scelta del materiale è caduta inevitabilmente su alcuni estratti dei vari notiziari dell'emittente sammarinese. La selezione è stata poi effettuata in base a due criteri: le tematiche comunemente toccate dalla RTV (politica, economia, cultura, sport e cronaca locale) e i fini della ricerca. Come si è già visto, per ottenere delle linee guida in materia di sottotitolazione per sordi è necessario anche operare una riduzione del TP ai fini di una piena accessibilità del prodotto audiovisivo da parte del pubblico a cui questi sottotitoli sono destinati. Non sapendo in che termini operare tale sintesi, è stato necessario selezionare

133 In particolare è stato preso in prestito e adattato l'approccio etnografico dei reception studies. Cfr. de Certeau 1990.

cinque notizie per ogni tematica, ognuna delle quali è stata sottotitolata secondo criteri di riformulazione[134] diversi. In particolare, si è optato per cinque livelli di difficoltà:

- difficoltà alta: trascrizione ortografica del TP;
- difficoltà medio-alta: riformulazione lessicale[135];
- difficoltà media: riformulazione lessicale e sintattica[136];
- difficoltà medio-bassa: riformulazione lessicale, sintattica e semantica[137];
- difficoltà bassa: riformulazione lessicale, sintattica e semantica con riduzione quantitativa del TP del 50%.

Prima di iniziare l'esperimento, è stato chiesto a tutti i sordi il permesso di video-registrare l'intero esperimento. Una volta ottenuto il consenso da tutti, è stato spiegato loro il funzionamento del test. Al singolo volontario sarebbe stato chiesto di scegliere un argomento tra i cinque disponibili (politica, economia, cultura, sport e cronaca locale) e sarebbero poi stati mostrati loro le cinque versioni precedentemente sottotitolate relative all'argomento scelto. Per visualizzare il primo sottotitolo e la scena o parte di scena ad esso corrispondente, il volontario avrebbe dovuto semplicemente cliccare il tasto sinistro del mouse o la freccia in basso o a destra presente sulla tastiera del computer. Per continuare la lettura dei sottotitoli e la visione delle rispettive scene o parti di scena, il volontario avrebbe dovuto ripetere l'operazione appena descritta fino alla fine del filmato tante volte quante erano le diapositive[138]. Quanto alla registrazione del tempo di lettura, il cronometro sarebbe partito nel momento in cui il volontario fosse passato dal primo sottotitolo, contenente esclusivamente il titolo del filmato in questione (e nessuna immagine associata), al secondo, contenente la prima scena sottotitolata. Allo stesso modo, il cronometro sarebbe stato

134 Il tentativo è stato di semplificare linguisticamente il testo per permettere una più immediata comprensione del TP senza peraltro spiegare i concetti. Così facendo il gruppo di ricerca era conscio di non riprodurre esattamente nel sottotitolo il TP (cfr. Gambier 1992). Tuttavia, si è tentato di giocare con la forma rispettando il più possibile le singole unità concettuali.

135 Questa tipologia di riformulazione è stata operata in linea con la Banca Dati del Vocabolario Di Base della lingua Italiana (De Mauro 1997) in cui vengono riportate le 5000 parole che ogni parlante italiano dovrebbe capire, conoscere e usare regolarmente. Sempre nel pieno rispetto delle unità concettuali, le parole tecniche e i nomi propri non sono stati riformulati.

136 Si è cercato di rispettare il più possibile la struttura sintattica di base (S-V-O) e di evitare troppi schemi sintattici.

137 Si è cercato di evitare inutili tropi, nel pieno rispetto di quelli più comuni o lessicalizzati.

138 Grazie a questa procedura, ogni filmato è stato reso totalmente dipendente dalla velocità di lettura del singolo. La sottotitolazione è stata effettuata rispettando le norme grafiche di impaginazione seguite dal servizio sottotitoli della RAI, intuitivamente a loro più familiari rispetto alle regole seguite da altri fornitori di sottotitoli da loro meno utilizzati. Il tutto è stato caricato su PowerPoint.

stoppato con il passaggio dalla penultima diapositiva, contenente l'ultima scena sottotitolata, all'ultima, contenente solo la parola 'fine'.

Alla fine della visione di ciascun filmato, al volontario sarebbe stato chiesto se avesse compreso o meno il significato generale del testo audiovisivo appena visto dopodiché gli/le sarebbero state poste dieci domande di comprensione generale nella lingua di sua scelta (italiano o LIS). Solo i risultati (velocità di lettura e comprensione del TA) di chi avrebbe risposto correttamente ad almeno sei domande su dieci sarebbero stati presi in considerazione nella valutazione finale. Terminato il proprio turno e alla fine di tutti i test da parte dei volontari di un gruppo, si sarebbe passati a una discussione di gruppo tra il gruppo di ricerca e il *focus group* circa le difficoltà riscontrate e le relative possibili soluzioni.

I risultati

Uno dei primissimi dati interessanti emersi dall'analisi quali-quantitativa è stata la sostanziale identità di approccio alla lettura e alla comprensione del TA da parte dei sordi segnanti pre-linguali, dei sordi peri-linguali e dei sordi oralisti[139]. Le differenze riscontrate sono da attribuirsi alle abitudini in materia di lettura e al grado di istruzione[140] più che alla lingua materna[141]. Questo dato ha comportato la non esclusione degli oralisti e dei sordi peri-linguali dal *focus group*.

Per quanto riguarda più da vicino i dati, dopo aver scelto la tematica di suo maggiore interesse tra le cinque disponibili e aver visionato le cinque versioni, ogni volontario ha risposto, in LIS o in italiano, a dieci domande di comprensione generale del testo ottenendo, per ogni grado di difficoltà, i seguenti risultati:

- difficoltà alta: solo 13 persone su 197 (il 6,6%) hanno risposto ad almeno sei domande su dieci. Si tratta perlopiù di persone abituate a utilizzare il servizio

139 Come accennato precedentemente, esiste una differenza culturale molto profonda tra coloro che hanno come lingua materna l'italiano e coloro che hanno da sempre parlato la lingua dei segni in casa. Anche se oggi la maggior parte dei sordi è bilingue, molti hanno come lingua dominante quella della comunità (udente o sorda segnante) in cui sono cresciuti e, nella maggior parte dei casi, vive. Tuttavia, studi del CNR (Caselli et al. 1994 e Volterra 1986) dimostrano che non esistono differenze sostanziali tra sordi segnanti e sordi oralisti dal punto di vista della competenza linguistica in italiano. Il risultato appena descritto sembra confermare questo dato.

140 Precedentemente è stato sottotilineato che i dati riguardanti le abitudini in materia di lettura sono stati esclusi dall'analisi finale dei dati. In realtà, la maggior parte dei sordi ha dato risposte poco precise limitandosi a scrivere parole come 'poco', 'per niente' o 'abbastanza' nella casella riservata al numero di ore settimanali e giornaliere dedicate alla lettura. Altri ancora non hanno risposto a questa domanda. Tuttavia, è stato possibile notare un ricorrente parallelismo tra il livello di istruzione e l'indicazione di massima circa le proprie abitudini di lettura.

141 Fino alla fine dei test, i dati riguardanti la lingua dominante e il metodo con cui sono stati educati i singoli sordi sono stati attribuiti a due categorie separate. Essi sono stati accorpati in un secondo momento, quando sono emerse delle sostanziali similitudini tra le due categorie che non giustificavano tale differenziazione.

sottotitoli della RAI e di Mediaset e a vedere film in DVD, sotto i quarant'anni e con almeno 13 anni di istruzione. La velocità media di lettura è di 3,9 secondi per riga di sottotitolo talvolta accompagnata da subvocalizzazione udibile o da traduzione in segni di alcune parole del testo letto. Dopo i primi casi di subvocalizzazione e di traduzione in segni è stato esplicitamente richiesto a ciascuno di evitare questi atteggiamenti, suscettibili di compromettere la velocità di lettura. Tuttavia, soprattutto in questo livello di difficoltà, pochi sono riusciti a resistere alla tentazione. La giustificazione più volte addotta è stata la difficoltà di comprensione di alcune parole e strutture sintattiche;

- difficoltà medio-alta: 27 persone (13,71%) hanno risposto correttamente ad almeno sei domande su dieci. Come nel caso delle trascrizioni, si tratta di persone con un alto grado di istruzione e abitudini in materia di lettura favorevoli. Durante le interviste e la discussione in gruppo, è risultato inoltre chiaro che, anche in questo caso, la difficoltà di comprensione del testo, seppur inferiore rispetto al caso precedente, causava una certa frustrazione nei singoli soggetti e conseguentemente un rallentamento nella lettura dei sottotitoli, accompagnata, anche in questo caso, da subvocalizzazioni e traduzioni in segni percepibili nella maggior parte dei volontari. Gli errori più ricorrenti sono derivati da una confusione generalizzata tra il soggetto e l'oggetto. La velocità media di lettura non si discosta molto dal dato precedente: 3,6 secondi per riga di sottotitoli;
- difficoltà media: in seguito alla riformulazione sintattica nei sottotitoli, 121 persone (61,42%) hanno risposto correttamente a una media di 7,4 domande su dieci. Si tratta di un dato straordinario che prova che uno dei fattori che maggiormente influiscono sulla comprensione del testo è il grado di complessità sintattica dello stesso. Parallelamente a questo dato, la velocità media di lettura è scesa a 2,74 secondi per riga di sottotitoli. Nessuno ha subvocalizzato in maniera percepibile e solo pochi hanno tradotto in segni alcune parole di cui non conoscevano probabilmente il significato;
- difficoltà medio-bassa: 154 persone (78,17%) hanno compreso il testo, rispondendo correttamente a una media di 8.3 domande su dieci e totalizzando una velocità media di lettura di 2,51 secondi per riga di sottotitoli. Anche in questo caso,

nessuno sembra aver subvocalizzato o tradotto in segni il TA. Diversamente dal caso precedente, tutti hanno avuto l'impressione di aver capito il testo[142];

- difficoltà bassa: 163 persone su 197 (82,74%) hanno dimostrato di aver compreso il TA rispondendo a una media di 8,9 domande su dieci. Le restanti 24 persone hanno dichiarato di aver compreso il testo mentre leggevano, ma di aver dimenticato alcune informazioni in seguito a uno scarso interesse per la notizia e conseguentemente a una maggiore attenzione alla lettura del testo rispetto ai concetti. La velocità media di lettura è stata di 2,46 secondi per riga di sottotitoli.

Discussione

Da una prima rapida analisi dei risultati appena descritti è stato possibile trarre alcune considerazioni. Per quanto riguarda la comprensione del testo, risulta immediatamente chiaro che la trascrizione esatta del TP, richiesta a gran voce da tutte le associazioni in difesa dei sordi[143], non è la strada migliore per l'accessibilità dei notiziari. Sempre in termini di comprensione del testo, è altrettanto vero che nemmeno una riduzione quantitativa comporta un significativo miglioramento. Questo è valido sia dal punto di vista della percezione della comprensione del testo, sia da quello dell'effettiva comprensione. Piuttosto, a influire positivamente sulla comprensione del testo risulta essere una riformulazione sintattica, seguita da lontano dalle riformulazioni semantica e lessicale.

Quanto alla velocità di lettura, la media delle ultime tre tipologie di testo (riformulazione lessicale e sintattica, riformulazione lessicale, sintattica e semantica e riassunto del 50% del testo) si attesta attorno ai 2.5 e i 3 secondi per riga. In vista di una tendenza alla riformulazione sintattica come principio fondante della sottotitolazione per sordi all'interno del progetto SALES, questa dovrebbe essere la media da prendere generalmente in considerazione.

In seguito a un'analisi più approfondita dei risultati, fatta a partire dallo studio dei filmati, alcuni dati interessanti spiccano sugli altri. Dal punto di vista della comprensione del testo, è stato ulteriormente confermato il dato precedentemente riportato circa la sostanziale identità di approccio al testo da parte dei sordi segnanti e dei sordi oralisti. Inoltre, si è potuto osservare che sottotitoli che casualmente rispecchiavano la struttura morfo-sintattica

142 Un elemento di cui ci si è subito rammaricati è stata l'impossibilità di constatare quanto la sola riformulazione semantica abbia influenzato la comprensione generale del testo. Intuitivamente, meno della riformulazione sintattica.
143 Cfr. Mereghetti 2006.

della LIS erano più facili da leggere, rispetto ad altri sottotitoli che non presentavano questa caratteristica, sia per i sordi segnanti, sia per gli oralisti. Questo potrebbe significare banalmente che le lingue dei segni sono più semplici delle lingue orali e che semplificando sintatticamente queste ultime si ottiene la sintassi delle lingue dei segni. Quindi i sordi, segnanti e oralisti, farebbero meno fatica a leggere dei testi semplici rispetto a testi più complessi. Per quanto verosimile possa essere, questa ipotesi è in parte smentita dalla constatazione che nei casi in cui i sottotitoli non rispecchiavano la struttura della LIS, pur essendo semplificati sintatticamente, la lettura delle singole righe di sottotitoli risultava essere più lenta, probabilmente più macchinosa. Una chiave di lettura forse più esatta e in sintonia con eminenti studiosi del settore[144], potrebbe essere che le lingue dei segni, esattamente come le lingue orali, sfruttano appieno il potenziale comunicativo dei parlanti per esprimere qualsiasi tipo di concetto. Nel caso specifico, questo varrebbe per ogni tipologia di sordo, vale a dire ogni persona che percepisce il mondo attraverso un canale in meno rispetto ai normoudenti e che, di conseguenza, sfrutta maggiormente altri canali, *in primis* quello visivo. Indipendentemente dai meriti o demeriti della lingua dei segni, ai fini del progetto SALES questo dato ha permesso di ottenere un'indicazione chiara del tipo di riformulazione da operare quando si sottotitolano testi informativi per sordi: nei limiti imposti dalla grammatica e dalla pragmatica italiana, rispecchiare il più possibile la struttura morfo-sintattica della LIS costituisce, in termini di accessibilità, la migliore forma di sottotitolazione mirata ai sordi segnanti in particolare e ai sordi gravi e profondi in generale.

Sempre osservando le riprese, è stato possibile ottenere dati interessanti anche circa la velocità di lettura. Se è vero che la velocità di lettura media oscilla, nel migliore dei casi, tra i 2,5 e i 3 secondi per riga, una precisazione deve essere fatta a proposito del concetto di riga. È stato osservato infatti che sottotitoli costituiti da una sola riga erano letti a una media di 2,1 secondi, mentre sottotitoli costituiti da due righe erano letti a una media di 2,57 secondi per riga[145]. Intuitivamente, come sostengono d'altronde anche molti studiosi del settore[146], un dato contrario sarebbe più naturale. In altre parole, se è vero che per accorgersi dell'arrivo di un sottotitolo, e quindi per iniziare a leggerlo, un qualsiasi spettatore necessita di un certo lasso di tempo che va a sommarsi al tempo di lettura di ogni riga di sottotitolo, la

144 Cfr. Stokoe 1981 e Volterra 1986.
145 Le medie in questione si riferiscono ai sottotitoli del livello più basso di difficoltà, ma scarti simili sono stati riscontrati a tutti i livelli.
146 Cfr. Groner et al. 1990, Karamitroglou 1998 e Perego 2005.

deduzione logica vuole che più testo si trova in un unico sottotitolo, più veloce sarà la sua lettura rispetto alla stessa quantità di testo spalmata su due o più sottotitoli. Nel primo caso, sarà presente un solo tempo morto dovuto all'attesa, la presa di coscienza dell'arrivo e la messa a fuoco del sottotitolo in questione, mentre nel secondo caso, oltre alla velocità di lettura del testo, sono da considerare anche tanti tempi morti quanti sono i sottotitoli. Tuttavia, questa deduzione non sembra essere valida per il *focus group* in questione. Oltre al dato ufficiale circa la velocità di lettura di ogni sottotitolo, infatti, si è riscontrato che, in più di un'occasione, i volontari ricominciavano la lettura delle due righe di sottotitolo daccapo oppure si fermavano a metà della lettura o ancora si avvicinavano di più allo schermo, come se i caratteri si fossero improvvisamente rimpiccioliti. Una plausibile spiegazione può essere la maggiore difficoltà di gestire visivamente un sottotitolo di due righe rispetto a un sottotitolo di una sola riga. Durante la discussione finale, alcuni sordi hanno infatti lamentato la presenza di troppe parole nei sottotitoli, soprattutto in quelli di due righe. In questa confusione, molti hanno dichiarato di essersi persi e di aver scelto di continuare comunque la lettura per non perdere troppo tempo, mentre alcuni hanno scelto di ricominciare dall'inizio.

In conclusione, dalla lunga ricerca qui brevemente riportata sono emersi alcuni dati utili ai fini del progetto SALES e in particolare:

- dal punto di vista qualitativo, è necessaria una riformulazione quanto meno sintattica del TP;
- tale riformulazione deve essere operata tenendo in considerazione il più possibile la struttura della LIS;
- dal punto di vista quantitativo, bisogna tenere in considerazione una velocità media di lettura di 2,5-3 secondi per riga di sottotitolo;
- è preferibile un sottotitolo di una riga rispetto a un sottotitolo di due righe.

Se dal punto di vista della forma, la situazione sembra essere risolta una volta per tutte, resta ancora da chiarire il secondo punto: in che modo deve essere sintatticamente riformulato il TP? Risulta a questo punto scontato che il passo successivo, sebbene non previsto dal gruppo di ricerca, è stato un rapido sguardo alla struttura della lingua dei segni in contesto giornalistico di modo da ricavare alcune indicazioni precise circa la necessaria riformulazione da attuare.

5.3.3 L'analisi della LIS in contesto giornalistico

L'ultimo passo verso la produzione di linee guide per il rispeakeraggio del notiziario della RTV è stato caratterizzato da alcune prime difficoltà nella pianificazione del lavoro. Per quanto riguarda la scelta del materiale da analizzare, infatti, non c'è stato immediato accordo. Se, da una parte, sembrava scontato proporre l'analisi dell'interpretazione in LIS dei TG di RAI e Mediaset, visto che lo scopo del rispeakeraggio rispecchia proprio quello dell'interpretazione in LIS, sia l'interprete LIS, sia gli insegnanti che componevano il gruppo di ricerca hanno immediatamente sottolineato la differenza esistente tra l'uso fatto della LIS da parte dei sordi nei contesti di comunicazione spontanea e l'uso della LIS da parte degli interpreti simultanei al TG nazionale delle emittenti televisive italiane. In particolare, è stata messa in risalto proprio la diversa struttura sintattica seguita dagli interpreti LIS rispetto alla struttura standard. Come è intuibile, infatti, l'interprete LIS tende maggiormente a calcare la struttura sintattica dell'italiano per evitare di perdere troppe informazioni nel passaggio in LIS, che segue una struttura sintattica totalmente differente. Malgrado l'esattezza di questa osservazione, i presidenti delle sedi provinciali dell'ENS dell'Emilia-Romagna hanno assicurato che comunque le interpretazioni sono comprensibili a un segnante medio.

Vista l'impossibilità di protrarre ulteriormente le ricerche e visto che un rispeaker non potrebbe veicolare un messaggio ai segnanti meglio di un interprete LIS, proprio perché si trova a dover lavorare con l'italiano, la scelta del materiale è caduta comunque sull'interpretazione in LIS dei TG di RAI e Mediaset. Dopo aver video-registrato 14 notiziari, si è passati all'analisi grammaticale del TP e del TA in chiave contrastiva e all'annotazione delle regolarità ricorrenti nel processo traduttivo. Alla fine dell'analisi, sono state selezionate le strategie più adeguate al rispeakeraggio in italiano, sia in termini grammaticali (la strategia non doveva contrastare con la grammatica e la pragmatica della lingua italiana), sia in termini funzionali (la strategia rientrava nell'ambito delle strategie di riformulazione selezionate).

Dal punto di vista lessicale, un primo dato è emerso su gli altri: l'uso limitato all'essenziale di morfemi grammaticali. In specifico, non sono state riscontrate trasposizioni regolari di articoli, preposizioni e forme flesse. Sono invece regolarmente trasposti la maggior parte dei connettori e degli avverbi di luogo, di causa-effetto e di modo. Il resto

della traduzione è svolto dalla lessicalizzazione, laddove possibile, dal rispetto della struttura sintattica di base e dalla coordinazione sintattica. Quanto ai morfemi lessicali, tutti i sinonimi periferici di un lessema usati per motivi stilistici e non referenziali (come incrementare invece di aumentare, eloquio invece di discorso, procrastinare invece di ritardare, ecc.), i nomi deverbali (quei nomi che sono il risultato della lessicalizzazione di un verbo, come passaggio invece di passare, viaggio invece di viaggiare, dormita invece di dormire, ecc.) e i termini tecnici sono 'semplificati', disambiguati, trasposti nella loro forma base o spiegati. Un ultimo aspetto più difficile da affrontare riguarda quei termini semitecnici, che non hanno un equivalente nella LIS e che sono tradotti con una circumlocuzione. Si tratta di termini abbastanza comuni la cui identificazione risulta ostacolata dalla velocità di eloquio del TP e un'eventuale perifrasi particolarmente macchinosa.

Dal punto di vista sintattico, tra le ricorrenze maggiori è stata riscontrata la quasi assenza di schemi sintattici (in particolar modo è evitato il più possibile l'uso del passivo, delle frasi incidentate e della tematizzazione) a cui viene invece preferito il rispetto della struttura sintattica di base della LIS (Coordinate spaziali e temporali-Oggetto-Soggetto-Verbo-Altri complementi). Questa 'linearità' nella forma è stata riscontrata anche a livello frastico. Oltre all'imperante presenza della coordinazione, è stata infatti registrata anche una regolarità nella presentazione degli eventi, che nella maggior parte dei casi avviene inevitabilmente secondo un ordine cronologico, l'unico in grado di rendere i rapporti transfrastici in assenza di connettori e di subordinazione:

TP: Un'autobomba è esplosa al passaggio di un convoglio americano.

TA[147]: 'CONVOGLIO' 'AMERICANO' 'PASSA' 'AUTOBOMBA' 'ESPLODE'.

Oltre ai già citati ordine sintattico e presentazione degli eventi, un ultimo settore di applicazione di questa linearità riguarda gli aggettivi e gli avverbi di luogo, presentati in ordine di grandezza dal più grande al più piccolo, come nell'esempio che segue:

147 Per riportare la traduzione dell'interpretazione in LIS è stato deciso di adottare le virgolette e le maiuscole per indicare che la traduzione in italiano qui riportata è una traduzione parola-per-segno, in cui, cioè, ogni parola traduce il concetto espresso da un segno.

TP: al 10 di Downing Street a Londra, dove vive il Primo Ministro britannico

TA: INGHILTERRA-PRIMO-MINISTRO-VIVE-LONDRA-DOWNING-STREET-10

Similmente a quanto appena affermato per la sintassi, dal punto di vista semantico, tutti i tropi tendono a essere evitati. Da questa regola seguirebbe una lista infinita di linee guida. Vista la non necessità, in questo contesto, di produrre linee guida così dettagliate, basta considerare che questa regola è così diffusa che si applica anche a termini che sono oramai lessicalizzati (come 'un sacco/una valanga di soldi', tradotto con 'SOLDI-MOLTI' o 'la caduta di Napoleone', tradotto con 'NAPOLEONE CAPO FRANCIA NON PIÙ). Questo però non significa che la LIS o qualsiasi lingua di segni sia una lingua povera, con poche sfumature e senza umorismo verbale. Si tratta semplicemente di un indice di come le interpretazioni in LIS di un telegiornale siano neutre e puntino maggiormente alla resa dell'informazione che non a soluzioni stilistiche di pregio.

Quanto alla pragmatica, la forza illocutoria del messaggio originale emerge sempre, anche nei casi in cui prevale, nel TP, l'aspetto retorico del discorso, particolarmente grazie al ricorso che si fa, nelle lingue dei segni, dell'espressione del volto. Aggiunto alla maggiore iconicità delle lingue dei segni rispetto alle lingue orali, in questo senso più immediate nel veicolare le informazioni, quest'ultimo aspetto permette agli interpreti LIS di non perdere nemmeno un'unità concettuale. Tuttavia, non può non passare inosservata la politica dei direttori dei telegiornali accessibili ai sordi segnanti circa la velocità di eloquio dei giornalisti. Questi ultimi sembrano infatti rallentare oltremodo il loro eloquio per consentire così una corretta e piena interpretazione, che avviene in tempo reale.

In seguito a tutte queste osservazioni, si è operata una selezione delle strategie più applicabili anche alla luce dei dati ottenuti dai due test sopra riportati. In particolare, è stato dato poco peso agli aspetti più specificatamente lessicali e semantico-pragmatici (alcuni dei quali di difficile trasposizione nel rispeakeraggio, come evitare un eccesso di morfemi grammaticali, riformulare ogni volta che si incontra un lessema complesso o un nome deverbale, evitare schemi e tropi, ecc.). È stato invece dato molto risalto alla 'linearità' e alla coordinazione che, oltre a risolvere molti problemi sintattici (uso del passivo, frasi

incidentate, tematizzazioni, ecc.), sono molto utili anche nei casi di scarsa densità lessicale. Per quanto riguarda la struttura sintattica di base, tuttavia, l'oggetto prima del soggetto e del verbo risulta di difficile decodifica in italiano. È stato quindi deciso di optare per un adattamento della struttura utilizzata in LIS, variandone la posizione dell'oggetto in maniera da rispettare la sintassi dell'italiano, dando così vita alla seguente struttura:

Coordinate spaziali e temporali-Soggetto-Verbo-Oggetto-Altri complementi.

Un'ultima considerazione va fatta sui termini tecnici. Previa consultazione con esperti nel settore (interpreti e docenti di LIS), è stato deciso di non spiegarli, dando modo così agli spettatori che li conoscessero di tenere in allenamento le proprie competenze linguistiche, a color che volessero arricchire il proprio bagaglio linguistico di andare a cercarne il significato su un dizionario e a coloro che volessero ignorarli di non essere trattati come persone bisognose di ulteriore aiuto oltre alla necessaria trasposizione del TP, dalla sua forma orale al sottotitolo.

Alla luce di quanto riportato nei capitoli precedenti e in seguito ai dati appena riportati è possibile riassumere le linee guida ad uso dei rispeaker italiani nella maniera che segue:

- competenza fonetica: il rispeaker deve pronunciare le singole parole nella maniera più chiara possibile ed evitare elementi non-lessicali[148];
- competenza psico-cognitiva: il rispeaker deve simultaneamente ascoltare e comprendere il TP ed elaborare e produrre il testo di arrivo nei limiti spazio temporali dettati dalla multimedialità del genere televisivo TG;
- competenza sintetica: ai fini di una piena accessibilità del TA, il rispeaker deve operare una sintesi quantitativa del TP, rispettando il principio delle unità concettuali, e una riformulazione qualitativa che tenga in considerazione i seguenti criteri:
  - velocità media di produzione di 2,5-3 secondi per riga;
  - un sottotitolo di una riga piuttosto che uno di due righe;
  - riformulazione morfo-sintattica attuando i principi di coordinazione (da preferire alla subordinazione) e di linearità nella forma (rispettando l'ordine Coordinate

148 Cfr. Savino et al. 1999.

spaziali e temporali-Soggetto-Verbo-Oggetto) e nel contenuto (rispettando l'ordine cronologico di presentazione delle idee).

Sulla base di questi risultati è stato possibile passare alla fase sperimentale, qui di seguito riportata, che prevedeva il rispeakeraggio dei due dibattiti televisivi tra Prodi e Berlusconi per la costituzione della quindicesima legislatura, oltre che il rispeakeraggio di conferenze e di altri eventi in diretta sul territorio italofono, per poi raggiungere l'obiettivo finale, l'insegnamento del rispeakeraggio come disciplina universitaria e la sua conseguente diffusione come modalità per garantire l'accessibilità degli audiolesi a tutti gli ambiti di applicazione.

5.4 La fase sperimentale

Sulla base delle linee guida risultanti dagli studi appena riportati, è iniziata la fase sperimentale del progetto SALES. Il primo test è stato effettuato da un team di due rispeaker, l'autore del presente lavoro e un'interprete professionista[149] con un *software* di riconoscimento del parlato che proietta sottotitoli una frase alla volta in modalità *pop-on* ad ogni pausa naturale percepita nel TM. L'evento sottotitolato è stato il secondo dibattito televisivo trasmesso dalla televisione pubblica tra i due candidati alle elezioni politiche del 9 aprile 2006, l'allora Presidente del Consiglio dei Ministri uscente, cavalier Silvio Berlusconi e il rappresentante dell'allora coalizione di opposizione, professor Romano Prodi. La sottotitolazione dell'evento è avvenuta nella sede provinciale dell'Ente Nazionale Sordi di Rimini in presenza di un pubblico composto da soli sordi associati all'Ente. Da un punto di vista socio-linguistico e professionale è da sottolineare la compresenza dei rispeaker e del resto del pubblico nella stessa sala (figura 16), senza che il rispeaker potesse godere di un isolamento acustico.

149 La formazione dell'autore è avvenuta in maniera composita. Per quanto concerne l'acquisizione di competenze d'uso del software di riconoscimento del parlato, la formazione è avvenuta da autodidatta, mentre l'acquisizione di competenze professionali è stato il frutto di due stage, uno più pratico e formativo presso la BBC e l'altro, più nozionistico e informativo presso la RAI. La formazione della collega è stata effettuata in maniera del tutto sperimentale dall'autore stesso.

Figura 16: il rispeaker lavora nello stesso ambiente in cui il pubblico riceve il TA.

Per motivi logistici e organizzativi, il rispeakeraggio è dovuto avvenire senza un ingresso audio nella cuffia del rispeaker. L'*input* audio è stato preso direttamente dagli altoparlanti del televisore. Questo ha causato alcuni iniziali problemi di riconoscimento, visto che il microfono dei rispeaker riceveva in ingresso oltre al TM anche i rumori ambientali provenienti dal televisore e dal pubblico in sala. Si tratta di un elemento di disturbo notevole per il *software* che deve compiere uno sforzo di elaborazione maggiore per discriminare il TM dal rumore di sottofondo. Lo stesso vale per il rispeaker, che deve riuscire a isolarsi oltre che dall'audio del televisore sovrapposto alla propria voce, anche da eventuali reazioni del pubblico a un errore o semplicemente a un'affermazione dell'oratore. La questione della presenza degli errori è stato un argomento trattato precedentemente alla sottotitolazione. Il pubblico, che aveva già assistito a prove di rispeakeraggio in condizioni simili a quelle riportate, è stato informato della probabile presenza di errori aggravata dal mancato isolamento acustico dell'operatore. Il risultato finale (96.4% di accuratezza) ha comunque soddisfatto il pubblico e i rispeaker stessi, le cui aspettative non puntavano certamente a un simile risultato in condizioni non ideali di lavoro.

Alla fine del rispeakeraggio, durato circa 90 minuti, è stata organizzata una tavola rotonda durante la quale sono state poste al pubblico alcune domande riguardo la loro personale esperienza di spettatori di sottotitoli in diretta e soprattutto riguardo la qualità dei sottotitoli stessi. In generale, il commento è stato di apprezzamento per lo sforzo compiuto e per una resa sufficientemente valida del TP. Uno dei commenti positivi più ricorrenti è stato circa il ritmo tenuto dai sottotitoli, che per tutto il corso della trasmissione sono riusciti a mantenere un *décalage* accettabile con il TP pur lasciando il tempo ai singoli spettatori (per la maggior parte di età superiore ai 50 anni) di leggerli. Come conseguenza di questa caratteristica del TA, la maggior parte degli spettatori ha dichiarato di aver compreso nei dettagli tutte le fasi del dibattito, mentre gli spettatori più giovani hanno lamentato un atteggiamento paternalistico nei loro confronti (il TA ha seguito un ritmo di 118 parole al minuto e un tempo medio di esposizione di ogni singola riga di sottotitolo di 2,65 secondi contro delle linee guida che propongono un ritmo di 120 parole al minuto e un tempo medio di esposizione di ogni singola riga di sottotitolo di 2,5 secondi). Anche questo aspetto era stato dibattuto nel corso degli incontri precedenti con il medesimo gruppo di sordi, ma le spiegazioni e gli scambi di idee ed esperienze non sono serviti a cambiare le opinioni di alcuni.

Un altro aspetto poco apprezzato è stata, in alcuni casi, l'assenza di sottotitoli nei momenti in cui gli oratori parlavano (figura 17). Come era già stato precedentemente spiegato loro, a parte alcune omissioni intenzionali, questi vuoti sono dovuti a ragioni tecniche e professionali. In particolare, il rispeaker ha bisogno di qualche secondo per comprendere il TP e pensare a una resa accettabile, così come il *software* ha bisogno di percepire una pausa naturale prima di proiettare il testo che è riuscito a riconoscere. Questo ha comportato problemi di sincronizzazione con il TA.

Figura 17: L'oratore muove le labbra, ma nessun sottotitolo compare sullo schermo.

La questione della sincronizzazione, che come si è già ampiamente visto è una delle maggiori differenze tra la sottotitolazione per non-udenti in tempo reale e quella pre-registrata, è stata molto dibattuta, sia precedentemente l'esperimento in questione, sia durante la tavola rotonda che l'ha seguito. Nella sottotitolazione di film, in particolare, la ricerca in materia è perlopiù ferma sulla perfetta sincronizzazione tra il TP e il TA, che nella pratica si traduce con la comparsa del sottotitolo nel momento in cui l'oratore a cui si riferisce inizia a parlare e con la sua scomparsa entro un lasso di tempo sufficiente alla sua lettura da parte dello spettatore. A tal proposito, una nota va evidenziata circa i fenomeni noti come *leading* e *lagging*, vale a dire la comparsa in anticipo e la scomparsa in ritardo del sottotitolo rispetto al tempo di parola dell'oratore a cui si riferisce. Neves, pur confermando che il pubblico non udente fa molto affidamento nella perfetta sincronizzazione dei sottotitoli per poter identificare chi sta parlando, una certa flessibilità "seems allowable with leading and there are guidelines that even suggest that subtitles can come in a few frames before the actual words start being heard" (Neves, 2005: 182). Sempre in materia di sincronizzazione, un altro aspetto da considerare è che "shot changes normally reflect the beginning or end of speech" (ITC, 1999: 12). Applicati alla sottotitolazione dal vivo, la sincronizzazione e i cambi di inquadratura assumono un significato diverso. Come è evidente, la sincronizzazione perfetta e il *leading* non possono avere alcuna possibilità di esistere, salvo quei rari casi in cui un rispeaker anticipa il TP.

Quanto al ruolo dei cambi di inquadratura, c'è da sottolineare che essi svolgono un ruolo molto importante, come confermato dall'analisi riportata nel capitolo precedente. Nel caso del dibattito tra Prodi e Berlusconi, l'emittente ha stabilito personalmente dei vincoli molto precisi volti ad assicurare pari opportunità ai due candidati. Prima di tutto, la regia aveva l'ordine di mantenere la telecamera fissa sull'oratore e l'inquadratura doveva essere a mezzo busto. Gli oratori erano cinque: il presentatore del programma e moderatore del dibattito, i due candidati e due giornalisti di quotidiani nazionali. Anche i turni di parola erano prestabiliti: 30 secondi ai giornalisti per formulare la domanda e 150 a ciascuno dei due candidati. A parte l'introduzione alla trasmissione, il moderatore, incaricato di introdurre il dibattito, di dare la parola e di far rispettare i tempi, occupava dei turni di parola che variavano da un secondo a tre secondi. Questa politica dei tempi di parola lunghi e delle inquadrature fisse sul volto dell'oratore di turno hanno agevolato notevolmente il lavoro

della sottotitolatura e del pubblico sordo, che ha dovuto concentrarsi sul contenuto dei sottotitoli più che sulla componente video del TA.

Tuttavia, visto l'inevitabile ritardo con cui comparivano e scomparivano i sottotitoli, la questione dell'identificazione del parlante è stata un aspetto che ha posto qualche problema. Visto che il *software* in uso non permetteva un cambio di colore immediato, la soluzione adottata è stata l'introduzione di una didascalia all'inizio di ogni turno di parola che conteneva, nel caso dei candidati, ben noti al pubblico, il cognome (es. BERLUSCONI:), mentre per i due giornalisti è stata preferita l'etichetta GIORNALISTA:. Quanto al moderatore, è stato necessario sottolineare la sua identità (con una didascalia contenente il suo cognome, anch'esso ben noto ai telespettatori) in un'unica situazione, alla fine della trasmissione, quando ha dichiarato chiuso il dibattito e ha augurato la buonanotte ai telespettatori. Negli altri casi in cui ha parlato, i turni erano talmente brevi che era impossibile sottotitolarli. Considerati inutili sia dai rispeaker, sia dal pubblico, è stato deciso unanimemente di omettere la sottotitolazione dei suoi interventi, prima ancora dell'esperimento. Quanto all'introduzione al dibattito, il suo è stato il primo intervento e quindi non è stato necessario identificarlo. Queste soluzioni sono state molto apprezzate dal pubblico, sebbene i telespettatori non udenti italiani siano abituati all'uso dei colori nelle sottotitolazioni intra-linguistiche piuttosto che alle didascalie, che necessitano di tempo per essere identificate e quindi lette.

Per quanto riguarda la componente linguistica dei sottotitoli, i rispeaker hanno cercato di aderire alle linee guida summenzionate. Come già detto precedentemente, i sottotitoli sono risultati comprensibili dall'intero pubblico. Grazie alla registrazione dell'evento è stato anche possibile riscontrare una partecipazione attiva dei telespettatori sordi alle dichiarazioni degli oratori, sottolineata da evidenti approvazioni e disapprovazioni.

Nel rispondere alle domande di comprensione durante la tavola rotonda, nessuno ha dichiarato di non aver capito il senso generale di ogni passaggio e soltanto pochi hanno dimostrato il contrario rispondendo in maniera inappropriata a domande specifiche su un determinato passaggio. Una volta spiegato il senso reale di un determinato passaggio, la reazione di questi spettatori è stata pressoché la medesima: affermare che nel momento in cui leggevano i sottotitoli stavano capendo il passaggio in questione, ma che poi hanno dimenticato perché poco interessante o perché troppo lontano nel tempo. Quest'ultimo dato è confermato dal numero superiore di risposte sbagliate a domande riguardanti i primi

concetti rispetto a quello a domande riguardanti concetti cronologicamente più vicini al momento dell'indagine.

Da un punto di vista puramente quantitativo, il testo dei sottotitoli è stato paragonato con la sottotitolazione del medesimo evento offerta dal Centro d'Ascolto[150] sul proprio sito web in differita di mezz'ora. Questa versione era la trascrizione ortografica del TP, prodotta tramite rispeakeraggio da Cedat 85[151], adattata ai sottotitoli e sincronizzata nella mezz'ora a disposizione tra la produzione e la messa in *streaming*. Innanzitutto, è possibile notare che i sottotitoli prodotti in diretta alla sede ENS di Rimini equivalgono al 60% circa delle parole contenute nel testo integrale. Tuttavia, grazie a un'analisi basata sulle unità concettuali, è possibile notare come soltanto il 18,2% dei concetti non sia stato reso. Si tratta di un risultato straordinario, vicino a quello dei rispeaker della BBC[152], che conferma la possibilità da parte dei rispeaker di ridurre il TP senza peraltro perderne il senso generale.

Quanto al rispetto delle linee guida, oltre ai dati già riportati sul ritmo e l'accuratezza, i risultati sono ovviamente meno aderenti rispetto a quanto richiesto, ma pur sempre apprezzabili. Dal punto di vista lessicale, i termini tecnici non sono mai stati spiegati, ma sempre lasciati nel TA od omessi per motivi spazio-temporali. La sinonimia sembra essere stata applicata, ma non con il rigore dovuto, così come le riformulazioni e le semplificazioni. In fase di confronto tra i due rispeaker, è emersa la naturalezza di molte operazioni, derivante più da un automatismo acquisito durante gli studi in interpretazione di entrambi i rispeaker, sia da una cosciente volontà di semplificare il TP nel pieno rispetto delle linee guida, come nel caso di 'entrambi' reso con il meno formale 'tutti e due' o ancora di 'gli ospiti avranno <u>due</u> minuti e mezzo per rispondere' reso con il più diretto 'gli ospiti avranno <u>massimo</u> due minuti e mezzo per rispondere'. Per concludere, degna di nota è la resa con il solo cognome di tutti i nomi ed eventuali titoli di persone note menzionate nel TP, che in italiano risulta in un abbassamento del registro, ma che in lingua dei segni è invece il modo più naturale per riferirsi a una persona:

TP: A<u>l Presidente Silvio</u> Berlusconi vorrei chiedere se <u>secondo lei</u> …

150 Il Centro d'Ascolto è l'istituto di monitoraggio dell'informazione radi otelevisiva del partito radicale italiano. Cfr. http://www.centrodiascolto.it/

151 Cedat 85 è una società privata di resocontazione stenografica che collabora con la Camera dei Deputati e con il Centro d'Ascolto.

152 Va però sottolineata la differenza nella velocità di eloquio tra BBC News e il dibattito in questione, nettamente più lento.

TA: A (…) (…) Berlusconi vorrei chiedere se (…) …

Per quanto riguarda la sintassi, l'adesione alle strategie proposte dalle linee guida è stata più marcata. Il compito più difficile da portare a termine è stato il rispetto dell'ordine sintattico di base. Visto che l'italiano orale non segue rigidamente un ordine prestabilito, in sei casi su tutto il testo non è stato possibile rispettarlo. Quanto alla coordinazione, applicarla in tutto il testo non è stato un compito troppo gravoso vista la bassa complessità lessicale del TP, preparato in anticipo dai due candidati durante un allenamento durato per alcuni giorni prima della trasmissione e quindi, salvo alcuni passaggi, poco improvvisato. Un'altra regola rispettata appieno è stata la trasformazione di tutte le forme passive in attive. Maggiori problemi sono stati registrati circa il rispetto della presentazione cronologica degli eventi. I migliori risultati sono stati ottenuti nei due esempi seguenti. Nel primo, una principale e una relativa con complemento di tempo incidentato, strutturate in ordine cronologico casuale, sono state rese con un'unica principale che segue l'ordine sintattico auspicato dalle linee guida (Coordinate temporali-Soggetto-Verbo-Complemento Oggetto indiretto), con un guadagno significativo in termini quantitativi e cognitivi, ma senza perdita di unità concettuali:

TP: Ma, vorrei cominciare dalla pena di morte che è diventata, a ridosso del terribile assassinio del bimbo Tommaso, l'ultimo argomento di questa campagna elettorale

TA: GIORNALISTA: dopo la morte del piccolo Tommaso, tutti (sic.) politici parlano di pena di morte

Nel secondo è riportato un caso emblematico di costruzione sintattica in italiano parlato reso nei sottotitoli in maniera molto più lineare anche se una micro-unità concettuale è stata sacrificata:

TP: Questo confronto andrà in onda di nuovo, stasera, a mezzanotte, nella versione LIS, su Raitre, per i non udenti, a cura di Rai news 24

TA: A mezzanotte di stasera, su Raitre, ci sarà la traduzione in lingua dei segni di questo dibattito (…)

Dal punto di vista semantico, ancora due esempi sono particolarmente significativi. Nel primo si fa uso della metafora del calcio, sistematicamente eliminata in quanto ritenuta non necessaria ai fini della resa delle informazioni contenute nel TP. Anche in questo caso due micro-unità sono state omesse ('cinque giornate' e 'arbitrata da Clemente Mimun'), ma il senso generale delle due macro-unità è rimasto intatto. Inoltre, la rispeaker ha tematizzato le coordinate spazio-temporali, propendendo quindi per una resa più orientata verso il pubblico di destinazione:

TP: Ultimo confronto stasera di questa Coppa dei Campioni della politica in cinque giornate che si è disputata nelle ultime settimane su Raiuno e partita di ritorno tra Silvio Berlusconi e Romano Prodi. La partita di andata si è disputata il 14 marzo ed è stata arbitrata da Clemente Mimun

TA: Stasera, su Raiuno, ultimo dibattito politico della campagna elettorale tra Berlusconi e Prodi. Il 14 marzo si è svolto il primo incontro.

Concludendo questa breve analisi con l'aspetto pragmatico, è interessante notare alcuni aspetti che differenziano enormemente la costruzione del discorso dell'italiano orale e quella della Lingua dei Segni Italiana:

TP: Come sapete al moderatore è affidato esclusivamente il ruolo di controllare il regolare svolgimento del confronto.

TA: Io dovrò far rispettare le regole del dibattito.

In questo esempio, il moderatore definisce i ruoli di ciascun partecipante all'evento. Nel delineare i propri compiti fa uso della strategia di *hedging*[153], usando la terza persona singolare per parlare di sé stesso e del passivo per spiegare il proprio compito. Nella fase di sottotitolazione, il rispeaker si rende conto che questa strategia è puramente discorsiva e per niente informativa. Per rendere il sottotitolo maggiormente comunicativo opta per una trasformazione della terza persona in prima persona e della forma passiva in forma attiva, rendendo così più immediata l'informazione. Inoltre, il sottotitolo è comparso sotto l'inquadratura dell'oratore permettendo alle componenti semiotiche del TA di interagire proficuamente a esclusivo vantaggio degli spettatori. Dal punto di vista quantitativo, un'analisi contrastiva tra i due testi mostra chiaramente la differenza di caratteri, ma un'analisi delle micro-unità concettuali non permette alcuna speculazione circa l'omissione di informazioni. Questo è un caso lampante di riformulazione sintattica senza perdita di informazioni, resa possibile dall'omissione della summenzionata strategia di *hedging* e soprattutto dalla compressione della perifrasi 'è affidato esclusivamente il ruolo di' nel più semplice e diretto 'dovrò', oltre che da aggiustamenti strutturali derivanti da queste due riformulazioni.

Un altro esempio di buona applicazione delle summenzionate linee guida in materia di riduzione senza perdita di informazioni riguarda sempre l'illustrazione delle regole. Per spiegare che, contrariamente a quanto prestabilito, l'ultima parola spetterà al candidato Silvio Berlusconi, il moderatore inizia il discorso illustrandone la ragione:

> TP: Per un'imprecisione dovuta ai meccanismi di rodaggio della prima puntata, Berlusconi avrebbe dovuto fare l'appello finale per ultimo che invece fu assegnato a Romano Prodi.
>
> TA (…) La volta scorsa, (…) Prodi ha fatto l'appello finale per ultimo.

Anche in questo caso, la riduzione meramente quantitativa è stata di peso: 11 parole nel TA al posto di 25 nel TP. Vista la natura informativa delle tre unità concettuali contenute nel TP, il rispeaker si è concentrato sul significato delle micro-unità in esse contenute più

153 L'hedging è la presa di distanze da quanto si afferma per sottolineare, in questo caso, l'imposizione di regole e la conseguente necessaria obbedienza alle stesse da parte di chi parla.

che sulla forma. Considerata la poca informatività della prima (la ragione non viene realmente spiegata, è soltanto fatta allusione a una non meglio definita 'imprecisione') e della seconda micro-unità (ridondante in quanto esplicita quanto è facilmente inferibile dal dato di fatto espresso nella terza), solo la terza macro-unità è stata espressa.

Per quanto riguarda il resto del dibattito, le riduzioni quantitative non sono state così evidenti. I giornalisti, infatti, avevano un tempo massimo di parola, per cui i loro interventi erano preparati in anticipo. Allo stesso modo, anche i candidati avevano un tempo massimo di parola, per cui le risposte sono state ben ponderate e ogni micro- e macro-unità concettuale erano quasi tutte lessicalmente dense e grammaticalmente poco intricate. Tuttavia, applicando la tassonomia sviluppata nel capitolo precedente, è possibile sottolineare come la riduzione quantitativa del TP abbia sfiorato il 40%, tramite una significativa riformulazione sintattica per tutto il corso della trasmissione e tramite l'omissione di unità concettuali considerate ridondanti o meno informative. Un'interessante annotazione è da fare circa l'espressione 'appello finale', che in italiano non ha un referente immediato e che quindi potrebbe essere esplicitato con sinonimi più immediati o circumlocuzioni. Nelle prove al rispeakeraggio di questa trasmissione, i due rispeaker si sono preparati grazie alla registrazione del primo dibattito televisivo tra Prodi e Berlusconi, avvenuto qualche giorno prima. Anche in quell'occasione il mediatore dell'evento aveva utilizzato il termine 'appello finale' per parlare dell'ultima parte del dibattito in cui ciascuno dei due candidati aveva un certo lasso di tempo per convincere gli elettori a votare per lui. È stato quindi considerato un termine tecnico e, salvo l'esempio in questione, è sempre stato lasciato nei sottotitoli.

Un ultimo esempio interessante riguarda un passaggio in cui i tempi di parola non sono stati rispettati perché i due candidati si sono prodigati in accuse reciproche facendo ricorso a forme di umorismo verbale più o meno ironiche ed esplicite. Nel caso specifico, con una lista di dati sciorinati, Silvio Berlusconi smentisce le precedenti accuse di Romano Prodi circa l'operato del suo governo ancora in carica in materia di istruzione. La replica di Romano Prodi, nel rispetto dei tempi attribuiti, è la seguente:

> TP: A me sembra che il Presidente del Consiglio si affidi ai numeri un po' come gli ubriachi si attaccano ai lampioni, non per farsi illuminare ma per farsi... [sorreggere]

Per motive temporali e a causa del continuo cambio di inquadratura dovuto a questo scontro verbale particolarmente animato, la rispeaker ha dovuto optare per una scelta pragmaticamente efficace. Il TA, perfettamente in sintonia con le immagini, è stato il seguente:

TA: Berlusconi è ubriaco

Di primo acchito, la soluzione potrebbe sembrare molto azzardata, in quanto una citazione contestualizzata, pur utilizzata a scopo umoristico, se non sarcastico, è stata resa con un esplicito insulto. Dal punto di vista prettamente traduttivo, la soluzione qui riportata sembra essere quanto meno temeraria, in quanto viola ogni norma che regola l'intervento del sottotitolatore sul TP. Tuttavia, dal punto di vista perlocutorio, la soluzione adottata dalla rispeaker sembra aver centrato nel segno. Se non si può avere la certezza dell'intento illocutorio dell'enunciato, lo stesso non si può dire dell'effetto che le parole di Romano Prodi hanno avuto su Silvio Berlusconi. Senza lasciargli il tempo di terminare la sua citazione, Berlusconi ha ripetutamente chiesto al moderatore di 'moderare' i toni del suo avversario politico, accusandolo di mancare di rispetto al capo del Governo italiano. Quanto agli spettatori sordi, le loro reazioni sono state contrastanti, ma certamente questo passaggio ha avuto il merito di ridestare il loro interesse per il dibattito, che, a sua volta, non ha deluso le loro aspettative. Calmate momentaneamente le acque, Berlusconi si vendica dell'umorismo del candidato del centrosinistra con dell'altro umorismo basato su una citazione, o meglio, un riferimento culturale al comunismo mondiale:

TP: Ricambio l'ubriaco del signor Prodi dicendo se non si vergogna […] di svolgere […] il ruolo […] dell'utile idiota.

E ancora:

TP: In questo momento lui presta la sua faccia di curato bonario ad una parte della sinistra composta per il 70% da comunisti o da ex-comunisti

Queste due reazioni sollevano la soluzione precedentemente adottata dalla rispeaker da dubbi circa la sua adeguatezza e pertinenza. È evidente che il candidato del centrodestra ha subito l'effetto del termine 'ubriacone' più di ogni altro termine contenuto nella battuta di Romano Prodi. Una conferma proviene dal termine utilizzato da Silvio Berlusconi nella sua reazione all'insulto appena subito, 'ricambiare'. Visto che le condizioni semiotiche del TP sono rimaste immutate rispetto al passaggio precedentemente riportato (repliche molto brevi e continuo cambiamento di inquadratura), la rispeaker utilizza la stessa strategia per entrambe le repliche:

TA: Prodi è un utile idiota

E ancora:

TA: Prodi si dovrebbe vergognare.
Faccia da prete!

Anche qui, le soluzioni della rispeaker potrebbero sembrare delle deliberate e inutili infrazioni alle regole, ma, forte del successo appena incassato, la sottotitolatrice ha optato per il trasferimento sia dell'effetto perlocutorio, sia dell'intento illocutorio del TP nei sottotitoli. Da notare, inoltre, che nel primo caso, l'insulto 'idiota' è mitigato, sia nel TP, sia nel TA, dall'aggettivo 'utile', che rimanda a quei personaggi di spicco della borghesia o del clero dei paesi non (o non ancora) comunisti, che, con aria ingenua e benevola, stemperavano l'anticomunismo senza destare sospetti in quanto appartenenti a classi sociali convenzionalmente considerate conservatrici. Contrariamente al caso dell'ubriacone, quindi, la prima resa non modifica lo *status* delle parole di Silvio Berlusconi, trasformando il suo riferimento storico in un insulto, ma lo rende solo più ambiguo e più diretto. Lo stesso dicasi della seconda resa, che è un insulto che traduce un altro insulto. Pur in maniera più verbosa, e quindi apparentemente più mitigata, l'enunciato pronunciato da Silvio Berlusconi è da considerarsi un insulto. E tale è rimasto nel TA.

5.5 Conclusioni

Dopo aver brevemente analizzato l'esperimento riportato in questo capitolo, è possibile constare come il rispeakeraggio di un dibattito politico per un pubblico di sordi segnanti sia un processo particolarmente complesso che prevede l'interazione tra le attività generalmente svolte nel rispeakeraggio e il rispetto di regole specifiche volte alla piena accessibilità del TA al pubblico di destinazione. In linea generale, dei rispeaker preparati sono in grado di far fronte a un discorso politico, pur in condizioni ideali per il rispeakeraggio (velocità di eloquio particolarmente bassa, complessità semiotica ridotta al minimo, lunghi turni di parola, TP scritto per essere letto od orale ma precedentemente preparato mentalmente). Altrettanto fattibile è l'applicazione delle strategie circa la forma dei sottotitoli (un riga per sottotitolo, rispetto delle 120 parole al minuto e tempo di esposizione di 2,5 secondi).

Tuttavia, il rispetto di alcune regole sembra essere impossibile. In particolare, dal punto di vista lessicale sembrano di non immediata applicazione la sinonimia semplificatrice, la moderazione nell'uso dei morfemi grammaticali e la verbalizzazione di un nome deverbale, sebbene il rispetto dei termini tecnici e la lessicalizzazione siano, nella maggior parte dei casi, applicati. Dal punto di vista sintattico, le strategie più laboriose risultano essere l'eliminazione delle forme passive e la tematizzazione e, in misura minore, l'eliminazione delle frasi incidentate e la linearità nella presentazione degli eventi e dei luoghi, mentre la coordinazione e il rispetto della struttura base (Coordinate spaziali e temporali-Soggetto- Verbo- Oggetto-Altri complementi) sembrano non costituire motivo di difficoltà per il rispeaker. Dal punto di vista semantico, l'eliminazione dei tropi è un'attività pressoché impossibile se prescinde da una riformulazione generale della macro-unità in cui è contenuta. Quest'ultima, invece, soprattutto se volta a una resa pragmatica delle unità concettuali, risulta essere abbastanza automatica a un rispeaker con formazione da interprete di simultanea. Il trasferimento dell'intento illocutorio è sempre garantito, anche se qualche soluzione potrebbe risultare più avventata di altre.

Un altro aspetto da considerare è il *décalage*, che risulta essere sempre troppo, soprattutto a un pubblico abituato a godere dei vantaggi della sottotitolazione di programmi pre-registrati e non a leggere sottotitoli in tempo reale. Eppure, quando il ruolo delle immagini è, come in questo esperimento, ancillare e collaborativo, un divario tra il TP e il TA anche di cinque secondi può non causare problemi di comprensione. Un altro aspetto di non secondaria importanza riguarda la pretesa volontà di soddisfare le esigenze di un

pubblico, che, per quanto ridotto nelle caratteristiche peculiari (sordi segnanti pre- e peri-linguali), è comunque troppo disomogeneo per presentare esigenze simili in termini di riduzione del TP e velocità di lettura.

In seguito a questa esperienza, due domande restano ancora inevase:

- il ruolo del rispeaker deve essere quello di 'travasare' il TP nel contenitore del TA senza filtrarlo o di garantire accessibilità al pubblico di destinazione e quindi mediare?
- la sottotitolazione del discorso politico deve puntare maggiormente alla resa dell'intento illocutorio o alla resa delle strategie retoriche e di *hedging* da parte degli oratori? In alte parole, è più importante la forma o il contenuto?

Quanto a una possibile applicazione delle linee guida appena presentate in un contesto anglofono, interessante è notare come molte di esse siano paragonabile a quelle della Plain English Campaign, per la diffusione del Plain English nel mondo, in vista di una comunicazione oiù agevole. In particolare, secondo la guida on-line della Plain English Campaign[154], un testo scritto in Plain English "is a message, written with the reader in mind and with the right tone of voice, that is clear and concise" (*ibid.*: 2). Nonostante il Plain English sia volto alla redazione o alla traduzione di testi principalmente burocratici nei settori assicurativo, amministrativo, bancario, ecc., "almost anything […] can be written in plain English without being patronising or over-simple" (*ibid.*: 4). Se si passano in rassegna i punti principali del Plain English, senza peraltro considerare quelli più strettamente burocratici, è facile rendersi conto quanto questi siano in linea con i risultati ottenuti dal progetto SALES. In particolare, dal punto di vista semantico è preferibile veicolare un'idea alla volta utilizzando termini chiari e diretti piuttosto che produrre periodi concettualmente complessi. Strutturalmente, è auspicabile produrre enunciati dalla forma sintattica di base (soggetto, verbo e complementi) e contenenti un massimo di venti parole, in luogo di periodi lunghi e complessi. Quando necessario sono poi da privilegiare liste lineari. Quanto al lessico, inoltre, sono da prediligere parole brevi e di origine germanica a parole più lunghe e di origine latina; termini chiari e diretti a termini specialistici o polisemici, a meno che questi ultimi non siano strettamente necessari; verbi attivi a verbi passivi, salvo eccezioni; pronomi che indichino chiaramente il referente a circumlocuzioni ambigue; imperativi a

154 http://www.plainenglish.co.uk/howto.pdf (ultima consultazione 9 novembre 2008).

perifrasi; parole base a nomi deverbali. Grammaticalmente, è infine possibile iniziare un enunciato con ‘and’, ‘but’, ‘because’, ‘so’, ecc.; separare un infinito (es.: ‘to boldly go’); utilizzare la stessa parola due volte, invece di un sinonimo oscuro; e terminare un enunciato con una preposizione (es.: ‘that’s something we should stand up for’). Così facendo “you get your message across more often, more easily and in a friendlier way” (*ibid.*: 9).

L’uso del Plain English sembra quindi un’ottima soluzione in linea con i risultati del progetto SALES, in vista di una piena accessibilità del programma televisivo. Tuttavia, la sua applicazione e la sua validità restano ancora da testare. Vediamo ora come questi risultati si riflettono in un quadro didattico pensato appositamente per il rispeakeraggio *verbatim* e *non verbatim*.

Capitolo 6 - Verso una didattica del rispeakeraggio televisivo

6.1 Introduzione

Come si è visto nel corso di questo lavoro, il rispeakeraggio non è stato molto investigato come ambito di ricerca. Dalle analisi svolte nei capitoli precedenti, si è riusciti a stabilire la posizione del rispeakeraggio all'interno degli studi sulla traduzione, ossia alla confluenza tra gi studi sull'interpretazione e quelli sulla traduzione audiovisiva. In particolare, è stata sottolineata l'unicità della disciplina in questione, pur evidenziandone il forte apparentamento in termini di processo all'interpretazione simultanea e in termini di prodotto alla sottotitolazione per non udenti. Sono state inoltre delineate le migliori prassi del rispeakeraggio *verbatim* da una parte e del rispeakeraggio *non-verbatim* dall'altra in ottica strategica.

Grazie a questo approccio, è stato prodotto sufficiente materiale per tentare, in questa sede, un abbozzo di didattica. Visto il lento, ma crescente, fiorire di corsi universitari e post-universitari, che si prefiggono l'obiettivo di formare o di introdurre gli studenti al rispeakeraggio, la questione della didattica sembra essere rilevante oltre che utile. Tuttavia, la ricerca in materia è limitata a poche comunicazioni[155] e a pochi contributi scritti[156], spesso basati su esperienze professionali o universitarie. Traendo il massimo profitto da questi contributi in particolare e dai contributi in materia di interpretazione simultanea e sottotitolazione per non-udenti in generale, si tenterà, in questo capitolo, di costruire un quadro teorico solido e ben strutturato all'interno del quale poter inserire la didattica del rispeakeraggio.

6.2 Le prime esperienze professionali

Uno dei primi tentativi nel settore dell'insegnamento del rispeakeraggio si è svolto nel quadro del summenzionato progetto SALES. Oltre ai punti precedentemente trattati, il progetto prevedeva la formazione di rispeaker in grado di poter collaborare con uno dei partner del progetto, l'emittente pubblica sammarinese, RTV. A tal fine, è stato importante organizzare un corso di formazione professionale. Vista l'esperienza didattica e formativa del dipartimento SITLeC[157] e della SSLMIT[158] dell'università di Bologna e le competenze

155 Cfr. Baaring 2006, Eugeni 2006b e Remael e van der Veer 2006 e 2007.
156 Cfr. Baaring 2006, Arma 2007, van der Veer 2008 e Arumí Ribas e Romero Fresco 2008.
157 Cfr. http://www.disitlec.unibo.it

tecniche e professionali sviluppate nel quadro del SubtitleProject[159] dell'università di Bologna, del dottorato in Lingua Inglese per Scopi Speciali[160] dell'università di Napoli Federico II e del progetto VOICE[161] del Centro Comune di Ricerca della Commissione Europea, si è potuto mettere in piedi in un contesto accademico un corso di formazione professionale al rispeakeraggio.

È necessario premettere fin da subito che, vista la natura del corso (professionalizzante ma in contesto accademico), non sono state prese in considerazione le linee direttrici della didattica universitaria *stricto sensu*, ma si è preferito basare il corso più sugli obiettivi da raggiungere che sulle caratteristiche iniziali degli studenti, comunque prese in considerazione. Per ridurre al minimo lo scarto di conoscenze tra i candidati rispeaker, evitando così di dover pianificare il corso sulle competenze specifiche di ognuno, è stato inizialmente ristretto il campo della ricerca dei volontari a studenti di interpretazione con almeno un anno di studi di simultanea alle spalle. Così facendo, è stato anche possibile non lavorare sulle competenze psico-cognitive di base necessarie all'interazione in tempo reale con il *software* di riconoscimento del parlato. Ai formatori restava quindi il compito di sviluppare nei candidati le competenze fonetiche, di genere e sintetiche definite nei capitoli precedenti. Prima di entrare nel dettaglio del corso, è forse importante sottolineare anche il tipo di materiali a disposizione degli studenti che hanno partecipato volontariamente al corso, vale a dire alcune edizioni del telegiornale della RTV (sotto forma di trascrizione e di *file* audiovisivo), una cabina di interpretazione simultanea, un *software* di riconoscimento del parlato, un *software* di proiezione dei sottotitoli in blocco (una volta che il *software* percepisce una pausa naturale nell'eloquio del rispeaker), un computer (per visualizzare il risultato e per registrare i risultati ottenuti da ciascuno) e un mixer video (per proiettare i sottotitoli direttamente sullo schermo del computer). Infine, non essendo finanziato, il corso non aveva limiti di tempo ed è stato organizzato sotto forma di moduli individuali.

Per quanto riguarda la prima competenza (competenza fonetica) sono stati necessari degli esercizi di riscaldamento, di respirazione, di modulazione della voce e di articolazione della parola. Innanzitutto, è stato importante spiegare la necessità di tali esercizi, che sono raramente presi in debita considerazione. Gli esercizi di riscaldamento sono degli esercizi

158 Cfr. http://www.ssit.unibo.it
159 Cfr. http://subtitle.agregat.net/
160 Cfr. http://www.dipstat.unina.it/dottorato/intro.htm
161 Cfr. http://voice.jrc.it/

volti al rilassamento del corpo e all'acquisizione di una postura adeguata al lavoro da svolgere, spesso stancante in quanto prolungato nel tempo. L'obiettivo finale di questi esercizi è l'ottimizzazione del corpo in quanto cassa di risonanza della voce del rispeaker. Migliore sarà infatti la cassa di risonanza e più netta e potente sarà la voce del rispeaker. Il secondo tipo di esercizi ha riguardato la respirazione addominale, l'unica in grado di azionare, tramite il diaframma, una produzione vocalica sostenibile nel breve e lungo periodo. In seconda battuta, la respirazione addominale permette anche di rilassare ulteriormente il corpo evitandogli lo stress fisico conseguente la gestione simultanea di numerosi sforzi. Evita infine l'ipo- e l'iper-ventilazione. Gli esercizi proposti si possono suddividere in due tipologie : respirazione pura e respirazione assistita dalla parola. La prima serve a sviluppare una coscienza della respirazione addominale e, con il ripetersi dell'esercizio, un automatismo della stessa. La seconda consente di applicare i vantaggi della respirazione addominale all'eloquio. Visto che il rispeaker deve parlare a lungo e ininterrottamente, questo tipo di respirazione gli eviterà di andare in apnea dopo pochi minuti dall'inizio del suo turno. Gli esercizi appena accennati sono strettamente collegati a quelli di modulazione della voce. Questi ultimi consentono di migliorare la propria produzione vocale e di memorizzare gli esercizi di respirazione. Si suddividono in cinque fasi essenziali, ognuna volta allo sviluppo di una competenza necessaria alla creazione di un ritmo interno e al potenziamento della propria voce:

- emissione di sillabe abbinata a movimenti della lingua, della bocca o della mascella a un ritmo di respirazione imposto;
- azionamento delle varie parti del corpo necessarie all'amplificazione dei suoni emessi;
- stimolazione degli organi interessati dalla produzione del suono (lingua, bocca e mascella) tramite simulazioni esagerate di sillabe;
- alternanza di 'oh' e di 'ah' a un ritmo imposto e mutevole, produzione di scale musicali, recitazione di un ritornello senza cantarlo, vocalizzi, lettura di un testo a un ritmo imposto indipendentemente dalle strutture sintattica e frastica;
- emissione di vocali sotto forma di risate forzate a un ritmo e a una tonalità dati, imitazione del verso di alcuni animali dalle caratteristiche fisiche ben definite, come il verso del gabbiano (suono acuto), del cervo (suono sordo), della cavalletta (suono

continuo) o del tacchino (suono irregolare) e riproduzione di canti tipici, come la tirolese, a un ritmo e a una tonalità sempre diversi.

L'ultima batteria di esercizi per l'acquisizione delle competenze fonetiche sono incentrate sull'articolazione. Si potrebbe inserirli nella terza serie di esercizi di modulazione della voce, ma si rischierebbe di far perdere di vista l'obiettivo principale dell'esercizio stesso e di relegare l'articolazione a un livello inferiore rispetto agli aspetti appena esaminati (riscaldamento, respirazione e modulazione della voce), mentre è probabilmente l'aspetto principale, perché da essa dipende il buon esito del riconoscimento del parlato e perché attinge da tutti gli altri esercizi per produrre un TM che non avrà motivo di non essere riconosciuto dal *software* in uso. Gli esercizi di articolazione sono infatti strettamente legati allo stimolo degli organi destinati all'articolazione più importanti (lingua, bocca e mascelle). È inoltre indispensabile poter contare su una respirazione corretta, impossibile senza una buona gestione del proprio corpo.

Tuttavia, l'articolazione non si limita a questo. Per articolare bene, è infatti importante insistere su due aspetti: la corretta pronuncia delle singole sillabe che compongono le parole e la definizione dei confini tra una parola e l'altra. Per rafforzarli, sono stati proposti esercizi mirati all'acquisizione, da parte, degli studenti, di una certa disciplina nel pronunciare bene tutte le sillabe. Grazie alla metrica è stato possibile ottenere risultati insperati (Eugeni 2006b): essendo obbligati a seguire un dato ritmo, gli studenti che hanno partecipato al corso hanno da subito colto l'importanza della pronuncia di tutte le sillabe. Quando pronunciavano male una sillaba, si rendevano immediatamente conto della stonatura causata e si correggevano. Tuttavia, questi esercizi hanno avuto un grande limite: una volta confrontati con la realtà, gli studenti facevano fatica ad applicare le regole di buona articolazione di tutte le sillabe visto che dovevano pensare a gestire molti altri sforzi giudicati più importanti. Gli studenti hanno quindi necessitato di un monitoraggio continuo e prolungato proprio su questo aspetto. Una volta automatizzata una tale competenza, sono stati registrati buoni progressi nella pronuncia di ciascuno e nel conseguente riconoscimento da parte del *software*.

Un ultimo aspetto da sottolineare riguarda i frequenti errori di assorbimento di un monosillabo vocalico all'ultima sillaba della parola precedente o alla prima della successiva ('tutti i politici' viene riconosciuto come 'tutti (…) politici; 'Silvio ha avuto ragione' viene

riconosciuto come 'Silvio avuto ragione', ecc.). Per risolvere questi problemi, l'unica soluzione possibile è il ricorso al colpo di glottide tra i due suoni vocalici in questione. Questa strategia permette di correggere in anticipo la maggior parte degli errori imputabili al *software*. Tuttavia, insistere troppo sul colpo di glottide, potrebbe causare problemi di riconoscimento. In particolare, il *software* potrebbe interpretare il colpo di glottide come un suono consonantico ("ho visto otto persone" può essere riconosciuto come "ho visto cotto persone"). Sempre nel quadro dell'articolazione, sono stati proposti anche degli esercizi di dizione, visto che le variabili diatopiche e diastratiche in particolare di ciascuno rischiano di compromettere il buon riconoscimento da parte del *software* di una determinata parola ("esce Zidane" può essere riconosciuto come "pesce Zidane", se il verbo è pronunciato con la 'e' chiusa).

Gli esercizi per sviluppare una buona competenza fonetica sono stati proposti all'inizio del corso e sono stati riproposti all'inizio di ogni lezione, visto che una buona articolazione e una buona respirazione devono diventare automatismi, di modo che non pesino troppo sugli altri sforzi. C'è da considerare infatti che la competenza fonetica, avendo una natura trasversale alle altre tre competenze, è la prima che risente dei contraccolpi della stanchezza e dello stress del rispeaker. Dopo questo primo ciclo di esercizi, gli studenti hanno imparato a parlare alla macchina anche dal punto di vista verbale (evitare le produzioni sonore linguistiche, come false partenze e auto-riformulazioni) e non-verbale (evitare le produzioni sonore non-linguistiche, come pause piene, rumore di sottofondo, una respirazione troppo rumorosa, ecc.). Una componente importante del lavoro del rispeaker è anche saper introdurre la punteggiatura, alternandola alla produzione di testo. Una volta tutte queste competenze acquisite, il primo ciclo di esercizi si conclude con l'apprendimento degli espedienti tecnici per migliorare il riconoscimento (uso delle macro, gestione dei vocabolari, ecc.). Tutti i candidati hanno dimostrato di poter dettare al *software* qualsiasi tipo di testo scritto, senza produrre errori di riconoscimento imputabili al rispeaker.

Per quanto riguarda la seconda competenza da sviluppare, la competenza del genere da sottotitolare, il corso non ha dovuto concentrarsi troppo su questa questione, visto che l'obiettivo a breve termine era di formare dei rispeaker per la sottotitolazione in diretta del telegiornale. Il materiale utilizzato durante il corso è stato quindi esclusivamente caratterizzato dalle edizioni del telegiornale della RTV. Fin dal principio, gli studenti hanno lavorato a partire dalla trascrizione di un'edizione del telegiornale, dettandoli alla macchina.

Una volta tutte le competenze fonetiche acquisite e automatizzate, sono passati agli originali audiovisivi. Innanzitutto, gli esercizi sono stati fatti ascoltando una porzione di testo, fermando il video e ripetendo la porzione di testo memorizzata. Dopo aver dimostrato di possedere e di aver automatizzato le competenze psico-cognitive necessarie ad ascoltare e comprendere il TP parallelamente all'elaborazione e alla produzione del TM, i candidati sono passati alla fase successiva: il rispeakeraggio *verbatim*, vale a dire la ripetizione fedele del TP e l'introduzione dei segni di punteggiatura laddove necessari. Immediatamente, è stata riscontrata una difficoltà: visto che il TP era a tratti troppo rapido, i candidati si lamentavano di non poterlo ridurre. Questo esercizio era tuttavia necessario ad automatizzare le competenze fonetiche apprese, ad aumentare la propria velocità di eloquio (dettare un testo a una macchina non corrisponde al lavoro in tempo reale del rispeaker, durante il quale la sua velocità di eloquio deve essere sensibilmente più elevata) e a memorizzare il genere da sottotitolare.

A tal proposito, è stato chiesto ai candidati di cronometrare la propria velocità di eloquio e di valutare le proprie prestazioni in termini di errori per 100 parole pronunciate. Questa doppia valutazione è stata condotta su ciascun passaggio (titoli, servizi pre-registrati e in diretta, previsioni meteorologiche) e su ciascun sotto-genere (interviste, narrazione, *reportage* in diretta, domande e risposte, ecc.) del TG. Come era facile prevedere, gli aspetti che hanno posto il maggior numero di problemi sono stati, innanzitutto, la velocità di eloquio del TP, in seguito il ritmo con cui si alteravano le immagini, poi la tecnicità del linguaggio e infine la presenza di parole sconosciute o straniere. Questi fattori non sono da considerare separatamente, visto che succede spesso che la velocità di eloquio e quella di proiezione delle immagini vadano di pari passo all'interno del medesimo passaggio (soprattutto durante le previsioni meteorologiche e i servizi pre-registrati), così come la presenza del lessico tecnico e di parole difficili o straniere all'interno di uno stesso sotto-genere (servizio sull'economia, intervista a un esperto, ecc.). Una piccola questione da considerare riguarda gli aspetti culturali: i *file* audiovisivi usati durante il corso riguardavano perlopiù la repubblica di San Marino, che è uno stato indipendente, con le proprie istituzioni e le proprie regole, con le proprie celebrità e i propri toponimi. Tutte le parole che hanno a che vedere con questi aspetti non sono conosciuti per forza da un cittadino italiano, che non è abituato a considerare l'italiano come una lingua ufficiale di altre nazioni rispetto all'Italia. Questo fa sì che parole, che non esistono in italiano d'Italia, ma che suonano come italiane

(come la parola 'guaita', che significa 'guardia'), parole che non si usano più in italiano d'Italia, ma che sono ancora molto comuni a San Marino (come la parola 'arengo', che si riferisce a una forma di decisione popolare) e parole che non evocano niente in italiano d'Italia, ma che si riferiscono a qualcosa di molto conosciuto a San Marino (come il 'Consiglio Grande e Generale', che è il governo della Repubblica), non siano immediatamente comprese da un rispeaker italiano. In questi casi, i problemi di memorizzazione e di resa nel rispeakeraggio *verbatim* sono stati più che mai evidenti. Ecco quindi che è stato necessario approfondire la storia e le istituzioni di San Marino e soprattutto spiegare la terminologia relativa. Questa fase del corso, non prevista, si è rivelata essere molto importante e utile. In seguito, sono stati introdotti altri piccoli approfondimenti nei settori più ricorrenti (economia, politica, diplomazia).

Quanto alle competenze sintetiche, esse sono state sviluppate soltanto nell'ultima parte del corso, pur essendo una condizione essenziale del lavoro del rispeaker che intende sottotitolare un telegiornale per non-udenti. Gli esercizi per sviluppare queste competenze hanno costituito più della metà del corso. Tuttavia, visto che i candidati erano interpreti e quindi già abituati a fare sintesi in una lingua di arrivo per tradurre qualsiasi TP in una lingua diversa, gli esercizi di sintesi non si sono concentrati sulla capacità di recuperare una porzione troppo veloce del TP, ma sono stati progettati sulla base dei risultati del progetto SALES. Secondo questi dati, era necessario che un testo veloce come il telegiornale fosse tradotto intra-linguisticamente rispettando l'ordine sintattico di base (coordinate spaziali e temporali, soggetto, verbo, complemento oggetto e altri complementi), eliminando quindi ogni figura di stile semantica e sintattica, riducendo al minimo l'uso di morfemi grammaticali e riducendo quantitativamente il TP in maniera da lasciare al pubblico il tempo di leggere facilmente i sottotitoli (2, 5-3 secondi per riga) e di guardare le immagini. Il primo ciclo di esercizi si è concentrato sull'eliminazione delle figure di stile del TP, pur cercando di non neutralizzarle. A tal proposito, sono stati utilizzati discorsi politici altamente retorici. Visto che le elezioni dei capi di Stato di San Marino si svolgono due volte l'anno, non è stato difficile trovare testi audiovisivi retorici. Questa fase non ha posto molti problemi ai candidati, che, grazie alla loro formazione da interpreti, non hanno incontrato problemi nello spiegare una frase troppo barocca o trovare sinonimi per tropi non immediatamente comprensibili. Tuttavia, l'imposizione da parte del formatore di non condensare il TP ha talvolta comportato il non riconoscimento da parte degli studenti di

alcune figure di stile, troppo impegnati a produrre un TA il più completo possibile.

Un altro passaggio che è sembrato essere immediato è stata la serie esercizi volti alla restituzione di una sola idea alla volta. Si tratta di un'operazione necessaria alla leggibilità del TA che gli interpreti non sono per forza abituati ad attuare. Visto che il *software* utilizzato per il corso proietta i sottotitoli in blocco (ogni sottotitolo sostituisce il precedente), è stato necessario far mantenere agli studenti un certo ritmo. Per fare questo, gli studenti hanno imparato a produrre sottotitoli più o meno della stessa lunghezza, in maniera da lasciare al pubblico il tempo di leggerli per intero[162]. Il solo ostacolo incontrato durante questi esercizi è stato posto dalla semplificazione sintattica (eliminazione del maggior numero possibile di morfemi grammaticali e coordinazione delle proposizioni S-V-O). Per produrre un'idea alla volta, bisogna infatti spostare o eliminare ogni incidentata e coordinare le subordinate. Oltre a un sovraccarico cognitivo importante, questi esercizi implicano un lavoro che non è sempre considerato nella didattica dell'interpretazione simultanea, visto che gli interpreti non si preoccupano troppo della comprensione del lessico e della sintassi del TA da parte del pubblico di arrivo. Tuttavia, l'imposizione di limiti spaziali e temporali (un sottotitolo di meno di tre secondi composto da una riga che traduce un'idea alla volta) per tutto il tempo dedicato a questi esercizi ha finito col modellare il ritmo di produzione del TM da parte dei candidati rispeaker.

L'ultimo aspetto su cui lavorare è stata la condensazione quantitativa del TP. Gli esercizi che sono appena stati menzionati, le imposizioni spaziali e temporali, lo sforzo cognitivo implicato e la simultaneità del lavoro del rispeakeraggio hanno fatto concentrare l'attenzione dello studente sulla forma più che sul contenuto. Di conseguenza, se da una parte i sottotitoli erano formalmente impeccabili (di una riga, rispettanti l'ordine sintattico di base, senza figure di stile poco conosciute e restanti sullo schermo tra i 2,5 e i 3 secondi), è stata spesso riscontrata un'assenza totale di elementi di coesione tra un sottotitolo e il successivo. Il candidato, concentrandosi troppo sulla forma del sottotitolo, perdeva di vista l'ordine logico delle idee che legavano un sottotitolo al successivo. In un contesto professionale, questo rischia di causare una mancanza di comprensione da parte del pubblico interessato.

162 Per comprendere questa osservazione è necessario sottolineare che un ritmo e una lunghezza regolari permettono ai sottotitoli di restare tutti lo stesso tempo sullo schermo. In caso di dettatura di un sottotitolo sensibilmente più breve rispetto al precedente, quest'ultimo resterebbe sullo schermo il tempo di pronunciare il sottotitolo breve, quindi poco tempo rispetto al necessario.

Ecco quindi che è stato necessario introdurre esercizi per far sì che i sottotitoli prodotti non rispettassero soltanto le linee direttrici del progetto SALES, ma anche i summenzionati criteri di qualità proposti da Gambier. Per far fronte a questa esigenza, è stato necessario chiedere agli studenti di accompagnare le strategie di omissione con strategie di compensazione qualora si manifestasse un calo evidente della coesione testuale e quindi una possibile diminuzione della comprensione da parte di un eventuale pubblico. Questi esercizi non hanno portato i risultati sperati, visto che il primo effetto è stato un calo delle prestazioni fonetiche. Dovendosi concentrare sulla coesione testuale, i candidati perdevano gli automatismi fonetici e sintetici acquisiti e non facevano più attenzione alle diverse fasi e sotto-fasi del telegiornale, traducendo più testo e morfemi grammaticali rispetto alle fasi precedenti. Si è rivelato quindi necessario insistere di più su questa parte delle competenze e consacrarvi più tempo rispetto a quanto era stato previsto.

Dopo un'immersione totale in questo tipo di esercizi, i candidati sono stati testati e valutati grazie a dei criteri severi imposti dalla televisione sammarinese: rispeakeraggio di un'edizione mai vista prima del proprio TG, con un massimo di errori del 5% e adesione totale alle linee guida proposte dal progetto SALES. I risultati sono stati talmente incoraggianti da spingere la RTV a utilizzare il rispeakeraggio come forma di sottotitolazione in diretta delle proprie edizioni del telegiornale della sera[163].

6.3 Le prime esperienze didattiche

Sulla base dei loro esperimenti didattici e di ricerca, ma anche sulla base delle indicazioni fornite da van der Veer (2008), Arumí Ribas e Romero Fresco (2008) elencano tutte le competenze che un rispeaker professionista dovrebbe possedere e propongono una serie di esercizi concreti ad uso dei formatori di candidati rispeaker. Innanzitutto, gli autori fanno la duplice distinzione tra competenze preparatorie e competenze da possedere durante il lavoro in diretta. Queste categorie sono ulteriormente suddivise in competenze per il trattamento del TP, competenze per il *crossover* (vale a dire l'elaborazione del TM) e competenze per la produzione del TA. Tutte queste categorie e relative suddivisioni interne sono, a loro volta, suddivise secondo la disciplina alla quale appartengono, vale a dire a seconda che siano competenze derivabili dalla sottotitolazione per non-udenti, dall'interpretazione simultanea o che siano tipiche del rispeakeraggio.

163 Purtroppo, per motivi organizzativi, non è stato possibile tradurre in pratica queste intenzioni.

Per quanto riguarda quelle competenze che devono essere possedute prima del lavoro in diretta e che derivano dalle altre discipline menzionate, gli autori parlano di una certa abilità nell'utilizzare un *software* di sottotitolazione standard (competenze provenienti dalla sottotitolazione per non-udenti in pre-registrato), di competenze strategiche, come la capacità di lavorare in squadra, e infine di competenze preparatorie, come la capacità di sviluppare glossari e banche dati, la padronanza di una terminologia specialistica e una conoscenza del codice deontologico della professione (competenze provenienti dall'interpretazione simultanea).

Per quanto riguarda le competenze tipiche del rispeakeraggio, Arumí Ribas e Romero Fresco fanno una distinzione ulteriore tra competenze generiche e competenze tecniche. Le prime riguardano la conoscenza del funzionamento del *software* di riconoscimento del parlato, la conoscenza dell'utilizzo del *software* stesso nei limiti tecnici imposti dalla professione (senza essere frustrati dai primi risultati) e la coscienza del quadro professionale più ampio all'interno del quale il rispeakeraggio si colloca. Le seconde riguardano la capacità di controllare la propria resa e di anticipare possibili errori, la costanza nell'aggiornamento dei vocabolari e del proprio profilo vocale e infine la capacità di gestire le diverse opzioni offerte da ciascun *software*.

Quanto alle competenze da possedere durante il lavoro in diretta, gli autori le suddividono a seconda della fase cognitiva. Per quanto riguarda il trattamento del TP, le competenze necessarie (competenze generali, competenze specifiche, comprensione del TP e capacità di analizzare e di riformulare) provengono tutte da altre discipline. Le prime due sono tipiche della sottotitolazione in pre-registrato per non-udenti. Per competenze generali, gli autori intendono la capacità di riformulare e di correggere il TP, la capacità di applicare le strategie di riduzione e la capacità di individuare le varie unità concettuali nelle quali il TP si scompone. Per competenze specifiche, intendono invece la capacità di gestire i turni di parola, i nomi propri e i diversi generi audiovisivi. Le ultime due sono competenze tipiche dell'interpretazione simultanea. Nella categoria comprensione del TP, gli autori introducono la capacità di concentrazione nella fase di ascolto, la familiarità con accenti e contesti culturali diversi, la capacità di sviluppare una memoria a breve termine e la capacità di reagire di fronte a situazioni difficili. Nella categoria capacità di analizzare e di riformulare, introducono la capacità di comprendere l'intento dell'oratore, la capacità di seguire il filo del discorso, la capacità di selezionare e di concentrarsi sulle informazioni più importanti e

di distinguerle dalle informazioni secondarie, la capacità di individuare i connettori, la capacità di comprendere il senso globale grazie al contesto degli elementi extra-linguistici, la capacità di condensare l'informazione e infine la capacità di segmentare le informazioni in unità di senso.

Per quanto riguarda l'elaborazione del TM, gli autori parlano di competenze di sincronizzazione, tipiche della sottotitolazione per non-udenti in pre-registrato, e di competenze *multitask* in diretta, tipiche dell'interpretazione simultanea e del rispeakeraggio. Per quanto riguarda quelle competenze che sono tipiche dell'interpretazione simultanea, per competenze *multitask*, Arumí Ribas e Romero Fresco intendono la capacità di fare due cose alla volta, come parlare e ascoltare, comprendere e analizzare, produrre il TM e sorvegliarne l'accurata traduzione da parte della macchina, produrre il TA recuperando il ritardo con il TP e produrre il TA mantenendo un *décalage* dato. Per competenze in diretta, intendono la capacità di mantenere calma e precisione anche sotto pressione, di gestire il proprio stress, di correggere gli errori e di tenere sempre a mente il pubblico di arrivo. Per quanto riguarda le competenze tipiche del rispeakeraggio, per competenze *multitask* intendono la capacità di ascoltare e di parlare allo stesso tempo, scrivere e leggere. Per competenze in diretta, bisogna inoltre intendere la capacità di cambiare il colore e la posizione dei sottotitoli, la capacità di prevenire eventuali errori, la capacità di gestire la componente video e la capacità di lavorare senza poter ricevere *feedback* da parte del pubblico e con le alee tecnologiche tipiche della professione.

Quanto alla produzione del TA, ci sono quattro tipi di competenze, di cui due tipiche della sottotitolazione per non-udenti in pre-registrato (sforzi di produzione e conoscenza del pubblico di arrivo), uno dell'interpretazione simultanea e uno del rispeakeraggio. Gli sforzi di produzione sono gli sforzi che si compiono nella produzione di un testo grammaticalmente e ortograficamente corretto e linguisticamente comprensibile, mentre la coscienza del pubblico di arrivo è la comprensione delle difficoltà che incontrerà il pubblico di destinazione nel vedere un programma televisivo sottotitolato, la conoscenza delle capacità di lettura dello stesso, la capacità di restituire le informazioni extra-linguistiche rilevanti e la conoscenza degli aspetti riguardanti il *layout*. Il tipo di competenze che derivano dall'interpretazione simultanea riguardano in particolare la resa: capacità di esprimere le idee in maniera chiara e concisa, capacità di creare e gestire un vocabolario vasto, capacità di controllare la propria voce, capacità di comunicare facilmente e

puntualmente, capacità di restituire il tono e il registro del TP, capacità di trasmettere fiducia, capacità di parlare senza sfumature e con una buona dizione. Infine, le competenze tipiche del rispeakeraggio sono, sempre secondo gli autori, la capacità di dettare la punteggiatura al momento giusto, di produrre segmenti di testo brevi e coincisi, di mantenere lo stesso ritmo di locuzione per tutto il corso del rispeakeraggio e la capacità di parlare in maniera comprensibile al *software*, cioè con una pronuncia piatta e neutra, anche a una velocità superiore alla media.

A partire da queste considerazioni, Arumí Ribas e Romero Fresco propongono una serie di esercizi che potrebbero essere utili ai candidati rispeaker. Questi esercizi hanno l'obiettivo di sviluppare ciascuna delle competenze elencate. All'inizio, gli autori propongono una lunga serie di esercizi preparatori al rispeakeraggio vero e proprio. I candidati rispeaker devono cominciare creando un profilo vocale, liste di parole e vocabolari specifici e introducendo la pronuncia delle parole sconosciute. Quest'ultima attenzione è necessaria visto che, una volta inserito nel vocabolario, un termine deve essere sempre pronunciato alla stessa maniera, affinché il *software* lo riconosca e non lo consideri come un nuovo termine sconosciuto[164]. Dopo questa fase, i candidati rispeaker passano alla dettatura di un testo scritto. La prima difficoltà sta negli errori potenziali che il *software* potrebbe commettere. Per evitarli, il candidato dovrà anticipare ogni errore e cercare di evitarlo. Alla fine dell'esercizio, sarà necessario dettare la lista di parole sconosciute al *software*, aggiornare il vocabolario e creare le macro necessarie.

Infine, il candidato rispeaker potrà imitare il lavoro del rispeaker professionista iniziando a dettare lo stesso testo che ha preparato e individuando gli errori che non sono ancora stati corretti. Per sviluppare le competenze di ascolto e di memoria, i candidati cominceranno innanzitutto ad ascoltare un testo senza prendere appunti. Saranno quindi chiamati a rispondere a domande di comprensione e infine a riprodurre il TP. Il passaggio successivo è caratterizzato dall'introduzione di alcune difficoltà nel testo da ascoltare, come tratti dell'oralità, errori di grammatica, di sintassi o di coesione o ancora vuoti da riempire. Per lo sviluppo di competenze sintetiche, i candidati dovranno ascoltare un testo orale e individuare le idee centrali, redigere una lista di parole-chiave e di connettori e creare una mappa concettuale delle idee più importanti. La resa del testo ascoltato potrebbe avvenire sia

164 Anche questa affermazione necessita di un'ulteriore precisazione: alcuni software offrono infatti la possibilità di attribuire a una stessa parola più di una pronuncia.

cambiando l'ordine delle frasi senza cambiare tuttavia il senso generale del discorso o modificando il registro del TP, o ancora modificando totalmente la grammatica del TP (evitare le subordinate, cambiare i passivi in attivi, negare il contrario di ciò che detto, ecc.). Per migliorare le proprie competenze in sottotitolazione, i candidati rispeaker possono segmentare un testo scritto e riscriverlo sulla base delle convenzioni linguistiche e stilistiche in atto. Questo esercizio può essere ripetuto passando alla modalità rispeakeraggio, cioè ascoltando lo stesso testo e producendo sottotitoli in tempo reale. Quanto alla capacità di concentrare la propria attenzione su due azioni diverse (ascoltare e parlare allo stesso tempo), gli autori propongono un esercizio che deriva dalla didattica dell'interpretazione: ascoltare un testo mentre si conta fino a 100 o si recita una poesia. Alla fine dell'esercizio, il candidato rispeaker dovrà fare una sintesi del testo ascoltato di fronte un pubblico.

Il candidato rispeaker è ora pronto a cominciare gli esercizi di rispeakeraggio: *shadowing* affiancato dalla dettatura della punteggiatura. Per evitare che sia tentato di ripetere il TP parola per parola[165], i formatori dovranno proporre il cosiddetto esercizio della distanza, vale a dire dettare allo studente un'unità concettuale che quest'ultimo deve ripetere, lasciargli un certo lasso di tempo per farlo e, senza aspettare la fine dell'esercizio, iniziare a dettare l'unità concettuale successiva[166]. Per rafforzare la fiducia in sé stessi e per migliorare l'uso della propria memoria a breve termine, gli studenti dovrebbero svolgere degli esercizi di produzione di un testo orale a partire da un'idea proposta da un collega sia oralmente, sia per iscritto. Una volta terminata la narrazione, i colleghi commentano il testo. Un'altra serie di esercizi proposta dagli autori ha come obiettivo il perfezionamento della capacità di dettatura. Gli studenti dovranno cronometrare la propria velocità di lettura e contare il numero degli errori commessi. In seguito, dovranno ripetere questo stesso esercizio parlando più velocemente e cercando di mantenere lo stesso numero di errori ogni 100 parole dettate. Questo esercizio deve essere ripetuto finché il livello minimo del 95% di accuratezza non viene raggiunto. Un ultimo esercizio consiste nel far proporre al formatore degli esercizi, da svolgere uno dopo l'altro, più aderenti alla realtà: ascoltare un testo più di una volta e sottotitolarlo in diretta; sottotitolare in diretta un testo che è stato ascoltato una sola volta e che contiene delle parole sconosciute; sottotitolare in diretta un testo che non è mai stato ascoltato prima.

165 Cfr. Lambert 1988: 381.
166 Cfr. Van Dam 1989: 170.

Una volta che lo studente ha acquisito la competenza di ascoltare e di dettare allo stesso tempo, Arumí Ribas e Romero Fresco propongono una serie di esercizi per affinare la tecnica del rispeakeraggio: controllare lo schermo mentre si detta il TM, che permette di imparare a distribuire la propria attenzione su diverse azioni. Per quanto riguarda la scelta dei programmi da utilizzare per questi esercizi, gli autori propongono di cominciare con lo sport. In effetti, la semiotica di questo genere di programma permette al rispeaker di eliminare una buona parte del TP, in particolare la descrizione dell'azione filmata, e di concentrarsi esclusivamente sui commenti personali dei telecronisti. D'altronde, questi ultimi non producono un testo che necessita una sincronizzazione perfetta con i sottotitoli, visto che non si riferisce all'azione bensì a un aspetto secondario (la vita di un giocatore, una partita o un altro evento sportivo, ecc.). Tuttavia, nei rari casi in cui i volti dei telecronisti sono inquadrati, la sincronizzazione si impone. La difficoltà di questi esercizi potrebbe essere aumentata passando dai programmi sportivi ai discorsi politici, che impongono un ritmo di sottotitolazione superiore, ma che continuano a non richiedere al rispeaker lo sforzo di ridurre il *décalage*, in quanto non c'è un cambiamento continuo di oratore e le immagini sono per lo più fisse. Una tappa successiva potrebbe essere rappresentata dal rispeakeraggio del telegiornale, che presenta diverse difficoltà: elevata velocità di eloquio, cambiamenti improvvisi di oratore e di argomento, necessità di ridurre il *décalage* il più possibile, ecc. L'ultimo gradino è costituito dai dibattiti parlamentari, che presentano l'ulteriore difficoltà di un testo orale non sempre preparato in anticipo, una dizione non perfetta degli oratori, un uso delle tecnologie da parte degli oratori non sempre corretto, sovrapposizioni, interruzioni, ecc.

6.4 Il modello di D'Hainaut

Le due proposte di insegnamento del rispeakeraggio appena presentate (una in contesto professionale, l'altra accademico) sono uniche nel loro genere, ma presentano dei limiti abbastanza evidenti: pur essendo dei tentativi pionieristici, mancano di un modello teorico di riferimento. In entrambi i casi, l'insegnamento è proposto sulla base di competenze da acquisire. Imparare ogni singola competenza una dopo l'altra permetterebbe a qualsiasi candidato rispeaker di sottotitolare in diretta attraverso il rispeakeraggio. Inoltre, il primo tentativo è evidentemente troppo concentrato su un programma specifico (il telegiornale), che degli studenti specifici (delle persone che hanno ricevuto una formazione

da interprete simultaneo) devono essere in grado di sottotitolare per un pubblico specifico (i sordi segnanti), nel pieno rispetto di regole che sono state testate e la cui efficacia è stata provata, ma che non sono mai state messe in discussione dal punto di vista pedagogico e deontologico (è giusto semplificare un testo per un pubblico che non ha competenze che non sono considerate come normali? Questa maniera di lavorare risponde alle aspettative del pubblico?). È quindi limitato dal punto di vista didattico e non può essere riproposto in un'altra situazione, visto che si applica soltanto ai casi descritti. Il secondo modello didattico analizzato è più incentrato sul contesto accademico, ma sembra considerare le diverse competenze necessarie al rispeakeraggio come degli aspetti isolati da imparare separatamente. Una volta tutti gli esercizi completati, il rispeakeraggio di qualsiasi genere testuale dovrebbe essere alla portata di qualsiasi persona formata. Infine, entrambi i modelli non considerano una questione importante nella produzione dei sottotitoli: la nozione di accessibilità. Se il primo modello si concentra su norme fisse e molto severe che valgono soltanto per un pubblico dato, il secondo sembra non considerare le differenze tra un genere testuale e l'altro e tra un tipo di pubblico e l'altro. I primi infatti sono valutati sulla base della difficoltà nel sottotitolarli e sono visti come delle semplici tappe da percorrere una dopo l'altra invece che come generi che richiedono sforzi diversi da parte del rispeaker. I secondi infine non sono stati nemmeno menzionati.

Ciò premesso, i modelli presentati non sono da trascurare, anzi, sono una fonte preziosa di ispirazione, visto che sono stati elaborati sulla base di competenze tecniche e professionali fino a quel momento ignote. Cercando di metterle insieme e di includerle all'interno di un quadro didattico più vasto ed elaborato, si farà ricorso a un modello pedagogico per l'elaborazione di un corso universitario sul rispeakeraggio a vocazione professionale. Il modello che sembra più adatto a rispondere in maniera efficace a queste esigenze è il modello proposto da Louis D'Hainaut nel 1975[167]. Questo modello sarà innanzitutto analizzato e poi si cercherà di coglierne l'essenza e di adattarla all'obiettivo del presente lavoro. Prima di iniziare la trattazione, però, è necessario premettere che questo modello non è volto alla creazione di un corso universitario, ma è stato concepito come un progetto educativo per la creazione di un modello di *curriculum* pedagogico. Tuttavia, l'elasticità che lo contraddistingue sembra permettere un adeguamento del modello in questione alla didattica del rispeakeraggio.

167 Cfr. Safar 1992 e 2006.

Il modello di D'Hainaut è composto da tre livelli principali:

- scopi e obiettivi;
- metodi e mezzi di insegnamento e strumenti;
- metodi e mezzi di valutazione.

A loro volta questi livelli sono strutturati in 14 tappe gerarchizzate. Il primo livello, quello degli scopi e degli obiettivi, è composto da cinque tappe:

- *definizione e analisi della politica educativa*: l'obiettivo è quello di plasmare il *curriculum* in maniera da non infrangere le convenzioni della società nel suo insieme. Questo passaggio permette di constatare se la politica educativa generale è ostacolata da vincoli più o meno severi, secondo le opzioni fondamentali che orientano l'istruzione, le sue priorità e le esigenze che intende soddisfare; secondo i valori sui quali si basa e la libertà lasciata alla persona che riceve l'insegnamento nella scelta di questi valori; e infine secondo la conoscenza e la cultura considerate come acquisite dai docenti;
- *attuazione degli scopi*: questo passaggio permette di fissare gli obiettivi che ogni persona che riceverà l'insegnamento deve raggiungere alla fine del *curriculum*, a seconda di quanto emerso dalla prima tappa; a seconda dei ruoli, delle funzioni, dei compiti e dei comportamenti che dovrà svolgere e infine a seconda delle situazioni professionali e sociali all'interno delle quali sarà chiamato a operare;
- *studio della popolazione che si intende formare*: questo passaggio permette di comprendere il punto di partenza dei discenti in termini psicologici, pedagogici, culturali, psicologici e linguistici da una parte e professionali dall'altra. Così facendo sarà possibile mettere a punto gli strumenti necessari da utilizzare durante i vari corsi;
- *individuazione e analisi dei contenuti*: questa tappa permette di individuare le conoscenze necessarie per raggiungere gli scopi prefissati. Per evitare di "se livrer aveuglément à la discipline pour elle même" (Safar 1992: 87), senza considerare tutto il resto, è necessario individuare non soltanto le nozioni indispensabili al raggiungimento degli scopi prefissati, ma anche la relazione tra di essi, gli operatori che la persona che riceverà l'insegnamento dovrà poter attuare, le situazioni nelle quali sarà chiamata a operare e i problemi che dovrà risolvere;

- *elaborazione di obiettivi formativi*: questa tappa è indispensabile per capire qual è il percorso giusto da seguire per far sì che i discenti entrino in possesso dei contenuti da acquisire per raggiungere gli scopi prefissati. Solo così sarà possibile, in un secondo momento, valutare in quale misura l'obiettivo è stato raggiunto. Questo percorso dovrà essere strutturato sulla base di percorsi cognitivi essenziali che, partendo da un oggetto o uno stadio dato, permetteranno agli studenti di acquisire competenze intellettuali più complesse, necessarie a raggiungere il risultato sperato (oggetto o stadio superiore o scopo finale).

Il secondo livello del modello di D'Hainaut, metodi e mezzi di insegnamento e strumenti, è composto da sei tappe:

- *inventario delle risorse e dei limiti*: si tratta di un passaggio molto concreto, ma molto importante nella realizzazione di un *curriculum* di studi, visto che dalle risorse a disposizione dell'istituto o semplicemente dell'insegnante o del formatore e dai vincoli finanziari, materiali, logistici, amministrativi e altri dipendono la natura e il successo del *curriculum* stesso, oltre che le tappe che seguono;
- *strategia dei metodi e dei mezzi*: una volta individuati i vincoli e le risorse a disposizione, risulta indispensabile scegliere quei mezzi e di adottare quei metodi che permetteranno il raggiungimento degli obiettivi prefissati, a seconda del livello e delle caratteristiche dei discenti;
- *studio delle condizioni di inserimento*: questa tappa ruota attorno al ruolo dell'insegnante o del formatore che attuerà gli strumenti a sua disposizione, personalizzando i metodi di insegnamento a seconda delle prospettive;
- *individuazione delle situazioni di apprendimento*: strettamente legata alla tappa precedente, questa si concentra sull'ambiente nel quale il discente si troverà nel momento della formazione. L'ambiente circostante deve permettergli di poter godere delle condizioni migliori per poter approfittare nella maniera più efficace possibile degli strumenti messi a sua disposizione;
- *specificazione precisa dei mezzi*: si tratta di una fase intermedia, durante la quale sono individuati i mezzi e le competenze necessarie all'attuazione delle situazioni essenziali all'apprendimento e fissati i compiti di ciascuno;
- *realizzazione e messa a punto dei mezzi*: questa tappa costituisce "la phase

préparatoire à l'action" (Safar 1992: 88) e permette il rodaggio dei mezzi proposti per raggiungere gli obiettivi prefissati. Si scompone in quattro fasi: la concezione dei mezzi, la loro fabbricazione, la loro sperimentazione e, in caso di risultati negativi, loro messa a punto.

Il terzo e ultimo aspetto del modello di D'Hainaut è rappresentato dalla valutazione, che si caratterizza di tre tappe fondamentali:

- *elaborazione del piano di valutazione*: durante questa fase vengono fissati gli obiettivi e le variabili di valutazione, da una parte, e i criteri, i metodi e gli strumenti necessari all'attuazione di una valutazione seria e precisa, dall'altra. "Dans cette évaluation fonctionnelle, le critère essentiel est la réalisation des objectifs visés, mais celle-ci est souvent parasitée par le jugement du mérite ou l'appréciation de la performance relative" (Safar 1992 : 89);
- *selezione e realizzazione degli strumenti*: questa tappa dipende essenzialmente dalla natura dei criteri. Se ci sono strumenti che sono in grado di misurare il tasso di successo del candidato, questi devono essere selezionati o eventualmente fabbricati;
- *messa a punto dei metodi e degli strumenti di valutazione*: prima di iniziare il *curriculum* è indispensabile testare non soltanto gli strumenti di valutazione, ma anche i metodi di valutazione su un campione ristretto ma il più rappresentativo possibile della tipologia di discenti che si dovranno formare.

Nonostante risalga a 33 anni fa, questo modello non sembra essere superato, visto che non impone una politica educativa, ma cerca di mettere in atto un sistema per il quale le persone da formare sviluppano un sapere, un saper fare e un sapere essere all'altezza delle aspettative e dei contesti sociale, politico e professionale nei quali andranno a operare. Il successo di questo modello dipenderà dalla capacità di individuare gli strumenti giusti per metterlo in pratica e gli obiettivi formativi indispensabili al raggiungimento dello scopo prefissato. Grazie a una valutazione *ad hoc*, che raffronti il profilo di ciascun candidato con il profilo ideale, sarà possibile sia per il docente, sia per il discente, sia per il resto della società, valutare il grado di adeguatezza di ciascuno ai ruoli, alle funzioni e ai compiti che è chiamato a svolgere.

Per valutare l'adeguatezza di questo modello di insegnamento del rispeakeraggio

come disciplina universitaria, sarà necessario adattarlo, da una parte, alle caratteristiche del rispeakeraggio e, dall'altra, al contesto accademico.

6.5. Per una didattica del rispeakeraggio

Sulla base del modello analizzato nel paragrafo precedente e tenendo sempre a mente la nozione di accessibilità già analizzata nei capitoli precedenti, si propone una didattica del rispeakeraggio intra-linguistico *non verbatim* basata su tre assi principali:

- scopi e obiettivi;
- strumenti e insegnamento;
- valutazione.

Scopi e obiettivi

Nonostante la definizione degli obiettivi in ambito didattico non sia più la prima preoccupazione dei teorici della formazione, in un corso che ha l'obiettivo di fornire agli studenti universitari una formazione che sia il più professionale possibile, risulta necessario porre i fini e gli obiettivi in cima alla lista delle priorità. L'abilità del formatore starà nella sua capacità di individuare il più appropriato quadro di inserimento del corso. A tal proposito, è opportuno che un corso sul rispeakeraggio sia proposto da una facoltà per interpreti e traduttori. Viste le caratteristiche del rispeakeraggio e le competenze che ogni rispeaker deve possedere per poter lavorare in maniera professionale, sembra, infatti, sensato proporre il rispeakeraggio come un'attività con numerosi aspetti in comune con le due forme di traduzione più ricercate, la traduzione scritta e l'interpretazione orale. Un corso di rispeakeraggio che abbia come obiettivo la produzione di sottotitoli intra-linguistici *verbatim* potrebbe essere inserito sia nel primo anno di formazione in interpretazione, in linea con la didattica preparatoria al perfezionamento delle competenze in interpretazione simultanea, sia all'inizio del secondo anno di formazione. In quest'ultimo caso, i risultati del primo anno d'interpretazione potrebbero fornire agli studenti le basi psico-cognitive necessarie. Per quanto riguarda un corso sul rispeakeraggio intra-linguistico *non verbatim*, esso potrebbe seguire il corso sul rispeakeraggio intra-linguistico *verbatim* e costituire la fase immediatamente precedente la fase di perfezionamento della tecnica della simultanea inter-linguistica. Infine, per quanto riguarda un corso sul rispeakeraggio inter-linguistico (per udenti e per non udenti), la collocazione più logica è alla fine degli studi

d'interpretazione simultanea, visto che alle difficoltà dell'interpretazione simultanea, si aggiungono i vincoli tecnici e fonetici imposti dalla disciplina.

Quanto alla politica educativa, questa non può che essere in linea con il mercato del lavoro, visto che il rispeakeraggio è una disciplina nuova strettamente legata alla nascita della figura professionale del sottotitolatore in diretta tramite *software* di riconoscimento del parlato. Ogni docente che proponga un corso sul rispeakeraggio all'università dovrà accordare al corso una quantità di ore sufficienti alla buona formazione degli studenti. Inoltre, le competenze da fornire loro dovranno rispettare le esigenze del mercato del lavoro e degli utenti finali della sottotitolazione, pur lasciando al docente la scelta della metodologia didattica da adottare. Quanto alle convenzioni sociali, è altresì importante considerare il ruolo che il rispeaker dovrà svolgere all'interno della società. Se quest'ultima lo riconosce come una figura professionale di alto profilo, al rango di un interprete di conferenza, allora il corso di rispeakeraggio dovrà godere delle stesse ore concesse al corso di interpretazione. Nel caso contrario, il rispeakeraggio sarà considerato soltanto come uno degli sbocchi professionali dell'interprete o forse uno tra i meno prestigiosi e sarà accordato al corso un tasso di ora sensibilmente inferiore.

Per quanto riguarda l'attuazione degli obiettivi, un corso di rispeakeraggio deve non soltanto offrire agli studenti una formazione teorica e una formazione pratica generali, ma deve mettere lo studente in contesto, permettendogli così di avere a che fare con materiali originali, strumenti originali e condizioni di lavoro professionali verosimili. In breve, gli studenti non devono soltanto sapere come sottotitolare in tempo reale un programma in diretta, ma devono anche saperlo fare confrontandosi con i piccoli ostacoli di tutti i giorni, che un corso universitario potrebbe non prendere in considerazione. La capacità di far fronte a questa serie di piccoli ostacoli, che compongono una parte importante della professione, permetterà allo studente di avere un approccio professionale non soltanto teorico al rispeakeraggio, ma già spendibile sul mercato del lavoro.

Uno degli ostacoli più frustranti del rispeakeraggio è la sensazione di fallimento che nasce nello studente che inizia i propri studi di rispeakeraggio (ma anche di interpretazione simultanea e, in parte, di sottotitolazione). Oltre allo stress dovuto alla difficoltà di gestire tutti i compiti impliciti nel processo traduttivo e di non poter riprodurre tutte le unità concettuali del TP (o tutte le parole), è necessario non sottovalutare lo *shock* causato dall'impatto con il TA. Rispetto al TP, quest'ultimo sarà non soltanto asincrono e

quantitativamente inferiore, ma anche caratterizzato da errori operativi, che scoraggiano il candidato-rispeaker, visto che non avrà utilizzato correttamente tutte le proprie competenze nella maniera più corretta possibile e, se lo avrà fatto, non capirà la presenza degli errori. Quanto agli errori imputabili al *software*, la frustrazione del candidato rispeaker è maggiore, visto che la ragione di questi errori risulta essere ignota allo studente.

Una soluzione consiste nello spiegare agli studenti i meccanismi alla base di questi errori. In particolare, all'inizio del corso, è normale che il *software* non riconosca la voce del rispeaker al 100%. Con il passare del tempo, il *software* riconoscerà sempre di più l'utente purché questi metta in atto tutti gli accorgimenti fonetici per raggiungere il risultato sperato. Eppure, qualche errore continuerà a esserci. Questo aspetto è altamente stressante per quelle persone che sono abituate a produrre testi impeccabili ritenuti tali e, nei quali, ogni scarto da questa versione ottimale è da considerarsi come un fallimento. Compito del docente sarà quindi quello di far capire agli studenti quel che ci si aspetta da loro in questa prima fase e che gli errori prodotti sono la norma.

Lo studio dello studente medio è una tappa fondamentale nella formazione, visto che permette di confezionare il corso sul livello generale del gruppo da formare. Formare studenti che hanno già basi di interpretazione simultanea permette di evitare tutti gli esercizi che mirano alla capacità di ascoltare e comprendere il TP simultaneamente all'elaborazione e alla produzione del TM. Formare persone che hanno già ricevuto la formazione in sottotitolazione permette di evitare tutti quegli esercizi che si concentrano sulle norme tecniche e che regolano la forma della sottotitolo e del testo nel suo insieme. Formare dei professionisti della voce come attori, doppiatori, giornalisti e *speaker*, permette infine di evitare una buona parte degli esercizi volti alla produzione di una voce il più riconoscibile possibile dal *software*.

Le ultime due tappe, l'individuazione e l'analisi dei contenuti e l'elaborazione degli obiettivi formativi, sono stati già in parte analizzate. Una volta fissati gli obiettivi, è necessario comprendere quali sono le informazioni da dare agli studenti perché abbiano un quadro teorico di riferimento che permetta loro di capire le modalità per raggiungere gli scopi prefissati. Per quanto riguarda il quadro teorico di riferimento, i primi capitoli di questo lavoro hanno già ampiamente coperto la questione. Quello che deve essere ricordato riguarda la natura del rispeakeraggio, che lo rende sia uno strumento molto elastico, sia uno strumento che richiede molta dedizione, metodo e metodologia nell'impiego. Gli studenti

che vorranno cimentarsi in questa disciplina devono sapere esattamente quali sono i compiti del rispeaker, quali sono gli standard che deve rispettare, quali sono le strategie che deve applicare per raggiungere questi standard, in quali situazioni (in termini di genere, sotto-genere, passaggio, ecc.) e per quali ragioni (velocità di eloquio, ritardo, difficoltà grammaticale o semiotica, ecc.). Quanto agli obiettivi formativi, la formazione del rispeaker è, come abbiamo visto, basata su una serie di esercizi efficaci soltanto se gli studenti comprendono chiaramente le finalità degli esercizi stessi e il loro ruolo nel contesto generale. Considerati i primi esperimenti in materia, si propone un corso sul rispeakeraggio basato sugli obiettivi seguenti:

- sapere che cos'è un rispeaker;
- saper utilizzare un *software* di riconoscimento del parlato;
- acquisire competenze fonetiche e psico-cognitive;
- acquisire competenze di genere e sintetiche;
- saper gestire il proprio stress (in caso di errori, ritardo, difficoltà inattese, ecc.);
- saper reagire in maniera professionale alle situazioni di stress.

Sulla base di questi obiettivi, è possibile mettere in atto un corso che si componga di due grandi parti: acquisizione delle competenze di base (sapere che cos'è un rispeaker, saper utilizzare un *software* di riconoscimento del parlato standard e acquisire competenze fonetiche psico-cognitive) e acquisizione di competenze specifiche (acquisire competenze di genere e sintetiche, saper gestire lo stress e saper reagire in maniera professionale alle situazioni di stress). Durante ognuna di queste fasi, si impara non soltanto a sviluppare una competenza o una serie di competenze, ma anche a utilizzarle nel momento giusto e insieme alle altre.

Per quanto riguarda il primo obiettivo, un quadro teorico di base è necessario perché gli studenti capiscano le diverse nature dell'applicazione del rispeakeraggio, la funzione del rispeakeraggio per non-udenti, il ruolo del rispeaker e il suo rapporto con la macchina. Queste nozioni dovranno essere approfondite durante ognuna delle fasi formative perché non siano dimenticate e perché gli studenti capiscano la posizione di ognuno di questo esercizi nel contesto generale.

Dopo questa fase introduttiva, è importante che gli studenti imparino a familiarizzare con il *software* di riconoscimento del parlato. Questa fase è molto importante

nell'approccio al rispeakeraggio, visto che quest'ultimo è una disciplina molto frustrante rispetto alle scienze più 'esatte': acquisire conoscenze concrete e avere immediatamente a che fare con il *software* potrebbe essere stimolante per degli studenti che, nel corso dei propri studi, dovranno essere immediatamente confrontati alla presenza di errori. A seconda della durata del corso, sarà importante che il docente mostri agli studenti come creare il proprio profilo vocale, come introdurre la punteggiatura, come impaginare qualsiasi tipo di testo e come cominciare a migliorare il proprio tasso di riconoscimento attraverso l'aggiornamento continuo del proprio profilo, la buona gestione dei vocabolari (generale e specifici) e la creazione di macro di dettatura. Un ultimo aspetto che potrebbe essere già trattato in questa prima fase riguarda l'uso del colpo glottide, necessario per separare una parola che inizia per vocale da una precedente terminante ugualmente per vocale o, peggio ancora, con la stessa vocale. Per ottenere dei risultati immediati, gli studenti possono essere invitati a dettare al *software* un testo e a prendere nota degli errori di riconoscimento commessi, tra i quali compariranno certamente degli errori dovuti a una non corretta separazione tra parole accomunate da un medesimo suono vocalico, rispettivamente in coda e in testa di parola.

Queste note saranno la base degli esercizi successivi, volti all'acquisizione di competenze fonetiche. A seconda del genere di errore (al posto della parola sperata, il *software* riconosce un quasi-omonimo, una parola che ha soltanto una parte in comune, una serie di parole totalmente diverse; il *software* introduce delle parole che non sono state pronunciate; il *software* non riconosce le macro o la parola appena introdotta nel profilo vocale; ecc.), sarà possibile proporre delle soluzioni specifiche che possono essere raggruppate in quattro categorie: esercizi di riscaldamento del corpo della voce, esercizi di respirazione, esercizi di modulazione della voce ed esercizi di articolazione. Una volta ultimati questi esercizi, gli studenti dovranno essere invitati a continuarli a casa per automatizzare la propria competenza fonetica. In una prima fase, senza il *software* di riconoscimento del parlato e, in seguito, dettando un testo scritto alla macchina. Visto che l'obiettivo è di automatizzare la competenza fonetica, è necessario che gli studenti facciano questi esercizi per tutto il resto del corso. Da parte sua, l'insegnante dovrà preoccuparsi di riproporre all'inizio di ogni lezione alcuni esercizi fonetici in maniera tale da incoraggiare gli studenti ad aggiornare i propri profili vocali e a mantenere elevato il loro grado di automatismo delle proprie competenze fonetiche. L'ultimo esercizio fonetico consiste nel

cronometrare il proprio standard di produzione, vale a dire il numero di parole al minuto che ognuno riesce a pronunciare al *software* senza che quest'ultimo commetta degli errori di riconoscimento imputabili a una cattiva pronuncia. Questo dato costituirà il tetto massimo che ogni candidato potrà raggiungere in situazione professionale senza che il *software* commetta errori. Inoltre, nel tentativo di non superare questo standard e in condizioni di velocità di eloquio del TP particolarmente elevate (superiori alle possibilità fonetiche del rispeaker), il rispeaker dovrà attuare strategie di recupero specifiche onde evitare di commettere inutili errori di riconoscimento. C'è infine da sottolineare che questa soglia può essere aumentata con l'esercizio e con l'automatizzazione delle competenze necessarie al rispeakeraggio.

L'acquisizione delle competenze psico-cognitive è l'ultima fase della prima parte. In questo contesto, lo studente deve sviluppare un *know how* articolato che costituisce il nodo del rispeakeraggio. È per questo motivo, che sarà necessario dedicare molto tempo alla acquisizione di questo genere di competenze. Gli esercizi che permettono di sviluppare la capacità di utilizzare lo stesso canale cognitivo (acustico-vocale) per ricevere il TP e per emettere il TA[168] derivano direttamente dalla didattica dell'interpretazione simultanea. Per riprendere gli esercizi analizzati precedentemente, si propone in questo paragrafo di iniziare con esercizi di *shadowing* fonemico, senza utilizzare il *software* di riconoscimento del parlato. È importante che, in questa fase introduttiva allo *shadowing*, sia il docente, sia gli studenti siano coscienti che questi esercizi sono soltanto la prima tappa verso un approccio più articolato al rispeakeraggio. Questo esercizio mira infatti a far nascere negli studenti una competenza rudimentale e puramente funzionale. La seconda tappa dovrà prevedere l'introduzione della punteggiatura, permettendo così una divisione concettuale del TP in frasi prima e in sintagmi poi. Questi esercizi dovranno essere ripetuti finché lo *shadowing* con introduzione della punteggiatura non viene automatizzato.

Gli studenti saranno quindi pronti per lo sviluppo della propria memoria a breve termine. A tal proposito, si propongono innanzitutto degli esercizi introduttivi all'interpretazione di trattativa, come ascoltare un testo suddiviso in idee concettuali e ripetere ogni idea alla fine della stessa nella pausa che la separa dalla successiva. Una difficoltà aggiunta, ma che dovrebbe essere automatizzata, è rappresentata dall'esercizio

168 In questa fase della formazione il software di riconoscimento del parlato non è utilizzato e quindi il TM corrisponde con il TA.

della distanza, grazie al quale si passa alla creazione di una base solida per il rispeakeraggio.

Come esercizi di transizione, si propone di svolgere gli stessi esercizi appena menzionati utilizzando il *software* di riconoscimento del parlato. Gli studenti, che dovrebbero essere capaci, a questo stadio, di svolgere questi esercizi senza particolare difficoltà, dovranno confrontarsi a qualche ostacolo operativo. Innanzitutto, dovranno confrontarsi con l'attuazione simultanea di due competenze, con possibili conseguenze negative su una delle due competenze[169]. Un altro ostacolo alla buona realizzazione di questi esercizi è la co-presenza del TP e del TM nel canale acustico-vocale e del TA (con tutti i suoi errori) e le componenti video verbali e non verbali nel canale visivo. La volontà di controllare la correttezza di quanto è stato dettato, da una parte, e la presenza di errori nel TA, dall'altra, costituiscono un ostacolo importante al buon esito del rispeakeraggio. Per controllare se il *software* riconosce correttamente la propria voce, lo studente tende a rallentare la produzione del TM, talvolta addirittura a fermarsi, per concentrarsi maggiormente sulla trascrizione. In caso di errori gravi, lo studente sarà frustrato dall'insuccesso che deriva da un errore o da uno meccanismo a lui ignoto. Se gli errori si ripetono spesso, la frustrazione aumenterà, così come lo stress nella voce e conseguentemente il tasso di errori nel TA che segue, inducendo lo studente alla tentazione di abbandonare il corso. Una prima soluzione a questo problema è costituito dall'imposizione, da parte dell'insegnante, di non guardare il TA. In seguito, il ruolo del docente sarà quello di non lasciare gli studenti scoraggiarsi, mostrando loro che un livello superiore può essere raggiunto tramite un esercizio continuo e che anche nel rispeakeraggio professionale possono esservi degli errori. L'importante è fare del proprio meglio, non lasciarsi scoraggiare e non voler seguire il TP, costi quel che costi, anche a discapito della qualità della riconoscimento.

A questo stadio, gli studenti non hanno ancora sviluppato una memoria a breve termine che permetta loro di fare del rispeakeraggio professionale, ma sono comunque capaci di dettare nella stessa lingua un testo alla macchina e di conferirgli una forma di trascrizione ortografica. Per iniziare con gli esercizi di memoria, sembra sensato continuare con gli esercizi di *shadowing*. In questo contesto, si propone di iniziare con lo *shadowing* a

169 Tendenzialmente, la competenza che viene trascurata di più è quella che viene considerata maggiormente acquisita. Nel caso in questione, gli aspetti fonetici. Inconsciamente, gli studenti tendono infatti a credere che il buon esito di questi esercizi e del rispeakeraggio in generale dipenda più dalla resa quantitativa del TP che dagli aspetti fonetici. Ovviamente, questo non può essere vero, visto che una carenza negli aspetti fonetici comporta l'aumento del numero di errori nella fase riconoscimento, mettendo così a repentaglio il TA.

décalage fisso. Questo genere di esercizi impone la comprensione del TP, visto che la sola memoria ecoica non basta a ripetere il TP parola per parola con un divario di cinque parole. È quindi necessario che lo studente comprenda il TP e che se lo ricordi. Questo esercizio è ovviamente una balzo in avanti molto lungo per degli studenti che sono stati abituati, fino a quel momento, a fare degli esercizi progressivi raggiungendo la tappa seguente passo dopo passo. Ecco quindi, che dopo aver mostrato loro la difficoltà di questo compito e dopo aver ottenuto un primo riscontro da parte degli studenti, il docente dovrà valutare le competenze di questi ultimi: se riescono a svolgere il compito senza particolari difficoltà, allora potrà decidere di passare alla seconda parte del corso; se, invece, si rende conto che gli studenti fanno troppa fatica a produrre un TA accettabile, allora dovrà proporre degli esercizi per rafforzare la loro memoria a breve termine.

Una prima tappa potrebbe essere costituita dagli esercizi di segmentazione del TP in unità concettuali. In particolare, gli studenti saranno chiamati ad ascoltare un *file* audiovisivo, a fermarlo alla fine dell'unità di senso o di un gruppo conciso e omogeneo di unità di senso e a dettare alla macchina quanto memorizzato. Così facendo, gli studenti imparano a ragionare in unità concettuali e non saranno quindi più frustrati dalla consapevolezza della perdita di una determinata parola. Tuttavia, gli studenti sono ancora lontani dall'obiettivo da raggiungere, cioè a dire la memorizzazione di almeno un'unità concettuale simultaneamente all'ascolto del TP e alla produzione del TM. Per avvicinarsi a questo obiettivo, dovranno essere proposti altri esercizi di memoria, come lo sviluppo di strategie di recupero attraverso esercizi *ad hoc*. In particolare, si possono aggiungere dei fattori di disturbo nel testo dettato dall'insegnante (omissione di parole inferibili dal contesto, interruzione della trasmissione del segnale acustico, introduzione di tratti dell'oralità come false partenze, auto-riformulazione, ecc., accelerazione improvvisa della velocità di eloquio del TP, introduzione di parole foneticamente appartenenti alla lingua di lavoro, ma di fatto inesistenti, introduzione di errori grammaticali, ecc.). Questi esercizi sono importanti, visto che permettono allo studente di gestire il proprio stress. In caso di difficoltà, si spera che lo studente non sia più totalmente assorbito dall'errore, come accadeva precedentemente, ma che riuscirà a prendere abbastanza distanza dal TA e a considerarlo nel suo insieme. L'ultima tappa degli esercizi volti allo sviluppo di competenze psico-cognitive è rappresentato dallo *shadowing* sintattico. Questo esercizio dovrà essere svolto, inizialmente, senza dettare il TM al *software* di riconoscimento del parlato,

operazione che sarà effettuata soltanto dopo aver compreso il meccanismo dell'esercizio.

Una volta acquisite le competenze fonetiche e psico-cognitive e una volta sviluppate le strategie di recupero, si può passare alla seconda fase del corso di rispeakeraggio. Questa fase è più concreta rispetto alla prima, visto che gli studenti dovranno mettere in pratica quello che hanno imparato in un contesto reale e non più astratto. Prima di arrivare a questi esercizi, sembra sensato iniziare con degli esercizi di transizione durante i quali tutto ciò che stato appreso nella prima parte del corso sia messo in atto, per sviluppare un primo pacchetto di competenze operative, le competenze sintetiche. Per raggiungere questo obiettivo, si propongono degli esercizi come ascoltare un testo che deve essere compreso in generale e contare da 100 a 1 o recitare una poesia. Alla fine dell'esercizio, il docente porrà agli studenti delle domande non solo di comprensione generale, ma anche riguardanti alcuni aspetti puntuali, importanti dal punto di vista informativo. Questi esercizi permetteranno agli studenti di abbandonare, poco a poco, l'ossessione spesso riscontrata di dover dire tutto, come se ogni parola non ripetuta costituisca una perdita troppo importante perché lo studente possa dirsi tranquillo con la propria coscienza. Altri esercizi per raggiungere questo obiettivo sono: riassumere in due righe un testo di quattro, poi di sei; generalizzare la descrizione approfondita di una nozione o di un evento; eliminare un inciso dalla resa di unità concettuali sinteticamente complesse e via dicendo. Rispetto agli esercizi precedenti, questo tipo di esercizi permette allo studente di reagire in maniera professionale a una situazione reale, come un TP verbalmente troppo rapido o troppo complesso in seguito alla presenza di ostacoli puntuali al processo traduttivo.

Dal punto di vista sintetico, è necessario iniziare a far familiarizzare gli studenti con le due modalità di rispeakeraggio analizzate nei capitoli precedenti. Ecco quindi che si può procedere passo dopo passo all'espletamento di esercizi volti al soddisfacimento delle summenzionate linee guida. In particolare, per quanto concerne il rispeakeraggio *verbatim*, si possono inizialmente proporre esercizi volti alla ripetizione fedele di un testo inserendovi la punteggiatura al posto giusto e abolendone tutti i tratti dell'oralità. In un passo successivo, si possono poi prevedere esercizi maggiormente impegnativi, diretti alla ripetizione programmata di un testo con la punteggiatura, senza i tratti dell'oralità e cercando di utilizzare una velocità di eloquio di massimo di 180 parole al minuto. Compresa la difficoltà del rispetto del limite massimo imposto, successivamente si potranno suggerire l'abolizione

delle macro- e delle micro-unità concettuali più ridondanti e in ultima analisi il raggiungimento dell'obiettivo delle 180 parole al minuto tramite l'introduzione di esercizi di compressione sintattica.

Quanto al rispeakeraggio *non verbatim* si possono proporre esercizi di riscrittura del TP in base alle indicazioni della Plain English Campaign e cioè

- flessiblità grammaticale: iniziare un enunciato con 'and', 'but', 'because', 'so', ecc.; separare un infinito; utilizzare la stessa parola due volte, invece di un sinonimo oscuro; terminare un enunciato con una preposizione;
- chiarezza lessicale: trasformare parole lunghe e di origine latina in parole brevi e di origine germanica; termini specialistici o polisemici in termini chiari e diretti; verbi passivi in attivi; circumlocuzioni ambigue in pronomi che indichino chiaramente il referente; perifrasi in imperativi; nomi deverbali in parole base;
- linearità sintattica: tradurre periodi lunghi e strutturalmente complessi in frasi dalla forma base (soggetto, verbo e complementi) composte da un massimo di venti parole ciascuna; e occasionalmente in liste lineari;
- semplicità semantica: tradurre periodi concettualmente complessi in enunciati contenenti un'idea ciascuno espressa con termini chiari e diretti.

Gli esercizi finora descritti servono solo a prendere dimestichezza con le varie strategie di resa testuale, ma non sono funzionali a un'applicazione concreta e contestualizzata. Per mettere in atto tutte le competenze acquisite e per sapere quale strategia utilizzare e in quale contesto, è necessario infatti che ogni rispeaker conosca il genere testuale che dovrà sottotitolare in diretta e le esigenze delle varie tipologie possibili di pubblico di arrivo. Per quanto riguarda il primo tipo di conoscenze, il docente deve presentare agli studenti diversi tipi di testo, come un telegiornale, competizioni sportive, sessioni parlamentari, interviste, cerimonie, ecc. Insieme all'insegnante, gli studenti dovranno concentrarsi sull'analisi del genere testuale proposto e soprattutto sulla velocità di eloquio di ogni passaggio, sulla velocità in cui le immagini si susseguono e sul ruolo della componente video all'interno del sistema semantico e semiotico generali. Grazie a queste informazioni, lo studente potrà iniziare a vestire i panni del rispeaker. In particolare, potrà decidere se ripetere il TP parola per parola o se la velocità dello stesso sia troppo elevata da consentirlo, se sottotitolare una porzione di testo che si riferisce a inquadrature troppo rapide

(come durante una partita di calcio, una cerimonia, ecc.), dei turni di parola troppo rapidi (come le domande e risposte in un quiz o un dibattito improvvisato, la descrizione di un'azione ben comprensibile dalle immagini (come durante una partita di calcio), se ridurre il *décalage*, se cambiare il registro del TP, ecc. Grazie a una analisi dettagliata di più testi di uno stesso genere, sarà anche possibile individuare un registro ricorrente o una terminologia o ancora delle strutture ricorrenti, che possono anche essere anticipate. Successivamente, gli studenti saranno formati alla preparazione alla diretta, creando liste di parole specifiche di un dato testo, allenandosi al trattamento di alcuni passaggi (eliminare tutti i tratti dell'oralità, riassumere un passaggio troppo rapido od omettere un passaggio ridondante, ecc.), creando macro, ecc.

Per quanto riguarda le esigenze del pubblico di destinazione, gli studenti devono sapere per chi stanno sottotitolando. Si tratta di un passaggio teorico che poteva anche essere inserito all'inizio del corso. Tuttavia, sapere che cos'è la sordità e che all'interno dell'utenza finale ci sono gruppi di persone che hanno necessità diverse (sordi bilingui, sordi oralisti, sordi segnanti, audiolesi di diverso grado, anziani, bambini sordi, stranieri, sordi colti, sordi analfabeti, sordi pre-linguali, sordi post-linguali, ecc.) induce gli studenti a ragionare in maniera concreta sulle possibili ripercussioni sul prodotto finale. Inoltre, questa introduzione teorica permette loro di comprendere che non è necessario tradurre ciò che è anche comprensibile dalle sole immagini, ma che anche componenti apparentemente poco rilevanti, come un rumore di sottofondo, possono essere da tradurre perché portatrici di significato. In seguito, questa introduzione teorica apre gli occhi sulla questione delle necessità da soddisfare, che sono spesso diverse rispetto alle rivendicazioni delle associazioni in difesa degli audiolesi. Infine, questa fase è importante per sapere che un contatto continuo con il pubblico di destinazione permette di avere *feedback* indispensabile al miglioramento di aspetti puntuali riguardanti sia gli aspetti tecnici, sia quelli linguistici.

L'ultima tappa di questa seconda parte del corso dovrà concentrarsi sull'applicazione di quello che è stato fin qui descritto. Gli studenti saranno confrontati al rispeakeraggio in tempo reale e a diversi tipi di generi testuali. Ogni genere ha infatti specificità diverse che richiedono reazioni diverse da parte del rispeaker. È per questo motivo che gli studenti dovranno essere confrontati a tutti i generi che potranno essere incontrati nel corso della loro professione. Per iniziare, gli studenti potrebbero ascoltare un testo, prendere nota delle caratteristiche specifiche e sottotitolarlo più volte finché non si

raggiunge l'*optimum* fissato dal docente. Con il passare del tempo, dovranno diminuire progressivamente il numero di tentativi prima di raggiungere lo stesso *optimum*. Infine, dovranno abituarsi a raggiungerlo al primo colpo, come avverrà nella prova finale, oltre che nella realtà.

Il docente, che fino a quest'ultimo stadio svolgeva un ruolo importante, in quest'ultima parte si limiterà a proporre i testi, a dare indicazioni generali e a correggere qualche errore. Visto che il corso è quasi giunto alla fine, il docente dovrà anche valutare il livello generale raggiunto dalla classe e assicurarsi che l'esame potrà essere superato dalla maggior parte degli studenti. In caso contrario, dovranno essere proposti esercizi compensatori perché lo scarto con il livello desiderato sia colmato. Abbassare il livello richiesto all'esame potrebbe essere un'altra soluzione, ma questo significherebbe che il corso non è stato efficace o ben strutturato. In ogni caso, il docente dovrà dare indicazioni specifiche a ogni studente perché tutti abbiano coscienza dei propri limiti oltre che la possibilità di superare l'esame al primo tentativo.

Strumenti e insegnamento

Si tratta di un aspetto molto importante, visto che la maggior parte dei dubbi in materia di rispeakeraggio e di promozione di corsi di rispeakeraggio dipendono dalla mancanza di strumenti adeguati (cfr. Arumí Ribas e Romero Fresco 2008). Il rispeakeraggio è infatti efficace soltanto se gli strumenti a disposizione del docente e degli studenti sono qualitativamente e quantitativamente adeguati. Ecco quindi che ancor prima dell'inizio del corso, è necessario verificare la disponibilità di risorse:

- *materiali*: per far lavorare bene ogni studente, deve esserci almeno una cuffia per ogni studente e un microfono, una tastiera e un computer ogni due studenti. In caso contrario, è indispensabile variare la tipologia degli esercizi, proponendo alle persone che non stanno lavorando al computer degli esercizi alternativi, che rischiano però di rallentare l'evoluzione dell'insegnamento;
- *informatiche*: ogni studente deve poter accedere al proprio profilo vocale e arricchirlo ogni volta che fa un esercizio con il *software* di riconoscimento del parlato. Inoltre, deve poter lavorare con una strumentazione semi-professionale, come un *software* di sottotitolazione che offra varie opzioni, un'interfaccia che permetta l'auto-/etero-correzione e dei filmati verosimili (per evitare di creare false

illusioni negli studenti, per spingerli a fare meglio e per abituarli a convivere con la frustrazione dell'imperfezione);

- *logistiche*: il corso deve svolgersi in un ambiente con infrastrutture adeguate, come una cabina insonorizzata per ogni computer e una *console* grazie alla quale il docente potrà controllare il lavoro degli studenti senza doversi spostare di cabina in cabina, disturbando e stressando ulteriormente gli studenti;
- *temporali*: a seconda del *curriculum* e degli obiettivi del corso, è necessario che siano dedicate all'insegnamento del rispeakeraggio almeno 30 ore, benché un corso completo e degno del mondo del lavoro necessiti di almeno 50 ore e benché la maggior parte dei corsi professionali all'interno delle aziende durino 100 ore;
- *finanziarie*: in caso di mancanza di una delle risorse precedentemente analizzate, dovrebbe essere possibile risolvere immediatamente il problema, per evitare di svalutare il corso. Tuttavia, non è impossibile proporre un corso di rispeakeraggio con delle risorse più limitate.

Una volta individuate le risorse e i vincoli e sulla base delle indicazioni provenienti dalla prima fase, i metodi e i mezzi da attuare per permettere l'organizzazione del corso dipenderanno dalle caratteristiche degli studenti, dal ruolo che è stato previsto per l'insegnante e dalle caratteristiche di quest'ultimo. L'insegnante deve possedere competenze professionali per poter penetrare le molteplici sfaccettature della disciplina e per poterle illustrarle agli studenti. Deve inoltre avere un'ottima padronanza degli aspetti fonetici e psico-cognitivi che sono alla base della professione, oltre che degli aspetti sintetici e di genere tipici di ogni singola situazione (sottotitolazione del telegiornale, di una partita di calcio, di una cerimonia religiosa, di un'intervista; con o senza correzione prima della messa in onda; con o senza collega; ecc.). Il docente deve altresì conoscere bene gli aspetti tecnici del *software* con il quale forma gli studenti, per permettere loro di approfittare al massimo dei supporti tecnologici che ogni *software* di riconoscimento del parlato offre (profilo vocale, macro e vocabolari) e delle opzioni specifiche del *software* con l'obiettivo di migliorare il processo di riconoscimento del parlato. Infine, deve poter reagire immediatamente a ogni esigenza puntuale da parte degli studenti.

Quanto all'ultima fase, la realizzazione dei mezzi, è necessario che il docente non si concentri troppo su un unico genere testuale, ma che preveda molteplici tipologie testuali

(telegiornali, competizioni sportive, sessioni parlamentari, cerimonie ufficiali, ecc.). È necessario che i filmati siano sotto forma digitale di modo che gli studenti possano lavorare in maniera professionale: a partire da uno stesso computer, uno studente deve poter ricevere la componente audio attraverso una cuffia e la componente video attraverso lo schermo, produrre il TM grazie a un microfono adeguato e grazie a una tastiera correggere, cambiare il colore del TM o spostarlo prima di mandarlo in onda. Dopo questa simulazione, gli studenti dovranno poter lavorare di nuovo sul testo prodotto, correggendo gli errori commessi sottotitolando nuovamente lo stesso testo. Inoltre, gli studenti devono poter esercitarsi a distanza e condividere *file*, vocabolari, macro, ecc. Infine, il docente deve poter correggere gli esercizi e controllare l'evoluzione di ciascuno, anche da una postazione remota.

Valutazione

La valutazione degli studenti di un corso di rispeakeraggio deve farsi *in itinere* visto che un'unica valutazione finale non permette la valutazione effettiva dei progressi compiuti e non offre la possibilità al docente di correggere la rotta se il corso non sta dando i risultati attesi. Durante l'elaborazione del piano di valutazione sarà quindi indispensabile fissare criteri funzionali, come:

- saper usare un *software* di riconoscimento del parlato standard;
- gestire ognuna delle quattro competenze richieste;
- gestire il proprio stress (in caso di errori, ritardi, difficoltà impreviste, ecc.);
- saper reagire in maniera professionale a ostacoli imprevisti.

Ognuna di queste competenze sarà oggetto di verifiche *in itinere* in maniera tale che non sarà necessario aspettare la fine del corso per rendersi conto che un passaggio dello stesso non ha prodotto i risultati attesi dall'insegnante, dall'istituzione e dagli studenti. Alla fine del corso, una volta considerate come acquisite tutte le competenze, sarà necessaria una valutazione generale delle competenze di ciascun candidato, senza peraltro tenere conto dei risultati ottenuti durante le valutazioni parziali. La ragione di questa scelta sta nell'esigenza funzionale di aver raggiunto o meno il livello professionale richiesto. Pertanto, lo studente dovrà conoscere l'argomento della prova in anticipo, in maniera da prepararsi dal punto di vista linguistico e tecnologico e da aggiornare il proprio profilo vocale. Il giorno dell'esame,

dovrà poter utilizzare l'ultimo aggiornamento del proprio profilo vocale. Una volta messo nelle stesse condizioni del corso, sarà pronto per il rispeakeraggio di un testo audiovisivo di cui conosce il genere e che quindi riesce a gestire dal punto di vista sia del genere, sia linguistico[170], oltre che psicologico.

Rimane in sospeso la questione dello stress causato da una situazione, quella dell'esame, diversa da una situazione professionale: nel mondo del lavoro, il rispeaker non sarà chiamato un giorno a dimostrare, senza esperienza, le competenze che ha acquisito di fronte a un professore universitario pronto a rilevare ogni suo errore. A tal proposito, l'incombenza maggiore spetta al docente che deve cercare di trovare il giusto mezzo nella valutazione e il miglior approccio psicologico possibile nei confronti degli studenti sia durante il corso, sia all'esame.

Per poter mettere in atto la valutazione così come è stata pianificata, lo studente dovrà essere messo nelle stesse condizioni del corso (cabina insonorizzata, insegnante che ascolta la prova da un'altra postazione, dotato dei *software* necessari, genere testuale da sottotitolare conosciuto, argomento del testo conosciuto, ecc.). Dal canto suo, il docente dovrà poter avere accesso alla prova sotto forma digitale per poterne valutare la qualità secondo i criteri summenzionati in maniera approfondita e non soltanto sulla base di un'impressione a caldo. Lo stesso testo dovrebbe essere utilizzato per tutti i candidati, in maniera da evitare inutili differenze.

L'ultima fase di questo passaggio implica la messa a punto dei metodi e degli strumenti di valutazione. Si tratta di un passaggio fondamentale, visto che non è sempre facile stabilire il livello di difficoltà che gli studenti possono affrontare. Anche se si è ripetuto più volte che lo studente deve poter lavorare una volta terminato il corso, è altrettanto vero che l'esperienza è una variabile importante nel rispeakeraggio e che alcuni testi sono più complessi da tradurre rispetto ad altri. L'esaminatore dovrà quindi optare per un testo fattibile per il livello professionale e psicologico del candidato[171]. Dovrà fare lui stesso la prova dell'esame e valutarne la difficoltà intrinseca. La valutazione del tasso di

170 Durante il corso di formazione al rispeakeraggio (ma anche all'interpretazione e alla traduzione), accade spesso che uno studente incontri difficoltà nell'espletamento del suo compito unicamente perché non ha una padronanza assoluta del contenuto trattato dal TP. In questi casi, il TA è qualitativamente inferiore alla media. La ragione è semplice: lo sforzo di comprensione grava maggiormente sul totale degli sforzi da gestire simultaneamente. Qualora venga dedicata maggiore attenzione a questo sforzo rispetto agli altri, questi ultimi potranno contare su meno attenzione da parte del traduttore, che otterrà risultati globalmente inferiori (cfr. Gile 1995).

171 Come nel caso delle conoscenze di genere, più il testo dell'esame è complesso, più difficile sarà per il candidato svolgere il compito richiesto, anche se l'argomento è ben noto e le competenze sono ben acquisite.

difficoltà di un testo prima di proporlo all'esame non è sempre presa in considerazione e spesso gli studenti si lamentano di una difficoltà diversa rispetto alla media del corso. Succede, infatti, che senza fare la prova precedentemente, l'insegnante consideri semplice e adeguato un testo che invece si rivela essere un grave ostacolo al buon esito dell'esame. Testare in anticipo la difficoltà di un testo potrebbe quindi risolvere sul nascere eventuali problemi in sede di esame. Si tratta quindi di una tappa importante, visto che soltanto alla fine di questa valutazione, il docente può essere certo che i metodi e gli strumenti messi a punto durante il corso permetteranno una valutazione corretta e obiettiva.

Il corso è pronto per partire. Tutti i professionisti della formazione dovranno verificare l'attuazione di tutti i punti stabiliti. In caso di problemi nell'attuazione del modello appena elaborato, ognuno dovrà prendere la decisione giusta per risolvere anticipatamente gli eventuali problemi e per far sì che gli studenti abbiano ricevuto una formazione professionale una volta terminato il corso. L'istituzione dovrà inoltre interessarsi alla carriera degli studenti una volta laureati, visto che anche la propria reputazione al di fuori dei confini accademici influenza l'introduzione degli ex-studenti nel mercato del lavoro.

Un aspetto che non è stato preso in debita considerazione fino a questo punto, ma che merita tuttavia di essere menzionato è la questione dello *stage*. Una volta terminato il corso, gli studenti che lo vorranno dovranno poter approfittare dei contatti che l'università ha con il mondo del lavoro, in maniera da far conoscere loro l'ambiente professionale e a quest'ultimo le possibilità offerte in termini di ricerca e di formazione del personale dalla collaborazione con l'università. Succede, infatti, che anche gli studenti di facoltà prestigiose si lamentano di essere abbandonati a sé stessi una volta terminato il ciclo di studi. Potersi orientare nel mondo del lavoro con contatti durante gli studi è quindi essenziale perché la formazione offerta allo studente non vada perduta. Inoltre, gli *stage* permettono agli studenti di mettere in gioco la propria formazione, di comprendere e approfondire alcuni aspetti e di scoprire il mondo del lavoro. Infine, gli *stage* di formazione presso le aziende rendono l'università un vero trampolino di lancio per il futuro dei propri studenti, un elemento questo di grande prestigio.

6.6 Conclusioni

Con lo sviluppo di una didattica *ad hoc* si conclude un percorso che è iniziato con la descrizione di una tecnica poco nota, ma che grazie all'evoluzione tecnologica permea sempre di più la vita di tutti e in particolare di quelle persone che, per motivi sensoriali, non hanno la possibilità di avere accesso a uno dei mezzi di comunicazione più diffuso al mondo, la televisione. Questa tecnica è altresì importante per le emittenti televisive, in quanto offre loro un mezzo di comunicazione e di inclusione sociale flessibile e veloce, in grado di soddisfare le imposizioni legislative e di dare contemporaneamente visibilità alla propria programmazione. Infine, il rispeakeraggio è un'ottima possibilità di impiego per i futuri laureati in interpretazione, che avranno così modo di applicare le proprie conoscenze in uno dei settori oggi in maggiore sviluppo, l'accessibilità. È per tutte queste ragioni che il rispeakeraggio televisivo necessitava di un'attenzione maggiore, di un'analisi approfondita e possibilmente completa. Nel tentativo di raggiungere questo ambizioso obiettivo, si è iniziato un lavoro di descrizione del rispeakeraggio, che ha dimostrato la sua unicità come tecnica, tanto da ipotizzarne una disciplinarizzazione.

In quest'ottica si sono operate due analisi comparative, che hanno portato all'esatto posizionamento del rispeakeraggio all'interno degli Studi sulla Traduzione, cioè a dire a cavallo tra gli studi sull'interpretazione (per il suo apparentamento, come processo, alla simultanea) e gli studi sulla traduzione audiovisiva (per lo *skopos* a cui mira il rispeakeraggio come prodotto). Grazie a questi primi risultati, si è cercato di derivare un quadro teorico il più esaustivo possibile dagli insegnamenti degli studi sull'interpretazione e dagli studi sulla TAV. Questo ha permesso la realizzazione di un modello di analisi strategica del rispeakeraggio come prodotto audiovisivo che poggia su tre pilastri: l'analisi di genere del testo (o dei testi) in esame, per comprenderne la caratteristiche ricorrenti e la funzione; l'analisi multimodale delle fasi e sotto-fasi del testo (o dei testi) in questione, per comprendere la composizione semiotica del messaggio da esso (o da essi) veicolato; una tassonomia delle strategie adottate dai rispeaker nel passaggio dal TP al TA, appositamente ideata per gli scopi del presente lavoro. Qusta tassonomia ha permesso di analizzare le strategie adottate dai rispeaker della BBC (*leader* nell'applicazione di questa tecnica ai fini della sottotitolazione in tempo reale dei programmi in diretta) nel sottotitolare in tempo reale *BBC News*. I risultati di questa analisi hanno permesso di stilare una lista di linee guida ad uso dei rispeaker in lingua inglese. Da questi insegnamenti sono scaturite due applicazioni: una nel mondo professionale italiano, l'altra nel mondo accademico. Per quanto riguarda la

prima, è stato tentato un uso del rispeakeraggio in Italia, fino a quel momento inesistente. Questo ha implicato numerose tappe che hanno portato alla redazione di buone prassi per il rispeakeraggio intralinguistico per sordi segnanti del TG e all'applicazione di queste buone prassi che hanno portato a un esperimento rimasto ancora unico nel suo genere: la sottotitolazione in tempo reale tramite rispeakeraggio di un dibattito politico in televisione tra candidati Premier alle elezioni legislative.

Quanto alla seconda applicazione, l'obiettivo era di completare il percorso iniziato con la descrizione del rispeakeraggio, vale a dire la costruzione di un modello di insegnamento del rispeakeraggio come disciplina universitaria. Inizialmente, si è tratto spunto dai pochi contributi scritti e orali in materia per poter comprendere gli aspetti più immediati e concreti della didattica in materia (competenze da acquisire ed esercizi *ad hoc*). In seguito, grazie al contributo del modello didattico di D'Hainaut, appositamente adattato ai fini dell'insegnamento delle discipline traduttive audiovisive da Safar, sono state sistematizzate queste conoscenze in un quadro immediatamente applicabile dalle facoltà che vorranno introdurre l'insegnamento del rispeakeraggio nel proprio programma di studi. Si tratta di un modello coerente in grado di affrontare tutte le esigenze formative di studenti universitari desiderosi di acquisire competenze professionali subito spendibili sul mercato, di guidare il docente durante tutto il corso e di colmare almeno parzialmente un divario ancora molto profondo all'interno delle società moderne, quello tra mondo del lavoro e università.

Conclusioni

Il rispeakeraggio è la ripetizione, la riformulazione o la traduzione in tempo reale della componente verbale di un testo multimodale prodotto oralmente. Essendo esso stesso prodotto oralmente, necessita di un software di riconoscimento del parlato per trascrivere sotto forma di grafemi la voce (meglio, il parlato) del sottotitolatore. Applicato alla televisione è una innovativa tecnica di sottotitolazione in tempo reale volta a garantire l'accessibilità dei non-udenti e dei parlanti una lingua diversa da quella trasmessa a un prodotto televisivo in diretta. Si distingue dall'interpretazione simultanea perché la finalità è differente. Di contro, si differenzia dalla sottotitolazione per non-udenti di programmi pre-registrati per l'immediatezza e la simultaneità con cui avviene il processo di sottotitolazione. Per quanto riguarda il canale, infine, a trasmettere il lavoro del sottotitolatore (in questo caso il rispeaker) ai destinatari è il software di riconoscimento del parlato e non la tastiera (come avviene per le altre forme concorrenti: stenotipia, velotipia e dattilografia). Vista la sua natura complessa (a metà strada tra l'interpretazione simultanea e la sottotitolazione per non-udenti in pre-registrato), nel contesto traduttologico, il rispeakeraggio si pone al crocevia tra due discipline affini, l'interpretazione e la traduzione audiovisiva. Attingere dagli studi di questi due già sufficientemente vasti e perlustrati ambiti di studio è sembrato opportuno, oltre che maggiormente economico in termini di tempo, ai due principali fini che questo lavoro si era prefissato: teorizzare il rispeakeraggio sia come processo, sia come prodotto in ottica strategica e applicare i risultati di tali modelli teorici allo sviluppo della professione e alla creazione di una didattica universitaria *ad hoc*.

Prima di muovere qualsiasi passo è stato però necessario appurare la validità delle ipotesi di ricerca circa l'apparentamento tra il rispeakeraggio e le discipline in questione. Per giungere a un tale risultato è stato prima di tutto necessario un approccio descrittivo (primo capitolo) al rispeakeraggio in grado di cogliere le caratteristiche principali della disciplina. Da quest'analisi è emersa una duplice natura del rispeakeraggio che lo contraddistingue sia come processo, sia come prodotto. Per quanto riguarda il processo, la complessità della professione è tale da renderlo simile all'interpretazione simultanea. Le competenze che deve applicare un rispeaker nella fase operativa sono infatti del tutto simili a quelle studiate dagli *interpreting studies* e in particolare:

- fonetiche;
- psico-cognitive;
  - diamesiche:
  - metalinguistiche;
- sintetiche;
- di genere.

Quanto al prodotto, a seconda della politica dell'emittente e di numerosi altri fattori secondari esterni (tra cui spiccano per rilevanza la politica delle associazioni in difesa dei consumatori in generale e degli audiolesi in particolare e la tradizione audiovisiva del paese in cui l'emittente televisiva trasmette) e interni (gestione del servizio di sottotitolazione, formazione dei rispeaker, difficoltà tecniche e operative), il rispeakeraggio può essere sia interlinguistico, sia intralinguistico. In questo ultimo caso, il rispeakeraggio può essere definito sia *verbatim*, o trascrizione ortografica del TP funzionale a una presunta parità di accesso al prodotto televisivo in generale, sia *non verbatim*, o riformulazione volta all'illusione della comprensione dell'utente finale. Alla luce di questa grande diversità di applicazioni, è stato innanzitutto limitato il campo di indagine al solo rispeakeraggio intralinguistico, in quanto maggiormente in linea con la realtà attuale delle emittenti televisive europee[172]. Si è poi deciso di affrontare due analisi contrastive parallele: tra rispeakeraggio intralinguistico *verbatim* e *shadowing* da una parte e rispeakeraggio *non verbatim* e interpretazione simultanea dall'altra; e tra rispeakeraggio intralinguistico (*verbatim* e *non verbatim* insieme) e sottotitolazione intralinguistica per non udenti di programmi pre-registrati.

Nel secondo capitolo, vista l'apparentemente straordinaria somiglianza psico-cognitiva tra il rispeakeraggio *verbatim* e lo *shadowing* e, per opposizione concettuale, tra il rispeakeraggio *non verbatim* e la tecnica simultanea dell'interpretazione, si è deciso di tenere separate le due tipologie di rispeakeraggio nell'analisi contrastiva concernente il processo. Grazie a un approccio meramente contrastivo nel primo caso e nel secondo caso a un approccio più socio-linguistico, si è potuto vedere come i quattro processi traduttivi non sono semplicemente diversi per finalità, come si era ipotizzato in un primo momento, ma anche e soprattutto per la maniera in cui il processo viene portato avanti dai rispeaker

172 Attualmente, solo alcuni canali della televisione olandese NOS e della britannica BBC offrono rispeakeraggio interlinguistico e solo in maniera molto sporadica.

professionali. In particolare, in base alle competenze necessarie individuate[173], essi si pongono in un *continuum* psico-cognitivo ideale nell'ordine crescente che segue: *shadowing*, rispeakeraggio *verbatim*, rispeakeraggio *non verbatim* e interpretazione simultanea. Completa idealmente la fila, il rispeakeraggio interlinguistico, che tuttavia non è stato preso in esame. Considerate tali differenze e nel tentativo di elaborare un quadro teorico di riferimento del rispeakeraggio come attività strategica, si è fatto abbondante ricorso alla letteratura degli studi sull'interpretazione e in particolare ai numerosi contributi sulle strategie traduttive. Il modello che più sembrava adattarsi a tale scopo è il modello di Kohn e Kalina (1996) rielaborato parzialmente e con meno strategie, ma pur sempre basato sul principio del rispeakeraggio come *strategic discourse processing* e sul modello strategico di comprensione del discorso di Kintsch e Van Dijk (1978) e van Dijk e Kintsch (1983)[174].

Questo modello non è stato testato in questo lavoro in quanto è particolarmente difficile registrare un rispeakeraggio (nel senso di processo, o TM) e decisamente sconveniente speculare sulle intenzioni del rispeaker senza una base scientifica e una metodologia congrue. Sorte dissimile è toccata al rispeakeraggio come prodotto, analizzato nel terzo capitolo. In questo capitolo, sono stati presi in esame i principali studi sulla sottotitolazione per non udenti e in particolare quelli inerenti gli aspetti semiotici da considerare nella fase di produzione. In particolare, si sono esaminate le caratteristiche principali della sottotitolazione per non-udenti, le più ricorrenti componenti non verbali (para- ed extra-linguistiche) e le migliori strategie finora elaborate per renderle nel TA. Il quadro che è emerso è stato considerato come l'*optimum* a cui mirare nel rispeakeraggio, con la consapevolezza che alcune caratteristiche sono impossibili da ottenere nella sottotitolazione in diretta in generale (come la sincronizzazione con le immagini e l'accuratezza nell'impaginazione dei sottotitoli); che alcune componenti sono poi tipiche della sottotitolazione filmica e per questo non necessarie nella sottotitolazione di programmi in diretta (come gli effetti sonori off e le musiche di sottofondo); e che altre sono infine difficili da rendere a causa dei numerosi vincoli tecnici e operativi. Le sole componenti non

173 Competenza fonetica, metalinguistica (ulteriormente ripartita in redazionale e psico-cognitiva), sintetica (valida solo nel caso del rispeakeraggio non verbatim e dell'interpretazione simultanea) e di (conoscenza del) genere (da sottotitolare).

174 Le strategie in questione sono le strategie di segmentazione, neutralizzazione, evasione, adattamento, monitoraggio, memorizzazione inferente e infine di emergenza.

verbali indispensabili alla comprensione del TA sono la punteggiatura e l'identificazione del parlante tramite il cambio di colore.

Quanto allo strumento di analisi necessario a esaminare le operazioni di resa della componente verbale, si sono presi in considerazione i maggiori contributi nel settore degli studi sulla traduzione e sulla traduzione audiovisiva in generale e sulle strategie in particolare. Anche in questo caso si è optato per la tassonomia che meglio risponde alle esigenze dell'oggetto in esame, vale a dire quella proposta da Gambier (2006)[175]. L'applicazione di tale strumento è tuttavia inutile e sterile senza una contestualizzazione dell'analisi che ne permetta una lettura intelligente e scientificamente esatta. Ecco quindi che si è passati alla valutazione dei più adeguati strumenti di analisi macrotestuale in vista dello studio delle strategie utilizzate dai rispeaker per rendere un prodotto audiovisivo. Vista la natura poliedrica del rispeakeraggio come prodotto, lo strumento più flessibile è parso essere la *genre analysis*, che permette di individuare con chiarezza la struttura generale del TP e la natura linguistica[176] delle fasi e sottofasi che lo compongono. Collateralmente consente altresì di individuare l'unità minima di analisi. Per affinare ulteriormente l'analisi è stato stimato molto utile l'approccio multimodale, che prevede l'analisi di tutte le componenti semiotiche del TP in vista della valutazione del grado di interazione tra la componente verbale e quella non verbale e conseguentemente della maggiore o minore possibilità per il sottotitolatore di eliminare alcuni elementi linguistici considerati ridondanti.

Il metodo di analisi appena proposto è stato utilizzato nel quarto capitolo per analizzare il corpus di riferimento: la trascrizione ortografica di otto ore di *BBC News* e dei relativi sottotitoli, che negli intenti degli *Access Services* dell'azienda sono catalogati come *verbatim*. Innanzitutto, si è provveduto all'analisi di genere di *BBC News*[177], che ha permesso, grazie anche al contributo dell'ITC e dell'Ofcom, l'identificazione dell'unità minima di analisi, vale a dire quella che è stata definita come macro-unità concettuale, in opposizione alle micro-unità concettuali che la compongono e costituiscono il secondo

175 La tassonomia di Gambier raggruppa le principali strategie traduttive in tre grandi gruppi: riduzione, espansione e semplificazione semantica.
176 Per natura linguistica si intendono qui i vari criteri di valutazione del testo orale in contrapposizione a quello scritto: complessità grammaticale, densità lessicale, velocità di eloquio, natura dell'informazione (nota o nuova), ecc.
177 BBC News è generalmente composto da cinque fasi non sempre consecutive: titoli, reportage dal vivo, servizi pre-registrati, previsioni meteorologiche e sommario. Ognuna di esse si scompone in numerose sotto-fasi e di punti di transizione

livello di analisi. Le due trascrizioni sono state quindi segmentate in macro-unità concettuali e successivamente allineate. Da una prima analisi, ci si è immediatamente resi conto che i sottotitoli erano molto simili al TP, ma che piccole variazioni erano abbastanza frequenti. Sulla base del modello nel capitolo precedente si è quindi provveduto alla realizzazione di una tassonomia *ad hoc* che permettesse di analizzare il materiale in esame. La prima grande ripartizione è stata quindi fatta tra macro-unità concettuali rese e macro-unità concettuali non rese. Nella prima categoria sono state catalogate soltanto quelle che erano state palesemente non rese nei sottotitoli (principalmente a seguito di una intenzionale omissione) e nella seconda tutte le altre macro-unità esaminate. In particolare, si è provveduto a una bipartizione che suddividesse le macro-unità ripetute letteralmente (tratti dell'oralità compresi) e quelle che invece sono state oggetto di una qualche alterazione. A loro volta queste sono state suddivise in espansioni, riduzioni ed errori. Nella categoria riduzioni sono state fatte confluire sia le omissioni di micro-unità concettuali, sia le compressioni. Sia le espansioni, sia le riduzioni sono state ulteriormente scomposte in strategie semantiche e non semantiche, ossia in strategie che hanno alterato la natura delle informazioni veicolate dal testo nella sua pluralità semiotica (componenti audio e video verbali e non verbali) e in strategie che hanno semplicemente modificato quantitativamente la natura del mero testo. Quanto agli errori, essi sono stati inclusi nelle strategie di alterazione, sebbene non siano l'oggetto di volontarie modifiche da parte del rispeaker. Tuttavia, a causa di un palese errore di interazione tra l'uomo e la macchina, essi risultano in un'alterazione evidente del TA e rischiano di compromettere la comprensione del pubblico[178].

Da questa dettagliata analisi, sono emersi numerosi dati riguardanti quelleche possono essere definite le migliori prassi in ambito di rispeakeraggio *verbatim*. In particolare, si è notata la capacità dei rispeaker della BBC di rendere più dell'80% delle macro-unità concettuali del TP. Questo dato sale oltre il 90% se si considera che una fetta consistente delle strategie di omissione di macro-unità concettuali sono, in realtà, funzionali al principio che potrebbe essere definito dell'economicità semiotica del rispeakeraggio, cioè a dire, della riduzione delle parole al minuto da ripetere tramite l'ellissi oculata delle ridondanze e di quelle che sono definite espressioni formulaiche. Nella categoria delle

178 In realtà, nemmeno gli errori più compromettenti sono stati considerati come degli errori implicanti la non resa della macro-unità concettuale in questione. La ragione di questa decisione risiede innanzitutto nella volontà di avere un dato unico inerente il tasso di errori e in seconda battuta nell'impossibilità di decidere aprioristicamente l'incapacità da parte di uno spettatore di cogliere il significato intenzionale di un sottotitolo.

macro-unità rese, la maggior parte delle strategie sono da considerarsi appannaggio della ripetizione, che caratterizza quasi la metà delle strategie traduttive dei rispeaker, sebbene sia possibile per un numero limitato di macro-unità concettuali consecutive. Ancora una volta, questo dato sale sensibilmente se vi si aggiungono tutte le operazioni di minima alterazione al TP (espansione, compressione e omissione di micro-unità semanticamente non rilevanti a fini informativi). Tra le strategie di alterazione, che compongono comunque un terzo delle strategie traduttive dei rispeaker della BBC, spicca la riduzione, sia semantica (sinonimia, sintesi, riformulazione, ecc.), sia non semantica (principalmente omissione dei tratti dell'oralità). Poco spazio spetta invece all'espansione, perlopiù non semantica, e agli errori, che hanno un'incidenza di circa l'1% sul totale delle macro-unità analizzate.

Come si è detto precedentemente, il rispeakeraggio di *BBC News* effettuato da parte dei rispeaker della BBC è da considerarsi *verbatim*. Come si è visto tuttavia, senza considerare l'inevitabile inserimento della punteggiatura, la resa *verbatim* del TP non è possibile per più della metà dello stesso, in quanto intervengono nella resa sia fattori esterni (velocità eccessiva di eloquio del TP, stretta interrelazione tra la componente verbale e le immagini del TP, ecc.), sia fattori interni (principio dell'economicità semiotica del TA, familiarità con l'argomento in questione, falla nell'interazione uomo-macchina, ecc.). ecco quindi che risulta necessario rivedere il concetto di rispeakeraggio *verbatim* in ottica funzionale e professionale, limitandone l'ambito d'uso al maggior numero possibile di macro-unità concettuali da ripetere, salvo i casi di eccessiva presenza di tratti dell'oralità (false partenze, auto-riformulazioni, esitazioni, ecc.).

Nel quinto capitolo, si è cercato di applicare le linee guida derivanti da questa analisi all'interno di un progetto internazionale (progetto SALES) avente come obiettivo finale l'accessibilità dei sordi segnanti italiani al telegiornale in lingua italiana. Dopo una prima sperimentazione che ha ottenuto risultati molto contraddittori tra di loro e in seguito alla lettura dell'abbondante letteratura in materia, ci si è immediatamente resi conto che la realtà tra le comunità sorde inglesi e italiane sono fortemente diverse in termini di alfabetismo e di dimestichezza con la lettura dei sottotitoli. Inoltre, un'ulteriore divario sempre di natura linguistica, forse ancor più incolmabile, esiste all'interno della comunità sorda italiana tra segnanti e oralisti, ossia tra sordi che comunicano tramite la lingua dei segni (linguaggio visivo-gestuale) e altri che comunicano tramite la lingua orale (linguaggio fonico-uditivo). Al loro interno un'ulteriore grossolana suddivisione degna di nota è da farsi

tra sordi pre-linguali e post-linguali. La ramificazione proseguirebbe praticamente all'infinito se si considerassero altri fattori ugualmente importanti come l'età, l'istruzione, la zona geografica, la professione, le abitudini di lettura, ecc. Vista questa marcata disomogeneità, si è cercato quindi di adottare un atteggiamento prudente: piuttosto che cercare di produrre sottotitoli in grado di soddisfare tutte le esigenze, con il forte rischio di non accontentare nessuno, si è provveduto ad analizzare le competenze linguistiche di un campione quantitativamente importante di sordi segnanti pre- o peri-linguali e di progettare successivamente dei sottotitoli a loro uso e consumo. Dai risultati è affiorato l'*identikit* dello spettatore sordo segnante medio pre- (o peri-) linguale medio, che ha permesso di indirizzare la ricerca verso forme di sottotitolazione *non verbatim*, maggiormente funzionali e quindi accessibili. Per giungere allo scopo finale del progetto, la creazione di linee guida valide per il rispeakeraggio *non-verbatim* di un telegiornale in lingua italiana accessibile ai sordi segnanti, si è proceduto alla somministrazione di numerosi test di lettura e all'analisi delle interpretazioni in LIS (Lingua dei Segni Italiana) delle edizioni italiane del telegiornale dedicato ai non-udenti. Grazie all'incrocio dei dati così ottenuti, si è riusciti a elaborare una serie completa di linee guida che prevedono, a ogni livello linguistico (lessicale, sintattico, semantico e pragmatico), una chiarezza e una linearità tali da garantire la comunicazione nei brevi vincoli spazio-temporali imposti dal prodotto televisivo. Tra le principali strategie da mettere in pratica spiccano per importanza una certa brevità della frase, la struttura sintattica di base, la coordinazione e l'uso di lessemi e di espressioni formalmente noti (possibilmente) e semanticamente non ambigui.

Queste linee guida sono quindi state testate e proposte in un contesto professionale reale: il rispeakeraggio del primo e del secondo dibattito televisivo avvenuto nel 2006 tra i due candidati di allora alla poltrona di Presidente del Consiglio. Appurata la validità, almeno in termini ricettivi e percettivi, delle linee guida appena derivate, si è pensato a una loro possibile applicazione in un contesto anglofono. Vista la resistenza mostrata dalla comunità sorda britannica a operazioni di questo genere (dettata in parte anche dalla maggiore familiarità che essa ha con la sottotitolazione in tempo reale) e in secondo luogo viste le difficoltà tecniche e organizzative nel procedere a una contro-verifica simile nel Regno Unito, è sembrato opportuno abbandonare temporaneamente questa strada per indicarne una già parzialmente percorsa e per certi versi vicina ai dati ottenuti nel quadro del progetto SALES: l'uso del Plain English come criterio del rispeakeraggio *non verbatim* volto

anch'esso alla piena accessibilità degli utenti al testo da leggere e comprendere. Alla luce di queste considerazioni, risulta evidente come la definizione del concetto di rispeakeraggio *non verbatim* rivesta totalmente un significato di piena accessibilità e non di semplice riduzione testuale, avvicinandosi molto di più di quanto non sembrerebbe a prima vista al rispeakeraggio *verbatim*.

Nel sesto e ultimo capitolo, il percorso iniziato con la descrizione funzionale del rispeakeraggio e continuato con la teorizzazione di modelli strategici e l'applicazione di uno di essi all'analisi testuale di un corpus di riferimento raggiunge il suo compimento nella progettazione di un quadro teorico per la didattica del rispeakeraggio in ambito accademico basato su importanti esperienze professionali. Per raggiungere questo obiettivo è stato scelto il modello pedagogico e didattico di Louis D'Hainaut del 1975 adattato da Safar nel 1992 per l'insegnamento dell'interpretazione e della traduzione e aggiornato nel 2006 per coprire anche aree oramai in diffusione nelle università europee come la traduzione audiovisiva. Sulla base degli insegnamenti derivanti dal progetto SALES e grazie ai primi esperimenti nel settore della formazione al rispeakeraggio in ambito sia accademico, sia professionale, è stato ideato un modello tripartito per la formazione professionale del rispeakeraggio in ambito accademico così composto:

- scopi e obiettivi;
- strumenti e insegnamento;
- valutazione.

Grazie alla centralità data agli obiettivi da raggiungere e grazie alla complementarietà in materia di competenze tra quelle del rispeaker e quelle dell'interprete di simultanea è stato possibile articolare un modello che potesse essere immediatamente funzionale evitando così le numerose tappe preparatorie indispensabili alla formazione in ambito professionale, che ne allungano altrettanto inevitabilmente i tempi. Dal punto di vista linguistico, l'aspetto maggiormente interessante è il confronto e l'oscillazione tra rispeakeraggio *verbatim* e rispeakeraggio *non verbatim*, da non considerare semplicemente come due punti di un *continuum* psico-cognitivo, come era stato inizialmente ipotizzato, ma come due macro-strategie comunicative volte entrambe alla piena ricezione da parte dell'utente finale del testo sottotitolato.

Bibliografia

Abercrombie, N. (1996) *Television and Society*. Cambridge: Polity Press

Accademia Aliprandi *et al.* (2007) *Atti del convegno "La resocontazione: competenze, tecniche, organizzazioni"* Povo (TN) 2 dicembre 2006.

Acero, A. Huang, X. Hon, H. (2001) Spoken Language Processing. Prentice Hall.

AENOR. (2003) *Norma Española UNE 153010-2003. Subtitulado para personas sordas y personas con discapacidad auditiva. Subtitulado a través del teletexto*. Madrid: AENOR.

AIIC (2006) *Code of professional ethics*. http://www.aiic.net/community/attachments/ViewAttachment.cfm/a24p54-1749.pdf?&filename=a24p54%2D1749%2Epdf&page_id=54

Alexieva, B. (1983) "Compression as a Means of Realisation of the Communicative Act in Simultaneous Interpreting". In Fremdsprachen 4.

Allen, R. (1989) "Bursting bubbles: "Soap opera" audiences and the limits of genre". In Seiter, E. Borchers, H. Kreutzner, G. e E.-M. Warth (a cura di) *Remote Control: Television, Audiences and Cultural Power*. London: Routledge. Pp. 44-55

Amatucci, L. (1995). "La scuola italiana e l'istruzione dei sordi". In Porcari Li Destri, G. e Volterra, V. a cura di (1995). Passato e presente: uno sguardo sull'educazione dei sordi in Italia. Napoli : Gnocchi.

Anderman, G. e Rogers M. a cura di (2003). Translation today. Trends and Perspectives. Clevedon: Multilingual Matters.

Angelelli, C. V. (2000) "Interpretation as a Communicative Event: A Look through Hymes' Lenses". In *Meta* XLV, 4. Pp. 580-592.

Angelelli, C. V. (2004) Medical Interpreting and Cross-cultural Communication. Cambridge: Cambridge University Press.

Araujo, V. (2004). "Closed subtitling in Brazil". In Orero, P. (a cura di) Topics in audiovisual translation. Amsterdam e Philadelphia: John Benjamins.

Arma (2007) *Dal parlato al (tra)scritto. La resocontazione parlamentare tra stenotipia e riconoscimento del parlato*. Tesi di laurea non pubblicata. Forlì: SSLMIT

Arumí Ribas, M. e P. Romero Fresco (2008) "A Practical Proposal for the Training of Respeakers 1". In JoSTrans, 10. http://www.jostrans.org/index.php

Aubry, P. (2000). The television without frontiers directive, cornerstone of the European broadcasting policy. European audiovisual observatory.

Auscap 1999. TBS (Digital Conversion) Act 1998 Draft Captioning Standards. www.dcita.gov.au

Australia (1998) *TBS (Digital Conversion) Act 1998*, Schedule 1, §38. http://scaleplus.law.gov.au/html/pasteact/2/3156/0/PA000070.htm#_Toc415905975

Baaring, I. (2006) "Respeaking-based online subtitling in Denmark". In Eugeni, C. e G. Mack (a cura di) *Proceedings of the first international seminar on real-time intralingual subtitling*. InTRAlinea special issue: respeaking. www.intralinea.it

Baker, R. *et al.* (1984) Oracle Subtitling for the Deaf and Hard-of-hearing. Southampton: Southampton University.

Baker, R. *et al.* (1986) "Television and video technology in the education of deaf children". *British Journal of Audiology 20*. Pp. 1-13.

Baker, M. (1992) *In other words: a coursebook on translation*. Londra e New York: Rouledge.

Baldry, A. and P. J. Thibault (2005) *Multimodal transcription and text analysis*. Oakville: Equinox publishing

Barik, H. (1971) "A description of various types of omissions, additions and errors encountered in simultaneous interpretation". In *Meta* 15:1. Pp. 199-210.

Bartoll, E. (2004) "Parameters for the classification of subtitles", in Orero P. (a cura di) Topics in Audiovisual Translation. Amsterdam e Philadelphia: John Benjamins.

Bassnett-McGuire, S. e A. Lefevere a cura di (1990) *Translation History and culture*. Londra e New York: Pinter Publishers.

Baum, L. E. Petri, T. Soules, G. e N. Weiss (1970) "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains". In *Annual Mathematical Statistics*, vol. 41.

Bazzanella, C. (1994). *Le facce del parlare: un approccio pragmatico all'italiano parlato*. Scandicci : La nuova Italia.

BBC *et al.* (1976) *Broadcast Teletext Specification*. Londra: BBC.

BBC. 1998. BBC Subtitling Guide. London: British Broadcasting Corporation.

Bell A. e P. Garrett a cura di (1998) *Approaches to Media Discourse*. Oxford: Blackwell.

Bell, A. (1984) "Language style as audience design". In *Language in society 13*.

Bernero, R. e H. Bothwell (1996) *Relationship of hearing impairment to educational needs*. Springfield: Illinois DPHI e Office SPI.

Bernuzzo, G. (1996) *E venne il giorno di santa Apollonia*. Milano : Editrice Nuovi Autori.

Bhatia V. K. (1993) *Analysing Genre. Language Use in Professional Settings*. London: Longman

Bhatia, V. K. (2002) "Applied genre analysis: a multi-perspective model". *Iberica 4*. Pp. 3-19.

Bignell, J. (2004). *An introduction to television studies*. Londra e New York. Routledge.

Blini, L. e F. Matte Bon (1996) "Osservazioni sui meccanismi di formazione dei sottotitoli". In Heiss, C. e Bollettieri Bosinelli, R.M. (a cura di). *Traduzione multimediale per il cinema, la televisione e la scena. Atti del convegno internazionale Forlì 26-28 ottobre 1995*. Bologna : CLUEB.

Blizzard, T. (2005). *Interview with Toby Blizzard*. www.subtitleproject.net

Bollettieri Bosinelli, R.M. et al. a cura di (2000). La traduzione multimediale: quale traduzione per quale testo? Atti del convegno internazionale Multimedia translation: which translation for which text? Forlì, 2-4 aprile 1998. Bologna : CLUEB.

Bordwell, D. (1989) *Making Meaning: Inference and Rhetoric in the Interpretation of Cinema*. Cambridge, MA: Harvard University Press.

Bowen, M. (1993) "Building Up Speed (Simultaneous Interpreting of Read Speech)". In Losa, E. F. (a cura di) *Keystones of Communication. Proceedings of the 34th Annual Conference of the American Translators Association. October 6th-10th 1993 Philadelphia, Pennsylvania*. Medford, NJ: Learned Information Inc.

Bowers, C. (1998) *The nature and constraints of subtitling with particular reference to intralingual subtitling and other forms of media access for the deaf and hard-of-hearing.* Tesi di dottorato non pubblicata. Manchester: University of Manchester.

Brette, N. (1982) "Sous-titres, le crève-cœur des réalisateurs". In *Cahiers du cinéma 338*.

Bristow, R. (1987). Teletext sub-titles. A preliminary exploration. Londra : BBC Broadcasting research.

Brown, G. e G. Yule (1983) Discourse analysis. Londra e New York: Cambridge university press.

Brunetto, F. (2005) *I sottotitoli per non udenti: considerazioni generali e applicazioni pratiche in Italia, Regno Unito e Belgio*. Tesi di Laurea non pubblicata in Traduzione dall'italiano all'inglese. Forlì: SSLMIT.

Bruti, S. e E. Perego (2005) "Translating the expressive function in subtitles: the case of vocatives". In Sanderson, J.D. (a cura di) *Research on Translation for Subtitling in Spain and Italy*. Alicante: Publicaciones de la Universidad de Alicante. Pp. 27-48.

CAB (2003) *Closed Captioning Standards and Protocol for Canadian English Language Broadcasters* http://www.cab-acr.ca/english/social/captioning/captioning.pdf

Cameron, D. (2001) *Working with spoken discourse*. Londra: Sage.

Candlin C. e M. Gotti a cura di (2004) "Intercultural Discourse in Domain-Specific English", special issue of *Textus, 17/1*.

Captionmax (2002) *Captioning Styles* http://www.captionmax.com/pages/ViewerInfo/VI_style.html

Carlson, M., *et al.* (1990) *Descriptive Video Service Style Manual*. Boston: WGBH Educational Foundation

Carroll, M. (2004). "Subtitling: changing standards for new media?" Lisa –The globalization insider. Newsletter XIII/3.3.

Carter, R. (1997). Investigating English discourse. Londra e New York: Routledge.

Caselli, M., Maragna, S., Rampelli, L. e V. Volterra (1994) *Linguaggio e sordità. Parole e segni per l'educazione dei sordi*. Firenze: La Nuova Italia.

Casetti, F. e F. Villa a cura di (1992) *La storia comune: funzioni, forma e generi della fiction televisiva*. Torino : RAI-Nuova ERI.

Cattryse, P. (2000) "Media translation. Plea for an interdisciplinary approach". In *VS 85-87*.

Cenelec (2003) *Standardisation requirements for access to digital TV and Interactive services for the disabled people*. www.cenelec.org

CFV (1996) *Captioning Key: Guidelines and Preferred Styles*. Spartanburg, SC: National Association of the Deaf.

Chandler, D. (1994) *The Grammar of television and film*. http://users.aber.ac.uk

Chatman, S. (1978). Story and discourse. Ithaca e Londra: Cornell university press.

Chaume, F. (1997) "Translating non-verbal information in dubbing". In Poyatos, F. (a cura di). *Non-verbal communication and translation. New perspectives and challenges in literature, interpretation and the media*. Amsterdam e Philadelphia: John Benjamins

Chaume, F. (2002) "Models of research in audiovisual translation". In *Babel 48 (1)*.

Chaume, F. (2004) "Film studies and translation studies: Two disciplines at stake in audiovisual translation". In *Meta. 49 (1)*. Pp. 12-24.

Chernov, G. V. (2004) *Inference and Anticipation in Simultaneous Interpreting: A Probability-Prediction Model*. Amsterdam e Philadelphia: John Benjamins.

Chesterman, A. (1993) "From «is» to «ought». Laws, norms and strategies in Translation Studies". In *Target 5 (1)*.

Chesterman, A. (1997) *Memes of Translation. The spread of Ideas in Translation Studies*. Amsterdam e Philadelphia: John Benjamins.

Clark, M. D. *et al.* (2001) Context, cognition and deafness. Washington: Gallaudet university press.

Corazza, S. e V. Volterra (1987) "Introduzione". In Volterra, V. a cura di (1987). *La lingua italiana dei segni: la comunicazione visivo-gestuale dei sordi*. Bologna : Il Mulino

Coro, G. (2004) "Modulation Spectrogram (MS) nel Riconoscimento Automatico del parlato". In *Proceedings of AISV*. Napoli: Università di Napoli Federico II.

Cortelazzo, M. A. (1985) "Dal parlato al (tra)scritto: i resoconti stenografici dei discorsi parlamentari". In Holtus, G and E. Radtke (eds.) *Gesprochenes Italienisch in Geschichte und Gegenwart*. Tübingen: Universität Tubingen.

Cosi P. e E. Magno Caldognetto (1996) "Lips and Jaw Movements for Vowels and Consonants: Spatio-Temporal Characteristics and Bimodal Recognition Applications" In D.G. Starke e M. E. Henneke (a cura di) *Speechreading by Humans and Machine: Models, Systems and Applications*, NATO ASI, vol.150, Springer-Verlag.

Coulthard, M. a cura di (1992) *Advances in spoken discourse analysis*. Londra e New York: Routledge.

Cowan, N. (1995) *Attention and Memory: An Integrated Framework*. New York: Oxford University Press

Cox, R. V. *et al.* (2000) "Speech and language processing for next-millennium communications services". In *Proceedings IEEE, vol. 88*.

CRTC (1995) "Public Notice CRTC 1995-48: Introduction to Decisions Renewing the Licences of Privately-Owned English-Language Television Stations".

Crystal (2001) *Language and the internet*. Cambridge : Cambridge university press.

Crystal D. e D. Davy (1985) *Investigating English Style*. London: Longman.

D'Hainaut, L. (1975) *Concepts et méthodes de la statistique, Vol. 1*. Bruxelles: Editions Labor

D'Ydewalle, G. (1999) "The psychology of film perception". In *Psicologia Italiana.* Rivista della società italiana di psicologia XVI (1-3).

D'Ydewalle, G. *et al*. (1987) "Reading a Message when the same Message Is available Auditorily in Another Language: The Case of Subtitling." In Regan e Lévy-Schoen (a cura di) *Eye Movements: From Physiology to Cognition*. Amsterdam.

D'Ydewalle, G. *et al.* (1991) "Watching subtitled television. Automatic reading behaviour". *Communication research* 18 (5).

Dam, H. V. (1993) "Text Condensing in Consecutive Interpreting". In Gambier, Y e Tommola, J. (a cura di) *Translation and Knowledge SSOTT IV-Scandinavian Symposium on Translation Theory Turku, Finland 4-6 June 1992.* Turku: University of Turku-Centre for Translation and Interpreting.

Dam, H. V. (1996) "Text Condensation in Consecutive Interpreting. Summary of a Ph.D. dissertation". In *Hermes* 17.

Dam, H. V. (2001) "On the Option between Form-based and Meaning-based Interpreting: The Effect of Source Text Difficulty on Lexical Target Text Form in Simultaneous Interpreting". In *The Interpreters' Newsletter 11*.

Darò, V. (1995) "Ricerche sulle componenti dell'interpretazione simultanea". In *Il Traduttore Nuovo* 1995/1.

Darò, V. (1997) "Experimental Studies on Memory in Conference Interpretation". In *Meta* 42 (4).

Davey, M. "Silenzio si parla", inserto *D La Repubblica delle Donne*. *La Repubblica*, 23 aprile 2005.

Davis, K. H. Biddulph, R. Balashek, S. (1952) "Automatic recognition of spoken digits". In *Journal of the Acoustical Society of America*, vol. 24, no. 6.

De Beaugrande, R. (1980). *Text, discourse and process. Towards a multidisciplinary science of text.* Londra: Longman.

De Beaugrande, R. e W. Dressler (1981) *Introduction to text linguistics*. New York : Longman.

De Certeau, M. (1980) *L'invention du quotidien 1. arts de faire* Paris: Editions Gallimard.

De Groot, A. M. B. e J. F. Kroll a cura di (1997) *Tutorials in Bilingualism: Psycholinguistic Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates

De Korte, T. (2006) "Live inter-lingual subtitling in the Netherlands". In Eugeni, C. e G. Mack (a cura di) *Proceedings of the first international seminar on real-time intralingual subtitling*. InTRAlinea special issue: respeaking. www.intralinea.it

De Linde, Z. (1996) "Le sous-titrage intralinguistique pour les sourds et les malentendants". In Gambier, Y. (a cura di) *Les transferts linguistiques dans les médias audiovisuels.* Paris : presses universitaires du septentrion.

De Linde e Kay (1999) *The semiotics of subtitling.* Manchester: St. Jerome

De Linde, Z. e N. Kay (1999). *The semiotics of subtitling*. Manchester: St. Jerome Publishing.

De Mauro, T. (1997) *Guida all'uso delle parole*. Roma: Editori Riuniti.

de Seriis (2006) "Il Servizio Sottotitoli RAI". In Eugeni, C. e G. Mack (a cura di) *Proceedings of the first international seminar on real-time intralingual subtitling*. InTRAlinea special issue: respeaking. www.intralinea.it

Deaf broadcasting council (2000) *Access to TV in Europe*. DBC www.deafbroadcastingcouncil.org.U.K

Deerwester, S. *et al*., (1990) "Indexing by latent semantic analysis". In *Journal of the American Society of Informatic Science*, vol. 41.

Delabastita, D. 1989. "Translation and mass-communication: film and TV translation as evidence of cultural dynamics". In *Babel* 35 (4).

Den Boer, C (2001) "Live interlingual subtitling". In Gambier, Y. e Gottlieb, H. (a cura di). *(Multi)media translation. Concepts, practices and research*. Amsterdam e Philadelphia: John Benjamins.

Denes, P. (1959) "The design and operation of the mechanical speech recognizer at University College London.". *Journal of the British Institute of Radio Engineers*, vol.19, no. 4.

Desmedt (2002) *Le sous-titrage pour sourds et malentendants*. Tesi di DESS non pubblicata. Bruxelles: ISTI-HEB

Díaz-Cintas (2003) "Audiovisual translation in the third millenium". In Anderman, G. e Rogers, M. (a cura di). *Translation today. Trends and perspectives*. Clevedon: multilingual matters.

Díaz-Cintas (2004) “In search of a theoretical framework for the study of audiovisual translation”. In Orero, P. (a cura di). *Topics in audiovisual translation*. Amsterdam e Philadelphia: John Benjamins.

Díaz-Cintas (2007) “The subtitler’s profession”. Intervento alla conferenza internazionale Multidimensional Translation: LSP Translation Scenarios 30/04-04/05 2007. Vienna.

Díaz-Cintas, J., Orero, P., Remael, A. a cura di (2005). *Atti della conferenza International conference on Audiovisual Translation Media for All 6-8 giugno 2005*. Barcelona.

Dimitriu, R. (2004) “Omission in Translation”. In *Perspectives* 12: 3.

Dittman, D. *et al.* (1989) *The Caption Center Manual of Style*. Boston: WGBH Educational Foundation

Dollerup, C. (1974) On Subtitles in TV programmes. *Babel* 20 (4). Pp. 197-202.

Donaldson, C. (2001) *SDI Grammar Guide* UK: SDI Media.

Donaldson (2004) Intervento alla tavola rotonda sull’accessibilità all’interno della conferenza internazionale Languages and the Media 2004. Berlino: InterContinental Hotel.

Dries, J. (1995). *Dubbing and subtitling. Guidelines for production and distribution*. Düsseldorf: European institute for the media.

Duda, R. O. Hart, P. E. (1973) *Pattern Classification and Scene Analysis*. New York: Wiley.

Dudley-Evans T. and St. John M.J. (1998) *Development in English for Specific Purposes, a Multi-disciplinary Approach*. Cambridge: Cambridge University Press

Durante, M.M. (2005) *La Televisione e la Sordità: Un'analisi della Sottotitolazione per Non Udenti delle Emittenti Televisive Italiane*. Bruxelles: Institut Supérieur de traducteurs et interprètes.

EBU (1997) “Towards a standardization of dubbing and subtitling procedures”. *EBU review. Programmes, administration, law.* XXXVIII (6), 31.

Eco (2003) *Dire (quasi) la stessa cosa*. Milano: Bompiani.

EIA (2002) “Recommended Practice for Line 21 Data Service.” EIA-608-A

Engel, F. *et al.* a cura di (1985). *Cognitive modelling and interactive environments in language learning*. Berlino e New York: Springer-Verlag.

ETSI (2002) *Digital Video Broadcasting (DVB), Subtitling systems, ETSI EN 300 743 V1.2.1 (2002-06)*. http://webapp.etsi.org

Eugeni, C. (2003) *Il teatro d'opera e l'adattamento linguistico simultaneo*. Tesi di laurea non pubblicata. Forlì: Università di Bologna.

Eugeni, C. (2006a) “Introduzione al rispeakeraggio televisivo”. In Eugeni, C. and G. Mack (a cura di) *Proceedings of the first international seminar on real-time intralingual subtitling*. InTRAlinea special issue: respeaking. www.intralinea.it

Eugeni, C. (2006b) “For a didactics of respeaking”. Intervento alla conferenza internazionale Languages and the Media. Berlino: Hotel InterContinentale.

Eugeni, C. (2007) “Il rispeakeraggio televisivo per sordi: per una sottotitolazione mirata del TG”. InTRAlinea, vol. 8 http://www.intralinea.it/volumes/eng_open.php?id=C0_60_2

Eugeni, C. (2008a) “A Sociolinguistic Approach to Real-time Subtitling: Respeaking vs. Shadowing and Simultaneous Interpreting”. C.J. Kellett Bidoli e E. Ochse (a cura di), English in

International Deaf Communication, Linguistic Insights series vol. 72, Bern: Peter Lang. Pp. 357-382.

Eugeni, C. (2008b) "Respeaking political debate for the deaf: the Italian case". In Baldry, A. e E. Montagna (a cura di). *Interdisciplinary Perspectives on Multimodality: Theory and practice*. Campobasso: Palladino Editore. Pp. 191-205

Eugeni, C. (in stampa). *Respeaking the TV for the Deaf. For a real special needs-oriented subtitling*.

Eugeni, C. e G. Mack a cura di (2006) *Proceedings of the first international conference on real-time intralingual subtitling*. InTRAlinea Special issue: respeaking www.intralinea.it

European Broadcasting Union (2005) *EBU tech 3295 - The EBU Metadata Exchange Scheme - P_META 1.2*. On-line publication at http://www.ebu.ch/CMSimages/en/tec_doc_t3295_v0102_tcm6-40957.pdf (ultimo accesso 5 October 2008)

Evans, M. J. (2003) "Speech Recognition in Assisted and Live Subtitling". In *BBC Research and Development White Papers*. www.bbc.co.uk

Even-Zohar, I. (1990) "Polysystem theory". *Poetics today 11* (1).

Facchini, G.M. (1981) "Riflessioni storiche sul metodo orale e il linguaggio dei segni in Italia". In Volterra, V. (a cura di). *I segni come parole: la comunicazione dei sordi.* Torino : Boringhieri.

Facchini, G.M. (1995) "Commenti al Congresso di Milano del 1880". In Porcari Li Destri, G. e Volterra, V. (a cura di). *Passato e presente: uno sguardo sull'educazione dei sordi in Italia.* Napoli : Gnocchi.

Færch, C. e G. Kasper (1983) "Plans and Strategies in Foreign language Communication". In Færch, C. e G. Kasper (a cura di) *Strategies in Interlanguage Communication.* Londra: Longman.

Fairclough, N. (1992) *Discourse and social change*. Cambridge: Polity press.

Falbo, C., Russo, M. e F. Straniero Sergio a cura di (1999). *Interpretazione simultanea e consecutiva. Problemi teorici e metodologie didattiche.* Milano : Hoepli.

Fawcett, P. (1997) *Translation and language – Linguistic theories explained*. Manchester: St Jerome.

FCC (1997) *Captioning*. Federal communications commission. www.fcc.gov

FCC (2003) *Captioning*. Federal communications commission. www.fcc.gov

Ferguson, J. a cura di (1980) *Hidden Markov Models for Speech*. Princeton: IDA.

Feuer, J. (1992) "Genre study and television". In Allen, R. (a cura di) *Channels of discourse reassembled*. Londra: Routledge.

Flanagan, J. L. (1972) *Speech Analysis, Synthesis and Perception*. New York: Springer-Verlag.

Flowerdew J. a cura di (2002) *Academic Discourse*. Londra: Longman

Forgie, J. W. e C. D Forgie (1959) "Results obtained from a vowel recognition computer program". In *Journal of the Acoustical Society of America*, vol. 31, no. 11.

Fowler, A. (1989) "Genre". In Barnouw, E. (a cura di) *International Encyclopaedia of Communications*, Vol. 2. New York: Oxford University Press, pp. 215-7

Franco, E. e Araújo, V. (2003) "Reading television". In *The translator 9*(2).

Freedman, A. e P. Medway a cura di (1994) *Genre and the New Rhetoric*. Londra: Taylor & Francis

Fry, D. B. (1959) “Theoretical aspects of the mechanical speech recognition”. In *Journal of the British Institute of Radio Engineers*. vol. 19, no. 4.

Furth, H. G. (1991) *Pensiero senza linguaggio: implicazioni psicologiche della sordità*. Roma : A. Armando.

Furui, S. (1981) “Cepstral analysis technique for automatic speaker verification”. In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol.ASSP-29.

Furui, S. (1986) “Speaker independent isolated word recognition using dynamic features of speech spectrum”. In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. ASSP-34.

Furui, S. (1997) “Recent advances in robust speech recognition”. In *Proceedings ESCA-NATO Workshop on Robust Speech Recognition for Unknown Communication Channels*, Pont-a-Mouson, France.

Gaell, R. a cura di (1999) *Subtitling consumer report*. Scotland: Royal national institute for deaf people.

Gambier, Y. (1992) “La reformulation – pratique intralinguistique et interlinguistique”. In *KOINÉ. Annali della Scuola Superiore per Interpreti e Traduttori “San Pellegrino”*, II, 1-2.

Gambier, Y. (1994) Audiovisual communication : “typological detour”. In Dollerup, C. e Lindengaard, A. (a cura di). *Teaching translation and interpreting 2: insights, aims, visions.* Amsterdam e Philadelphia: John Benjamins.

Gambier, Y. (1996). *Les transferts linguistiques dans les medias audiovisuels.* Villeneuve d’Asq : Presses universitaires du Septentrion.

Gambier, Y. a cura di (1998) *Translating for the media*. Turku : university of Turku.

Gambier, Y. (1999) “Qualité dans le sous-titrage: paramètres et implications”, in *Traduction-Transition-Translation.* Proceedings of the XV World Congress of FIT, Mons, 1999. Paris: FIT.

Gambier, Y. (2003) “Screen Transadaptation. Perception and Reception”. In *The Translator* 9 (2).

Gambier, Y. (2005a) “Screen translation : Subtitling”. In *The Encyclopedia of Languages and Linguistics* (2nd edition). Oxford: Elsevier.

Gambier, Y. (2005b) “Orientations de la recherche en traduction audiovisuelle”. In *Target* 17 (1)

Gambier, Y. (2006) “Le sous-titrage: une traduction sélective?”. In Tommola J. and Y. Gambier (eds.) *Translation and Interpreting. Training and Research*. Turku: University of Turku. Pp. 21-37

Gautier, G.L. (1981). “La traduction au cinéma. Nécessité et trahison”. *Image et son/Ecran. La revue du cinéma* 363.

Gee, J.P. (1991) An Introduction to Discourse Analysis. London: Routledge

Gerver, D. (1974). “The effects of noise on the performance of simultaneous interpreters: accuracy of performance”. In *Acta Psychologica* 38, pp. 159-167.

Gerver, D. (1976) “Empirical Studies of Simultaneous Interpretation: A Review and a Model”. In Brislin, R. W. (a cura di) *Translation. Applications and Research.* New York: Gardner Press.

Gile, D. (1985) “Le Modèle d’Efforts et l’équilibre d’interprétation en interprétation simultanée”. In *Meta 30* (1).

Gile, D. (1995) *Regards sur la recherche en interprétation de conférence*. Lille: Presses Universitaires de Lille.

Gotti, M. (1991) *I linguaggi specialistici: caratteristiche linguistiche e criteri pragmatici*. Firenze: La Nuova Italia.

Gotti, M. (2003) *Specialized discourse: linguistic features and changing conventions*. Berna: Peter Lang

Gotti, M. e M. Dossena (2001) *Modality in Specialised Texts. Selected Papers*. Berna: Peter Lang.

Gottlieb, H. (1991) *Tekstning. Synkron billedmedieoversaettelse*. Copenhague: Center for Oversaettelse (tesi di dottorato non pubblicata).

Gottlieb, H. (1992) "Subtitling. A new University discipline", in Dollerup C. e Loddegaard A. (a cura di) *Teaching Translation and Interpreting. Training, Talent and Experience*. Amsterdam e Philadelphia: J. Benjamins.

Gottlieb, H. (1994) "Subtitling: Diagonal translation", in *Perspectives* 2 (1).

Gottlieb, H. (1997) *Subtitles, translation and idioms*. Tesi di dottorato non pubblicata. Copenaghen: university of Copenhagen

Gottlieb, H. (2005) "Multidimensional Translation: Semantics turned Semiotics". *MuTra: challenges of multidimensional translation. Conference Proceedings*. http://www.euroconferences.info/proceedings/2005_Proceedings/2005_proceedings.html

Gran, L. (1998) "In-Training Development of Interpreting Strategies and Creativity". In Beylard-Ozeroff A., Kralová, J., Moser-Mercer, B. (a cura di) *Translators' Strategies and Creativity* – Selected Papers from the 9th International Conference on Translation and Interpreting, Prague, September 1995 Amsterdam e Philadelphia: John Benjamins.

Gran, L. (1999) "L'interpretazione simultanea: premesse di neurolinguistica". In Falbo, C., Russo, M., Straniero Sergio, F. (a cura di). *Interpretazione simultanea e consecutiva. Problemi teorici e metodologie didattiche.* Milano : Hoepli.

Greenberg, S. (1996) "Understanding speech understanding: Towards a unified theory of speech perception". In *ESCA Workshop on Auditory Basis of Speech Perception*. Keele, UK.

Greenberg, S. Kingsbury, B. E. D. (1997) "The Modulation Spectrogram: In pursuit of an invariant representation of speech". In *ICASSP, Vol. 3*. Monaco.

Gregory, S. e J. Sancho-Aldridge (1996) *Dial 888: subtitling for deaf children.* Londra: ITC.

Grice, H. P. (1957) "Meaning". In *The Philosophical Review, 66*. Pp. 377-88.

Grice, H. P. (1975) "Logic and conversation". In Cole, P. e Morgan, J. (a cura di) *Sintax and semantics III: speech acts.* New York: Academic press

Grice, H. P. (1978) "Further notes on logic and conversation". In Cole, P. e Morgan, J. (a cura di) *Sintax and semantics IV: speech acts.* New York: Academic press

Groner, R., D'Ydewalle, G. e R. Parham (1990). *From eye to mind: information acquisition in perception, search and reading*. Amsterdam : North-Holland.

Gullo, M. "L'orecchio bionico". Inserto *Salute*. *La Repubblica,* 16 giugno 2005. Pp. 16-17.

Gutt, E-A. (1991) "Translation as interlingual interpretive use". In Venuti, L. (a cura di). *The translation studies reader*. Londra e New York: Routledge.

Halliday, M. A. K. (2002) "On Grammar". In Webster, J. (a cura di) *The collected works of M. A. K. Halliday*. Volume 1. Londra e New York: Continuum.

Halliday, M. A. K. e R. Hasan (1976) *Cohesion in English*. London: Longman.

Halliday, M. A. K. e R. Hasan (1985) *Language, Context and Text: a Social Semiotic Perspective*. Oxford: Oxford University Press.

Harris, R. (1996) *Signs, language and communication*. Londra e New York: Routledge.

Hatim, B. e Mason, I. (1990) *Discourse and the translator*. Londra e New York: Longman.

Hatim, B. e Mason, I. (1990) *The translator as communicator*. Manchester: St. Jerome.

Heiss, C. e R. M. Bollettieri Bosinelli a cura di (1996) *Traduzione multimediale per il cinema, la televisione e la scena*. Atti del convegno internazionale. Forlì 26-28 ottobre 1995. Bologna : CLUEB.

Hempel, C. G. (1952) *Fundamentals of Concept Formation in Empirical Science*. Chicago: University of Chicago Press.

Herbert, J. (1952) *Le manuel de l'interprète*. Ginevra: Georg.

Hermans, T. (1999) *Translation in systems. Descriptive and system-oriented approaches explained*. Manchester: St. Jerome.

Hermansky, H. (1990) "Perceptual linear predictive (PLP) analysis of speech". In *Journal of the Acoustical Society of America*, vol. 87, no. 4.

Hermansky, H. Morgan, N. (1994) "RASTA processing of speech". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 2.

Higgs, C. (2006) "Subtitles for the deaf and the hard of hearing on TV". In Eugeni, C. e G. Mack (a cura di) *Proceedings of the first international conference on real-time intralingual subtitling*. InTRAlinea Special issue: respeaking www.intralinea.it

Hillier, H. (2004) *Analysing Real Texts*. Basingstoke: Palgrave Macmillan

Hindmarsh, R. (1985) *Language problems in European television. A feasibility study*. Manchester: St Jerome.

HMSO (2003) *Communications Act*, §303. http://www.opsi.gov.uk/acts/acts2003

Hodge R. e G. Kress (1993) *Language as Ideology*. Londra: Routledge.

Holman M. e Boase-Beier J. (1999) "Writing, rewriting and translation through constraint to creativity". In Boase-Beier J. e Holman M. (a cura di) *The practice of literary translation. Contraints and creativity*. Manchester: St Jerome.

Holmes, J. (1987) "The Name and Nature of Translation Studies". In *Indian Journal of Applied Linguistics 13* (2). Pp. 9-24.

HTKBook (2002) *Manuale del tool HTK*. Cambridge University Engineering Department.http://htk.eng.cam.ac.uk/

Hymes, D. (1974) *Foundations in Sociolinguistics – An Ethnographic Approach*. Londra: Tavistock Publications Ltd.

Ilg, G. (1959) "L'enseignement de l'interprétation à l'Ecole d'Interprètes de Genève", L'interprète N.1, Ginevra: Université de Genève.

Itakura, F. (1975) "Minimum prediction residual applied to speech recognition". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol.ASSP-23.

ITC (1999) *Guidance on standards for subtitling*. Londra: ITC. http://www.ofcom.org.uk/tv/ifi/guidance/tv_access_serv/archive/subtitling_stnds/itc_stnds_subtitling_word.doc

ITC (2001) *ITC code on subtitling, sign language and audiodescription*. Londra: ITC.

ITC (2002) The watershed and ITC family viewing policy. Londra: ITC.

Ivarsson, J. (1992) *Subtitling for the media: a handbook of an art.* Stockholm: Transedit.

Ivarsson, J. e M. Carroll (1998) *Subtitling*. Simrishamn: TransEdit HB.

Jackobson, R. (2000) "On linguistics aspects of translation". In Venuti, L. (a cura di). The translation studies reader. Londra e New York: Routledge.

Jelinek, F. (1997) *Statistical Methods for Speech Recognition.* Cambridge, MA: MIT Press.

Jelinek, F. "The development of an experimental discrete dictation recognizer". In *Proceedings IEEE*, vol. 73, Nov. 1985.

Jelinek, F. Bahl, L. R. Mercer, R. L. (1975) "Design of a linguistic statistical decoder for the recognition of continuous speech". In *IEEE Transactions of Information Theory*, vol. IT-21.

Jensema, C. (1998), "Viewer reaction to different television captioning speeds". In American Annals of the Deaf .Volume 143, No. 4.

Jensema, C. (1999) *Caption speed and viewer comprehension of television programs final report.* US Department of Education: Educational Resources Information Center.

Jensema, C., McCann, R. e S. Ramsey (1996). "Closed captioned television, presentation speed and vocabulary". In American Annals of the Deaf Volume 141, No.4.

Juang, B. H. (1996) "Automatic speech recognition: Problems, progress e prospects". In *IEEE Workshop Neural Networks for Signal Processing*, Kyoto.

Juang, B. H., Chou W., e C. H. Lee (1997) "Minimum classification error rate methods for speech recognition". In *IEEE Transactions Speech Audio Processing*, vol. 5.

Juang, B. H. e S. Katagiri (1992) "Discriminative training". In *Journal of the Acoustical Society of Japan* (E), vol. 13, no. 6.

Juang, B. H. e L. R. Rabiner, L. R. (1985) "A probabilistic distance measure for hidden Markov models". In *ATeT Technology Journal*, vol. 64.

Juang, B. H. (1985) "Maximum likelihood estimation for mixture multivariate stochastic observations of Markov chains". In *ATeT Technology Journal*, vol. 64, Morristown: Association for Computational Linguistics.

Juang, B.-H. e S. Furui (2000)"Automatic recognition and understanding of spoken language – A first step towards natural human-machine communication". In *Proceedings IEEE*, 88, 8.

Junqua, J.-C. Haton, J.-P. (1996) *Robustness in Automatic Speech Recognition.* Boston, MA: Kluwer.

Kalina, S. (1992) "Discourse Processing and Interpreting Strategies. An Approach to the Teaching of Interpreting". In Dollerup, C. e Loddegaard, A. (a cura di) *Teaching Translation and Interpreting – Training, Talent and Experience*. Papers from the First Language International Conference, Elsinore, Denmark, May 31st – June 2nd 1991 Amsterdam e Philadelphia: John Benjamins.

Kane, J. (1990) "Writing spoken English: the process of subtitling". In *Media information Australia* 56.

Karamitroglou, F. (1998) "A proposed set of subtitling standards in Europe". In *Translation Journal* 2(2)

Kawahara, T. Lee, C. H. e B. H. Juang (1997) "Combining key-phrase detection and subword based verification for flexible speech understanding". In *Proceedings IEEE ICASSP97*.

Kiesling S., Paulston C. (2005) *Intercultural Discourse and Communication*. Oxford: Blackwell.

Kintsch, W. e T. A. van Dijk (1978) "Toward a Model of Text Comprehension and Production". In *Psychological Review 85*.

Kirby, J. P. (1992) "On the Use of Strategies in Translation". In Lewandowska-Tomaszczyk, B. e M. Thelen (a cura di) *Translation and Meaning, Part 2.* Maastricht: Rijkshogeschool Maastricht – Faculty of Translation and Interpreting.

Kohn, K. e Kalina, S. (1996) "The Strategic Dimension of Interpreting". In *Meta 41* (1).

Kova, I. (1996) "Subtitling strategies: A flexible hierarchy of priorities". In Heiss, C. e R. M. Bollettieri Bosinelli (a cura di). *Traduzione multimediale per il cinema, la televisione e la scena*. Atti del convegno internazionale. Forlì 26-28 ottobre 1995. Bologna : CLUEB.

Kovačič, I. (1995) "Reception of subtitles. The non-existent ideal viewer". *Translatio* XIV (3-4).

Kovačič, I. (1996) "Reinforcing or changing norms in subtitling". In Dollerup C. e V. Appel (a cura di) *Teaching translation and interpeting 3*. Amsterdam e Philadelphia: John Benjamins.

Kovačič, I. (2000) "Thinking-aloud protocol. Interview. Text analysis". In Tirkkonen-Condit S. e R. M. Bosinelli (a cura di) *Traduzione multimediale per il cinema, la televisione e la scena*. Bologna: CLUEB.

Kurz I. (1995) "Interdisciplinary research - Difficulties and benefits". In *Target* 7:1. Pp. 165-179.

Kussmaul, P. (1995) *Training the translator*. Amsterdam e Philadelphia: John Benjamins.

Kutz, W. (1997) "Compression of the Source Message during Simultaneous Interpretation". In *La Traduzione. Saggi e documenti III. Quaderni di Libri e Riviste d'Italia 33*. Ministero per i beni culturali e ambientali.

Kyle, J. (1996) Switched on : deaf people's view on television subtitling previous reports www.deafstudiestrust.org.U.K.

Lacey, N. (2000) *Narrative and genre : key concepts in media studies*. Basingstoke: Macmillan.

Lakoff, G. (1972) "Hedges: A study of meaning criteria and the logic of fuzzy concepts". In Peranteau, P., Levi, J. e G. Phares (a cura di) *Papers from the Eighth Regional Meeting of Chicago Linguistic Society*. Chicago: Chicago University Press. Pp. 183-228.

Lambert, J. e D. Delabastita (1996). "La traduction de textes audiovisuels: modes en enjeux culturels". In Gambier Y. (a cura di) *Transferts linguistiques dans les médias audiovisuels*. Villeneuve d'Ascq: Septentrion.

Lambert, S. (1988) "A human information processing and cognitive approach to the training of simultaneous interpreters". In *Languages at crossroads: proceedings of the 29th annual conference of the American Translators Association*. Deanna Lindberg Hammond Ed. Medford, MJ: Learned Information Inc. Pp. 379-387.

Lambourne, A. (2006) "Subtitle respeaking". In Eugeni, C. e G. Mack (a cura di) *Proceedings of the first international conference on real-time intralingual subtitling*. InTRAlinea Special issue: respeaking www.intralinea.it

Lambourne, A. (2007) "Real time subtitling - extreme audiovisual translation". Intervento alla conferenza internazionale Multidimensional Translation: LSP Translation Scenarios, Vienna.

Laudanna, A. (1987) "Ordine dei segni nella frase". In Volterra, V. a cura di (1987). *La lingua italiana dei segni: la comunicazione visivo-gestuale dei sordi.* Bologna : Il Mulino.

Lawson, E. (1967). "Attention and Simultaneous Translation". In *Language and Speech* 10:1. Pp. 29-35.

Lederer, M. (1981) *La traduction simultanée – Experience et théorie.* Paris: Minard Lettres Modernes.

Lederer, M. (2003) "Le rôle de l'implicite dans la langue et le discours - les conséquences pour la traduction et l'interprétation". In *Forum 1*.

Lee, C. H. e Q. Huo (2000) "On adaptive decision rules and decision parameter adaptation for automatic speech recognition". In *Proceedings IEEE*, vol. 88.

Lee, T.-H. (2002) "Ear-Voice Span in English into Korean Simultaneous Interpretation". In *Meta 47* (4).

Lefevere, A. (1992) *Translation. Rewriting and the manipulation of literary frame.* Londra: Routledge.

Lepot-Froment, C. (1986) *Vivre Sourd: Communication et Surdité: aujourd'hui…et demain?* In Stocchero, I. a cura di (1994). *Dentro il segno*. Padova : CLEUP.

Lesser, V. R. Fennell, R. D. Erman, L. D. e D. R. Reddy (1975) "Organization of the hearsay—II: Speech understanding system". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. ASSP-23.

Lomheim, S. (1995) "L'écriture sur l'écran. Stratégies de sous-titrage à NRK". In Gambier Y. (a cura di) *Communication audiovisuelle et transferts linguistiques. Audiovisual communication and language transfer*. Numéro spécial de *Translatio/Nouvelles de la FIT/FIT Newsletter* 14 (3-4).

Lörscher, W. (1991) *Translation performance. Translation process and translation strategies. A psycholinguistic investigation*. Tübingen: Narr.

Loyola University Health System www.luhs.org/health/topics/ent/glossary.htm

Luyken, G.-M. *et al*. (1991) *Overcoming Language Barriers in Television: Dubbing and Subtitling for the European Audience*. Düsseldorf: EIM.

Lyons J. (1981) *Language and Linguistics: an Introduction*. Cambridge: Cambridge University Press

Mack, G. (2002) "New Perspectives and Challenges for Interpretation: The Example of Television". In Garzone, G. e M. Viezzi (a cura di) *Interpreting in the 21st Century - Challenges and Opportunities*. Selected Papers from the 1st Forlì Conference on Interpreting Studies, 9-11 November 2000 Amsterdam e Philadelphia: John Benjamins. Pp. 203-213.

Mack, G. (2006) "Detto scritto: un fenomeno, tanti nomi". In Eugeni, C. e G. Mack a (cura di) *Proceedings of the first international conference on real-time intralingual subtitling*. InTRAlinea Special issue: respeaking www.intralinea.it

Mackintosh, J. (1985) "The Kintsch and Van Dijk Model of Discourse Comprehension and Production Applied to the Interpretation Process". In *Meta* 30 (1).

Mellor, B. (1999) "Real-Time Speech Input for Subtitling". Aberystwyth: Aberystwyth University http://www.aber.ac.uk/mercator/images/barry.pdf

Manfredi, M. M. (1995) "Dall'Istituto all'esperienza delle Scuole Speciali in Emilia Romagna". In Porcari Li Destri, G. e V. Volterra (a cura di). *Passato e presente: uno sguardo sull'educazione dei sordi in Italia.* Napoli : Gnocchi.

Maragna, S. (2000) *La sordità : educazione, scuola, lavoro e integrazione sociale*. Milano : Hoepli.

Margareth, A. (2006) "Audiolesi: diktat nelle scuole?", inserto *Salute*. *La Repubblica*, 9/11.

Markel, J. D. e A. H. Gray Jr. (1976) *Linear Prediction of Speech*. Berlino: Springer-Verlag.

Marks, M. (2003) "A Distributed Live Subtitling System" In *BBC Research and Development White Papers*. www.bbc.co.uk

Marleau, L. (1980) "Les sous-titres…un mal nécessaire". In *Meta* 27 (3), 271-285.

Marsh, A. (2004). *Simultaneous interpreting and respeaking: a comparison*. Tesi di MA non pubblicata. University of Westminster.

Marsh, A. (2005). *Interview with Alison Marsh*. www.subtitleproject.net

Marsh, A. (2006) "Respeaking for the BBC". In Eugeni, C. Mack, G. (a cura di) *Proceedings of the First International Seminar on Real Time Intralingual Subtitling.* InTRAlinea, Special Issue on Respeaking. www.intralinea.it

Marshark, M. e M. D. Clark (1999) *Psychological perspectives on deafness*. New York: Lawrence Erbaum Association.

Martin, T. B. Nelson, A. L. e H. J. Zadell (1964) "Speech recognition by feature abstraction techniques". In *Air Force Avionics Lab, Tech. Rep. AL-TDR*.

Marzocchi, C. (1998) "The Case for an Institution-specific Component in Interpreting Research". In *The Interpreters' Newsletter 8*.

Mason, I. (1989) "Speaker meaning and reader meaning: preserving coherence in screen translating". In Kölmer, R. e J. Payne (a cura di) *Babel: the cultural and linguistic barriers between nations*. Aberdeen: Aberdeen university press.

Mason, I. (2000) "Audience design in translation". In *The translator 6 (1)*.

Massaro, D. W. (1970) "Preperceptual auditory images". In *Journal of Experimental Psycology*,85(3).

Massaro, D. W. (1972) "Perceptual images, processing time and perceptual units in auditory perception". In *Psychological Review*, 79 (2).

Mayoral, R., Kelly, D. e N. Gallardo (1988). "Concept of constrained translation. Non-linguistic perspectives of translation". In *Meta* 33 (3).

Mereghetti, E. (2006) "Le necessità dei sordi: TV e vita quotidiana". In Eugeni, C. and G. Mack (eds) *Proceedings of the First International Seminar on Real Time Intralingual Subtitling*. InTRAlinea, Special Issue on Respeaking. http://www.intralinea.it/

Merkle, D. a cura di (2002) "Censure et traduction dans le monde occidental. Censorship and translation in the Western world". In *TTR (Traduction, Terminologie, Rédaction)* 15 (2).

Miller, C. R. (1984) "Genre as social action". In *Quarterly Journal of Speech 70*. Pp. 151-67

Miller, L. G. e A. Gorin (1993) "Structured networks for adaptive language acquisition". In *International Journal of the Pattern Recognition and Artificial Intelligence* (Special Issue on Neural Networks), vol. 7, no. 4.

Minchinton, J. (1993) *Sub-titling*. Hertfordshire: Minchinton J.

Moore, R. C. (1997) "Using natural-language knowledge sources in speech recognition". In *Computational Models of Speech Pattern Processing.* Berlino: Springer-Verlag.

Moser-Mercer, B. (1996) "Koenraad Kuiper. 1996. Smooth talkers: The linguistic performance of auctioneers and sportscasters". In *Interpreting 1:2*

Myklebust, H. R. (1964) *The psychology of deafness*. New York: Grune and Stretton.

Nadas, A. Nahamoo, D. e M. A. Picheny, M. A. "On a model-robust training method for speech recognition". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 36.

Nadeu, C. Paches-Leal, P. e B. H. Juang (1997) "Filtering the time sequences of spectral parameters for speech recognition". In *Speech Communication*, vol. 22.

Nagata, K. Kato, Y. e S. Chiba (1963) "Spoken digit recognizer for Japanese language". In *NEC research and development*, no. 6.

Neale, Stephen (1980) "Genre". Londra: British Film Institute

Nencioni, G. (1983) *Di scritto e di parlato, discorsi linguistici*. Bologna : Zanichelli.

Nergaard, S. a cura di (1995) *Teorie contemporanee della traduzione*. Milano : Bompiani.

Neves, J. (2004) "Subtitling: written interpretation?". *Génesis – Revista de traduçao do ISAI 4*.

Neves, J. (2005) *Audiovisual translation. Subtitling for the Deaf and Hard-of-Hearing*. School of Arts, Roehampton University, University of Surrey.

Newmark, P. (1988) *Approaches to translation*. Hemel Hempsted: Prentice hall international.

Newmark, P. (1998) "Translation theory in the year 2000 amd its role in the translation schools". *Rivista internazionale di tecnica della traduzione 3*.

Ney, H. e S. Ortmanns (2000) "Progress in dynamic programming search for LVCSR". In *Proceedings IEEE*, vol. 88.

Nida, E. A. (1964) *Towards a Science of Translating*. Leiden: Brill.

Nord, C. (2000) "What do we know about the target-text receiver?". In Beeby, A. *et al*. (a cura di) *Investigating Translation*. Amsterdam: John Benjamins.

Norman, D. (1976) *Memory and attention*. New York: Wiley.

Norns, A.M. (1999) "For an abusive subtitling". In *Film Quarterly* 52 (3).

O'Connel, D. e S. Kowal (1994) "Some Current Transcription Systems for Spoken Discourse: A Critical Analysis". In *Pragmatics 4*.

Ofcom (2003) *ITC Guidance on Standards for Subtitling.* Londra: ITC. www.itc.org.uk

Oléron, P. e H. Nanpon (1964) "Recherches sur la Traduction Simultanée". In *Journal de Psychologie Normale et Pathologique*, 62. Pp. 73 - 94

Olson, H. F. e H. Belar (1956) "Phonetic typewriter". In *Journal of the Acoustical Society of America*, vol. 28, no. 6.

Orero, P. (2006) "Real-time subtitling in Spain". In Eugeni, C. e G. Mack (a cura di) Proceedings of the first international conference on real-time intralingual subtitling. InTRAlinea Special issue: respeaking www.intralinea.it

Orero, P. a cura di (2006) *Topics in audiovisual translation*. Amsterdam e Philadelphia: John Benjamins.

Palmer F.R. (2001) *Mood and Modality*. Cambridge: Cambridge University Press

Paneth, E. (1957) "An investigation into conference interpreting". In Pöchhacker, F. e M. Shlesinger a cura di (2002) *The interpreting studies reader*. Londra e New York: Routledge. Pp. 30-40.

Paradis, M. (1994) "Towards a neurolinguistic theory of simultaneous translation: the framework". In *International Journal of Psycholinguistics, 10* (3) [29]. Pp. 319-335.

Partington A. (1998) *Patterns and Meaning*. Amsterdam: John Benjamins.

Paul, P. (2001) *Language and deafness*. San Diego: Singular publishing group.

Perego, E. (2005). *La traduzione audiovisiva*. Roma: Carocci.

Petrillo, M. (1999) *APA an object oriented system for automatic prosodic analysis*. Tesi di dottorato non pubblicata. Università di Napoli Federico II.

Pieraccini, R. e E. Levin (1992) "Stochastic representation of semantic structure for speech understanding". In *Speech Communication*, vol. 11.

Pigliacampo, R. (1998). *Lingua e linguaggio del sordo: analisi e problemi di una lingua visivo-manuale.* Armando : Roma.

Pirelli, G. (2006) "Le necessità dei sordi: la sottotitolazione in tempo reale all'università". In Eugeni, C. e G. Mack (a cura di) Proceedings of the first international conference on real-time intralingual subtitling. InTRAlinea Special issue: respeaking www.intralinea.it

Pöchhacker, F. (1992) "The Role of Theory in Simultaneous Interpreting". In Dollerup, C e Loddegaard, A. (a cura di) *Teaching Translation and Interpreting – Training, Talent and Experience*. Papers from the First Language International Conference, Elsinore, Denmark, May 31st – June 2nd 1991. Amsterdam e Philadelphia: John Benjamins.

Pöchhacker, F. (2002) "Researching interpreting quality: models and methods". In Garzone, G. e M. Viezzi (a cura di) *Interpreting in the 21st Century - Challenges and Opportunities*. Selected Papers from the 1st Forlì Conference on Interpreting Studies, 9-11 November 2000 Amsterdam e Philadelphia: John Benjamins. Pp. 95-106.

Pöchhacker, F. (2004) *Introducing Interpreting Studies.* Londra e New York: Routledge.

Porcari Li Destri, G. e Volterra, V. a cura di (1995). *Passato e presente: uno sguardo sull'educazione dei sordi in Italia.* Napoli : Gnocchi.

Poyatos, F. a cura di (1997) *Non-verbal communication and translation. New perspectives and challenges in literature, Interpretation and the media*. Amsterdam e Philadelphia: John Benjamins.

Praet, C., Verfaillie, K., De Graef, P., Van Rensbergen, J. e G. D'Ydewalle, G. (1990). "A one line text is not half a two line text". In Groner, R., D'Ydewalle, G. e R. Parham. (1990). *From eye to mind: information acquisition in perception, search and reading*. Amsterdam : North-Holland.

Prandi, M. (2004) "Riformulazione e Condivisione". In *Rassegna Italiana di Linguistica Applicata* n.1, Bulzoni editore.

Quigley, S. P. e P. V. Paul (1984) *Language and deafness.* San Diego, CA : College-Hill Press.

Rabiner, L. R. (1989) "A tutorial on hidden Markov models and selected applications in speech recognition". In *Proceedings IEEE*, vol. 77.

Rabiner, L. R. Juang, B. H. (1993) *Fundamentals of Speech Recognition.* Englewood Cliffs, NJ: Prentice-Hall.

Rabiner, L. R. Levinson, S. E. Rosenberg, A. E. Wilpon, J. G. (1979) "Speaker independent recognition of isolated words using clustering techniques". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol.ASSP-27.

Radutzky, E. (1995). "Cenni storici sull'educazione dei sordi in Italia dall'antichità alla fine del Settecento". In Porcari Li Destri, G. e Volterra, V. a cura di (1995). *Passato e presente: uno sguardo sull'educazione dei sordi in Italia.* Napoli : Gnocchi.

RAI, a cura di (2002). *Scripta Volant – La Rai per i sordi*. Roma : Rai-Eri.

Rampelli, S. (1990). *Aspetti linguistici dei sottotitoli e fruibilità da parte delle persone sorde. Prospettive teoriche e ambiti applicativi.* Tesi non pubblicata, Università "La Sapienza" di Roma, Facoltà di Filosofia A/A 1988-1989.

RCQ (1983) La loi sur le cinéma, c. 37, a. 83; 1991, c. 21, a. 14. http://www.rcq.gouv.qc.ca/la_regie/classement.asp#83

Reason, P. e H. Bradbury (2001) *The handbook of Action Research: participative inquiry and practice.* Londra : Sage

Reddy, D. R. (1966) *An approach to computer speech recognition by direct analysis of the speech wave*. Computer Science Department, Stanford University, Tech. Rep. C549.

Ree, J. (1999) *I see a voice : a philosophical history of language, deafness and the senses.* London : Collins.

Reid, H. (1978) "Subtitling, the intelligent solution". In Horguelin, P.A (a cura di.) *La traduction: une profession. Translating, a profession.* Proceedings of the VIII World Congress of FIT, Montréal 1977. Ottawa: Council of Translators and Interpreters of Canada.

Reiss, K. (1971) *Möglichkeiten und Grenzen der Übersetzungskritik: Kategorien und Kriteren für eine sachgerechte Beurteilung von Übersetzungen*. Monaco: Max Hueber.

Reiss, K. (1983) *Texttyp und Übersetzungs methode. Der operative Text*. Heidelberg: Julius Gross.

Remael, A. (2007) *Sampling subtitling for the Deaf and the Hard of Hearing in Europe*. In Díaz-Cintas, J., Orero, P. e A. Remael (2007) *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language*. Amsterdam: Rodopi.

Remael, A. e B. van der Veer (2006) "Real-Time Subtitling in Flanders: Needs and Teaching". In Eugeni, C. e G. Mack 2006 (a cura di). *Proceedings of the first international conference on real-time intralingual subtitling*. InTRAlinea Special issue: respeaking www.intralinea.it

Riccardi, A. (1996) "Language-specific Strategies in Simultaneous Interpreting". In Dollerup, C. e Appel, V. (a cura di) *Teaching Translation and Interpreting 3. New Horizons* – Papers from the Third Language International Conference. Elsinore, Denmark 9th-11th June 1995. Amsterdam e Philadelphia: John Benjamins.

Riccardi, A. (1998) "Interpreting Strategies and Creativity". In Beylard-Ozeroff A., Kralová, J., Moser-Mercer, B. (a cura di) *Translators' Strategies and Creativity* – Selected Papers from the 9th International Conference on Translation and Interpreting, Prague, September 1995 Amsterdam e Philadelphia: John Benjamins.

Riccardi, A. (1999) "Interpretazione simultanea: Strategie generali e specifiche". In Falbo, C., Russo, M. e Straniero Sergio, F. (a cura di) *Interpretazione Simultanea e Consecutiva – Problemi teorici e metodologie didattiche.* Milano: Hoepli.

Riccardi, A. (2003) "The Relevance of Interpreting Strategies for Defining Quality in Simultaneous Interpreting". In Collados Aís, A., Fernández Sánchez, M. M. e D. Gile (a cura di) *La evaluación de la calidad en interpretación: investigación* - Actas del I Congreso Internacional sobre Evaluación de la Calidad en Interpretación de Conferencias: Almuñecar, 2001 Granada: Comares.

Robson, G. D. (1997) *Inside Captioning*. CyberDawg Publishing.

Robson, G. D. (2004). *The closed captioning handbook*. Oxford: Elsevier.

Rodda, M. e Grove, C. (1987) Language, cognition and deafness. Londra: Lawrence Arlbaum Associates.

Ross, D. (1997) “La struttura linguistica e l’elaborazione sintattica: strategie generali e specifiche”. In Gran, L. e Riccardi, A. (a cura di) *Nuovi orientamenti negli studi sull’interpretazione*. Giornata di studi, 19 Aprile 1996 Padova: C.L.E.U.P.

Rubbi, M. (2003) “Sottotitoli in Europa : I teleAspettatori”. In *HP/Accaparlante n.3*. Centro Documentazione Handicap: Bologna.

Rundle, C. (2008)

Russo, M. (1999) “La Conferenza come Evento Comunicativo”. In Falbo, C. Russo, M. e F. Straniero Sergio (a cura di) *Interpretazione Simultanea e Consecutiva. Problemi Teorici e Metodologie Didattiche*, Milano: Editore Ulrico Hoepli. Pp. 89-102.

Sacks, O. (1990). *Vedere voci. Un viaggio nel mondo dei sordi*. Milano : Adelphi.

Safar 1992

Safar, H. (2006) Intervento alla conferenza internazionale Languages and the Media. Berlino: Hotel InterContinentale.

Sakai, T. e S. Doshita (1962) “The phonetic typewriter, information processing”. In *Proceedings IFIP Congr*ess. Monaco.

Sakoe, H. e S. Chiba (1978) “Dynamic programming algorithm optimization for spoken word recognition”. In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. ASSP-26.

Sancho-Aldridge, J. e IFF Research ltd. (1996) *Good news for deaf people: subtitling of national news programmes*. Londra: ITC.

Sasso, C. “Sordomuti, la guerra dei segni”. *La Repubblica*, 28 marzo 2007.

Savino, M. Refice, M. e L. Cerrato (1999) “Individuazione di correlati acustici per la classificazione di intenzioni comunicative nell’interazione uomo-macchina”. In *Atti del convegno AI*IA,* Genoa.

Schwarz, B. (2003). “Translation in a confined space”. In *Translation Journal* 6 (4) e 7 (1).

Schweda-Nicholson, N. (1990) “The role of shadowing in interpreter training”. In *The Interpreters’ Newsletter, n. 3.*

Schweda-Nicholson, N. (1991) “Self-Monitoring Strategies in Simultaneous Interpretation”. In Picken, C (a cura di) *ITI Conference 5 – Windows on the World* – Proceedings of the 5th Conference of the Institute of Translation and Interpreting 25th-26th April 1991 at the Hotel Russell, Russell Square, London WC1 Londra: Aslib.

Searle, J. R. (1969) Speech acts : an essay in the philosophy of language. Cambridge: Cambridge university press.

Seleskovitch, D. (1968) *L’interprète dans les conferences internationales.* Paris: Minard, Lettres Modernes.

Seleskovitch, D. (1982) “Impromptu Speech and Oral Translation”. In Enkvist, N. E. (a cura di) *Impromptu Speech: A Symposium.* Åbo: Åbo Akademi.

Seleskovitch, D. e M. Lederer (1989) *Pédagogie raisonnée de l’interprétation.* Paris : Didier Erudition.

Setton, R. (1999) *Simultaneous Interpretation - A Cognitive-Pragmatic Analysis.* Amsterdam e Philadelphia: John Benjamins.

Shire, M. L. (1997). *Syllable onsets detection from acoustics.* Teso di MA non pubblicata. UC Berkeley.

Shlesinger, M. (1999) “Norms, Strategies and Constraints: How Do We Tell Them Apart?”. In Alvarez Lugris, A. e A. Fernandez Ocampo (a cura di) *Anovar Estudios de Traducción e Interpretación* vol. I. Vigo: Universidade de Vigo.

Silver, J. *et al.* (2000). *A new font for digital television subtitles*. www.tiresias.org

Snell-Hornby, M. a cura di (1994). *Translation studies – an integrated approach*. Amsterdam e Philadelphia: John Benjamins.

Stallard, G. (2003) *Final report to CENELEC on TV for All. Standardization requirements for access to digital TV and interactive services by disabled people*. www.cenelec.org

Stam, R. (2000) *Film Theory*. Oxford: Blackwell.

Steiner, E. H. e R. Veltman a cura di (1988) *Pragmatics, discourse and text*. Londra: Pinter.

Stocchero, I. (1994) *Dentro il segno*. Padova : CLEUP.

Stokoe, W. (1981) “Linguaggio, segnato e parlato”. In Volterra, V. (a cura di) *I segni come parole: la comunicazione dei sordi.* Torino : Boringhieri.

Sunnari, M. (1995a) “Processing Strategies in Simultaneous Interpreting: Experts vs. Novices”. In Krawutschke, P. W. (a cura di) *Connections.* Proceedings of the 36th annual conference of ATA, Nashville (Tennessee), November 8th –12th 1995, Information Today Inc: Medford, N.J.

Sunnari, M. (1995b) “Processing Strategies. In Tommola, J. (a cura di) *Topics in Interpreting Research.* Turku: University of Turku – Centre for Translation and Interpreting.

Suzuki, J. Nakata, K. (1961)“Recognition of Japanese vowels—Preliminary to the recognition of speech”. *Journal of the British Institute of Radio Engineers*, vol. 37, no. 8.

Swales J. M. (1990) *Genre Analysis. English in Academic and Research Settings*. Cambridge: Cambridge University Press

Taylor, C. (1998) “In defence of the word: Subtitles as conveyors of meaning and guardians of culture”. In Bollettieri Bosinelli, R. M. *et al.* a cura di (2000). *La traduzione multimediale: quale traduzione per quale testo?* Atti del convegno internazionale Multimedia translation: which translation for which text? Forlì, 2-4 aprile 1998. Bologna : CLUEB.

Tirkkonen-Condit, S. e R. Jääskeläinen a cura di (2000) *Tapping and mapping the processes of translation and interpreting*. Amsterdam e Philadelphia: John Benjamins.

Titford, C. (1982) “Sub-titling – Constrained Translation”. In *Lebende Sprachen* 27 (3).

Tommola *et al.* (2001) “Images of shadowing and interpreting”. In *Interpreting* vol. 5, n. 2. Pp. 147-169

Toury, G. (1986) “Translation”. In Sebeok, T. (a cura di) *Encyclopedic dictionary of semiotics. Volume 2.* Berlino: Mouton de Gruyter.

Toury, G. (1995) *Descriptive Translation Studies and Beyond*. Amsterdam e Philadelphia: John Benjamins.

Trivulzio, G. (2000) *Da Tirone al Riconoscimento del parlato*. Milano: Asfor.

Trivulzio, G. (2006) “Natura non facit saltus”. In Eugeni, C. e G. Mack (a cura di) Proceedings of the first international conference on real-time intralingual subtitling. InTRAlinea Special issue: respeaking www.intralinea.it

Trosborg A. a cura di (2000) *Analysing Professional Genres*. Amsterdam: John Benjamins.

Tudor, A. (1974): Image and Influence: Studies in the Sociology of Film. London: George Allen & Unwin.

Tveit, J.E. 2004. *Translating for television*. Oslo: Kolofon AS.

Van Basien, F. (1999) "Anticipation in Simultaneous Interpretation". In *Meta 44* (2).

van Dam, I. M. (1989) "Strategies of Simultaneous Interpretation". In Gran, L. e J. Dodds (a cura di) *The theoretical and practical aspects of teaching conference interpretation.* Udine: Campanotto.

van der Veer, B. (2007) "De tolk als respeaker: een kwestie van training". In *Linguistica Antverpiensia, LA NS6*. Pp. 315-328.

van Dijk, T. A. (1984) "Strategic Discourse Comprehension". In Coveri, L., Beretta. L. *et. al.* a cura di (1984). *Linguistica Testuale*: Atti del 15° Congresso internazionale di studi. Genova – Santa Margherita Ligure, 8-10 Maggio 1981 SLI, Società di Linguistica Italiana. Roma: Bulzoni.

van Djik, T. a cura di (1997) Discourse as Social Interaction. London: Sage

van Dijk, T. A. e W. Kintsch (1983) *Strategies of Discourse Comprehension.* Orlando: Academic Press.

Venuti, L. (1995) *The translator's invisibility*. Londra e New York : Routledge.

Venuti, L. a cura di (1992) *Rethinking translation: discourse, subjectivity, ideology*. Londra e New York: Routledge.

Verdirosi, M. L. (1987) "Luoghi". In Volterra, V. (a cura di) *La lingua italiana dei segni: la comunicazione visivo-gestuale dei sordi.* Bologna : Il Mulino

Verlinde, R. e P. Schragle (1986) *How to Write and Caption for Deaf People*. Silver Spring, MD: T.J. Publishers.

Vermeer, H. J. (1989) "Skopos and commission in translation action". In Venuti, L. a cura di (2000) *The translation studies reader.* Londra e New York: Routledge.

Viaggio S. (1992) "Teaching Beginners to Shut up and Listen". In *The Interpreters' Newsletter 4*.

Viezzi, M. (1999) "Interpretazione Simultanea: Attività specifica per coppie di lingue?". In *Settentrione 11*.

Viezzi, M. (2001) "Interpretazione e comunicazione politica". In Garzone, G. e M. Viezzi (a cura di). *Comunicazione specialistica e interpretazione di conferenza*. Trieste: Edizioni Università di Trieste.

Viezzi, M. (2002) "La non-selezione del congiuntivo quale opzione strategica nell'interpretazione simultanea dall'inglese in italiano". In Schena, L., Prandi, M. e M. Mazzoleni (a cura di) *Intorno al congiuntivo*. Atti del convegno di studi, Forlì, 2-3 Marzo 2000 Bologna: CLUEB

Vinay J. P. e J. Darbelnet (1958). *Stylistique comparée du français et de l'anglais*. Paris: Didier e Montréal: Beauchemin .

Vintsyuk, T. K. (1968) "Speech discrimination by dynamic programming". In *Kibernetika*, vol. 4.

Volterra, V. (1986) *Il linguaggio dei sottotitoli e gli spettatori sordi*. Roma, RAI-Televideo, n. 5.

Volterra, V. (1988). *Le reazioni di un campione di non udenti nei confronti di due diversi tipi di sottotitolatura adottati nei programmi televisivi*, ricerca n. 870901 RAI Televideo.

Volterra, V. a cura di (1981). *I segni come parole: la comunicazione dei sordi.* Torino : Boringhieri.

Volterra, V. a cura di (1987). *La lingua italiana dei segni: la comunicazione visivo-gestuale dei sordi.* Bologna : Il Mulino

Volterra, V., Romeo, C., Zargoni, A. e L. Tucci (1987) *Transferring information from auditory to visual mode, TV subtitles for deaf viewers*, RAI Televideo.

W3C (2003) Timed Text Working Group http://www.w3.org/AudioVideo/TT/

Wales, K. (1989) *A Dictionary of Stylistics*. Londra: Longman

Wang, H. C. Chen, M.-S. e T. Yang (1993) "A novel approach to the speaker identification over telephone networks". In *Proceedings ICASSP-93*. Minneapolis, MN, vol. 2.

Wilpon, J. G. e L. R. Rabiner (1985) "A modified K-means clustering algorithm for use in isolated word recognition". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. ASSP-33.

Wilpon, J. G. Rabiner, L. R. Lee, C. H. e E. Goldman (1990) "Automatic recognition of keywords in unconstrained speech using hidden Markov models". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 38.

Zanetti, R. (1999) "Relevance of Anticipation and Possible Strategies in the Simultaneous Interpretation from English into Italian". In *The Interpreters' Newsletter 9*.

Zatini, F. (1995) "Storia degli istituti per sordomuti in Italia". In Porcari Li Destri, G. e Volterra, V. a cura di (1995). *Passato e presente: uno sguardo sull'educazione dei sordi in Italia.* Napoli : Gnocchi.

Allegato – Trascrizione di BBC News del 5 luglio 2005, ore 10.15

1	It's quarter past ten.	(…)
2	Let's update you on with the headlines this morning:	(…) The headlines this morning:
3	Scores of anti-capitalist demonstrators are due to appear in court following yesterday's violent protest in Edinburgh against the G8 summit.	Scores of anti-capitalist demonstrators are due to appear in court following yesterday's violent protests in Edinburgh against the G8 summit.
4	It's the final day of lobbying in Singapore for the five cities hoping to host the 2012 Olympics.	It's the final day of lobbying in Singapore for the (…) cities hoping to host the 2012 Olympics.
5	And a man suspected of shooting a police officer	(…) A man (…)
6	has been found dead at a house in Lancashire.	has been found dead at a house in Lancashire.
7	In sport, Steven Gerrard transfer saga's certainly hotting up.	In sport, Steven Gerrard transfer saga(…) certainly hotting up.
8	Liverpool have received a £32 offer from Chelsea for the star,	Liv were received a £32 offer (…) for the star,
9	which they say they will turn down.	which they say they will turn down.
10	West Ham have been busy on the transfer of Cardiff defender Danny Gabbidon, one of three players to join West Ham today.	(…) Cardiff defender Danny Gabbidon, one of three players to join West Ham today.
11	And the Lions have beaten Auckland, but only just.	And the Lions have beaten Auckland, but only just.
12	They won their match this morning by 17 points to 13.	They won their match this morning by 17 points to 13.
13	Full details in half an hour.	Full details in half an hour.
14	Thank you Sue.	(…)
15	Allegations that Iran's newly elected president was a 1979 hostage-taker have been strongly rejected.	Allegations that Iran's newly elected president was a 1979 hostage-taker have been strongly rejected.
16	The authorities say the allegations about President Mahmoud Ahmedinejad are false	The authorities say the allegations about President Mahmoud Ahmedinejad are false

	and part of a smear campaign.	and part of a smear campaign.
17	As the BBC's Francis Harrison reports from the capital Teheran, it's feared the claims could increase support for the ultra-conservative	As the BBC's France's Harrison reports from the capital Teheran, it's feared the claims could increase support for the ultra-conservative
18	and help further isolate Iran.	and help further isolate Iran.
19	Radical Iranian students seizing the American embassy a quarter of a century ago.	Radical Iranian students seizing the American embassy a quarter of a century ago.
20	Six of the Americans held hostage now say they remember Iran's new president as one of their captors.	Six of the Americans held hostage now say they remember Iran's new president as one of their captors.
21	The controversy started when an opposition website published this photograph,	The controversy started when an opposition website published this photograph,
22	alleging the bearded man was Mr (...) Ahmedinejad (...).	alleging the man was Mr Mahmoud Ahmedinejad ack.
23	The US State Department now says it was not him.	The US Department says it was not him.
24	And (…) hostage takers, (…) are also adamant he was not one of them.	(…) The hostage takers are also adamant he was not one of them.
25	(many of whom are politically opposed to Mr Ahmedinejad,)	(…)
26	TRANSLATION: Another name I might not have been so sure,	TRANSLATION: (Yellow) Another name I might not have been so sure,
27	but his name I am 100% positive about.	but (…) am a hundred% positive about.
28	He could not have been there.	He could not have been there.
29	Even if 40 hostages claim it was Ahmedinejad I know it definitely could not have been him.	(…) I know it definitely could not have been him.
30	Either they're lying or they're mistaken.	Either they're lying or (…) mistaken.
31	Today, the former American embassy is known as the den of spies.	(White) Today, the former American embassy is known as the den of spies.
32	America is demonised in the same way that some now fear Iran's new president is being demonised abroad.	America is demonised in the same way some now fear Iran's new president is being demonised abroad.
33	The concern is this could create a more isolated and radical Iran.	The concern is this could create a more isolated and radical Iran.

34	Were Iran's new president one of the students who climbed over the wall of the American embassy here, it's likely he would wear it as a badge of honour.	Were Iran's new president one of the students who climbed over the wall of the American embassy here, it's unlikely he would wear it as a badge of honour.
35	As one headline of a newspaper put it: It would increase his popularity in Iran.	(…) It would increase his popularity in Iran.
36	Abroad the hostage takers are divided as terrorists.	Abroad the hostage takers are divided as terrorists.
37	Here it's a different story.	Here it's a different story.
38	TRANSLATION: I don't know if Mr (...) Ahmedinejad was one of the students here,	TRANSLATION: (Yellow) I don't know if Mr Mahmoud Ahmedinejad was one of the students here,
39	but if he was, we are very thankful to him	but if he was, we are very thankful to him
40	and like him even more.	and like him even more.
41	TRANSLATION: The Americans just want to paint a negative picture of our new president	(White) TRANSLATION: (Yellow) The Americans just want to paint a negative picture of our new president
42	and show that Islam is violent	and show that Islam is violent
43	and Iranians are terrorists.	and Iranians are terrorists.
44	That's because they said nobody would go out and vote in the elections,	That's because they said nobody would go out and vote in the elections,
45	but when we did, it was a slap in the face for them.	but when we did, it was a slap in the face for them.
46	Every year, the anniversary of the hostage taking is celebrated,	(White) Every year, the anniversary of the hostage taking is celebrated,
47	a reminder of how Iran humiliated the world superpower.	a reminder of how Iran humiliated the world's superpower.
48	Before he's even taken office, Iran's new president is being pitted against the Americans	Before he's even taken office, Iran's new president is being pitted against the Americans
49	and branded as an extremist.	and branded as an extremist.
50	The worry is that he will only make relations with the outside world more difficult.	The worry is that (…) will (…) make relations with the outside world more difficult.
51	Francis Harrison, BBC News, Teheran.	(…)

52	A massive relief operation is under way in India and Pakistan	A massive relief operation is under way in India and Pakistan
53	after flooding caused by monsoon rain left tens of thousands of people homeless.	after flooding caused by (…) rain left tens of thousands of people homeless.
54	One of the worse-hit areas is Gujarat, in West India,	One of the worse-hit areas is Gujarat (…),
55	where more than 7000 villages are cut off by the floodwater.	where (…) villages are cut off by the floodwater.
56	Relief for the flood victims in Gujarat.	Relief for the flood victims in Gujarat.
57	Indian air force dropped supplies to villagers marooned by the rising water.	The (…) air force dropped supplies to villagers marooned by the rising water.
58	The monsoon rains hit at the flood weekend, submerging huge areas.	The Monday soon rains hit at the weekend, submerging (…) areas.
59	The worst affected are thousands of tiny villages	Thousands of tiny villages (…) are affected.
60	where people have lost their homes and their livelihoods.	(…)
61	The leader of the Congress Party, Sonia Gandhi, visited the flooded area	The leader of the congress party (…) visited the flooded area
62	and promised aid.	and promised aid.
63	Gujarat is one of India's richest states,	Gujarat is one of India's richest states,
64	but the flooding has caused chaos, with everyday life virtually brought to a standstill.	but the flooding has caused chaos, with every (…) day life (…) brought to a stand still.
65	The monsoon rains (…) also swept into neighbouring Pakistan,	The (…) rains have (…) swept into neighbouring Pakistan,
66	causing the worst flooding in more than a decade.	causing the worst flooding in more than a decade.
67	The army has been distributing aid to tens of thousands of people left homeless in the north-west of the country.	The army has been distributing aid to tens of thousands of people left homeless in the north-west of the country.
68	Thousands of homes are reported to have been flooded,	Thousands of homes are reported to have been flooded,
69	but so far, there have been no casualties.	but so far, there have been no casualties.

70	William Forrest, BBC News.	(…)
71	A permanent ban on a group of chemicals (…) has been voted on by the European Parliament today.	A permanent ban on a group of chemicals (…) has been voted on by the European Parliament today.
72	(which are used to soften children's toys)	(which are used to soften children's toys)
73	The chemicals, (…) have been under a temporary ban since 1999.	The chemicals, (…) have been under a temporary ban since 1999.
74	(known as phthalates,)	(known as phthalates,)
75	Scientific studies have found links to an increased risk of asthma and allergies,	(…) Studies have found links to an increased risk of asthma and allergies,
76	as well as damage to the reproductive system.	as well as damage to the reproductive system.
77	In this tender age, anything that can be grabbed goes straight into the mouth.	(…) This tender age, anything that can be grabbed goes straight into the mouth.
78	While that might be good for teething,	While that might be good for teething,
79	scientists worry it may be harmful for children's health.	scientists worry it may be harmful for children's health.
80	Many chewy and rubbery toys contain special softening chemicals called phthalates,	Many chewy and rubbery toys contain (…) phthalates,
82	which laboratory tests have linked to allergies such as asthma and even certain types of cancer.	which (…) tests have linked to allergies such as asthma and even certain types of cancer.
83	Since 1999, phthalates have been under a temporary ban in the European Union.	Since 1999, phthalates have been under a temporary ban in the European Union.
84	Now the European Parliament is debating whether to make that ban permanent.	Now the European Parliament is debating whether to make that ban permanent.
85	Toy manufacturers have criticised the move,	(Line) Toy manufacturers have criticised the move,
86	claiming the research has been contradictory.	claiming the research has been contradictory.
87	MEPs insist that where children's health is concerned, all precautions must be taken.	MEPs insist that where children's health is concerned, all precautions must be taken.
88	Nobody is going to be killed by chewing them,	(Yellow) Nobody is going to be killed by chewing them,

89	but potentially, there is a carcinogenic threat,	but potentially, there is a carcinogenic threat,
90	and that's why, under the precautionary principle, with scientific advice, the European Parliament takes action,	and that's why, under the precautionary principle, with scientific advice, (…)
91	and we'll be waiting of course for the scientists to produce their conclusions, and that's now before us.	(…) waiting (…) for the scientists to produce their conclusions, and that's now before us.
92	The supporters of this ban here at the European Parliament have been pushing for this legislation for the last five years.	(White) the supporters of this ban here at the European parliament have been pushing for this legislation for the last five years.
93	It does now look as though they will be successful.	It (…) looks as though they will be successful.
94	But some MEPs want the European Commission to go even further.	Some MEPs want the European Commission to go even further.
95	They're calling for tests to be carried out on other types of materials which contain phthalates	They're calling for tests to be carried out on other types (…) materials which contain phthalates.
96	including bed coverings and food packaging.	(…)
97	??? BBC news, Strasbourg.	(…)
98	China has more than four million people who are addicted to computer games on the Internet.	China has more than four million people who are addicted to computer games on the Internet.
99	Official statistics suggest most of them are teenagers	Official statistics suggest most of them are teenagers
100	who are hooked on chat rooms and online gaming.	who are hooked on chat rooms and online gaming.
101	The Government has been an enthusiastic supporter of Internet for business and education,	The Government has been an enthusiastic supporter of Internet for business (…),
102	but says Internet cafes are eroding public morality.	but says Internet cafes are eroding public morality.
103	Now Beijing is opening its first clinic to help treat the addicts (...).	(…) The first clinic to (…) treat the addicts has opened.
104	From the Chinese capital here is our correspondent Daniel Griffiths	(…)
105	China has a new addiction: the Internet.	China has a new addiction: the Internet.

106	And for those who can't tear themselves away from computer games and chatrooms, this is where to kick the habit.	(…) For those who can't tear themselves away from computer games and chatrooms, it is where to kick the habit.
107	TRANSLATION: I can find myself again in computer games.	TRANSLATION: (Yellow) I can find myself again in computer games.
108	In real life, you are nothing but a small potato	In real life, you are nothing but a small potato
109	but in computer games you can be a Superman.	(…) in computer games you can be a Superman.
110	I want to be a Superman,	I want to be a Superman,
111	so I have to play, play, (...) play.	so I have to play, play, and play.
112	The clinic only has room for ten people,	(White) The clinic only has room for ten people,
113	but they get first-class care.	but they get first-class care.
114	TRANSLATION: We do all kinds of treatment.	TRANSLATION: (Yellow) We do all kinds of treatment.
115	One big part is using drugs.	One big part is using drugs.
116	We also do psychological consulting	We also do psychological consulting
117	and let the kids play sports.	and let the kids play sports.
118	So it's a comprehensive way of treating patients.	So it's a comprehensive way of treating patients.
119	Most of the time this computer addiction is a psychological issue,	Most of the time the computer addiction is a psychological issue,
120	so we figure out what kind of problem it is	so we figure out what kind of problem it is
121	and use the right way to deal with it.	and use the right way to deal with it.
122	Each patient stays just two weeks.	(White) Each patient stays just two weeks.
123	Then it's back to normal life and the real test for these computer obsessives.	Then it's back to (…) life and the real test for these computer obsessives.
124	China has fallen in love with the Net.	China has fallen in love with the Netherlands.
125	There's a hundred million web surfers here already, and many more to follow.	There's a hundred million web surfers here already, and many more to follow.

126	The country's first Internet clinic is going to have its hands full.	The country's first Internet clinic is going to have its hands full.
127	Daniel Griffiths, BBC News, Beijing.	(…)
128	Now, let's just tell you what's coming up in the next few minutes or so.	(…) Let's (…) tell you what's coming up (…).
129	At about 10.30 we are expecting a news conference to come out of Singapore.	At about 10.30 we are expecting a news conference to come out of Singapore.
130	This is the British bid team lining up to try and make another bid to win the 2012 Olympics.	This is the British bid team lining up to try and make another bid to win the 2012 Olympics
131	There's the scene now.	ment there's the scene now.
132	The journalists are all gathered there.	The journalists are all gathered there.
133	We are expecting a few high-profile team members there.	We are expecting a few high-profile team members there.
134	David Beckham amongst them, Matthew Pinsent, Steve Redgrave, Tanni Grey Thompson it's going to be quite a line-up.	David Beckham amongst them, (…) it's going to be quite a line-up.
135	So, we're taking up live in quite a few minutes time.	(…)
136	An earthquake (…) has struck the Indonesian island of Sumatra,	An earthquake (…) has struck the Indonesian island of Sumatra.
137	which was badly hit by the tsunami.	(…)
138	(measuring 6.8 on the Richter scale)	(measuring 6.8 on the Richter scale)
139	Tall buildings were shaken	Tall buildings were shaken
140	and people rushed out of their homes,	and people rushed out of their homes,
141	but no immediate reports of serious damages or injuries.	but no immediate reports of serious damages or injuries.
142	The quake was also felt on Nias island.	The quake was also felt on Nias island.
143	The Canadian sex killer has been freed after 12 years behind bars	The Canadian sex killer has been freed after 12 years behind bars
144	for her involvement in the rape, torture, and murder of three teenage(...) girls.	for her involvement in the rape, torture, and murder of three teenager girls.
145	Carla Homolka, (…) was released from a	Carla Hamalka (…) was released (…)

	prison in Montreal	
146	after serving a reduced sentence in return for testifying against her ex-husband.	after serving a reduced sentence in return for testifying against her ex-husband.
147	(who is 35,)	(…)
148	She's been described as the most reviled woman in Canada,	She's been described as the most reviled woman in Canada,
149	but today, Carla Homolka was freed.	but today, Carla Hamalka was freed.
150	She served 12 years in connection with the kidnapping, rape, sexual torture and murder of three teenagers.	She served 12 years in connection with the kidnapping, rape, sexual torture and murder of three teenagers.
151	One of the victims was her 15-year-old sister.	One of the victims was her 15-year-old sister.
152	She choked to death after being drugged and sexually violated by both (...) Homolka and her husband.	She choked to death after being drugged and sexually violated by both the Hamalka and her husband.
153	The two were charged together,	The two were charged together,
154	but Homolka made a deal with the prosecutors.	but Homolka made a deal with (…) prosecutors.
155	She testified against her husband	(…)
156	who's now serving a life sentence	(…)
157	and in exchange she got off with manslaughter and a 12-year jail term.	(…) She got off with manslaughter and a 12-year jail term.
158	Soon after, home-made videotapes of the crimes emerged.	Soon after, home-made videotapes of the crimes emerged.
159	They showed Homolka as a willing participant,	They showed her as a willing participant,
	outraging many Canadians.	outraging many Canadians.
160	Just days before her release, Homolka sought a court order	Just days before her release, Homolka sought a court order
161	to stop the media from reporting her whereabouts.	to stop the media from reporting her whereabouts.
162	It was refused,	It was refused,
163	but the families of her victims saying she should now face up to her crimes in public.	but the families of her victims saying she should now face up to her crimes in public.

164	Despite keen media interest in her release,	Despite keen media interest in her release,
165	Homolka slipped through a crowd of protesters outside the Montréal prison with a minimum of fuss.	Homolka slipped through a crowd of protesters outside the Montréal prison with a minimum of fuss.
166	Prison officials say she's changed her name	Prison officials say she's changed her name
167	and intends starting a new life in Montréal.	and intends starting a new life in Montréal.
168	???? ???? BBC News.	(…)
169	Deaths from lung cancer, cancer of the throat and larynx are higher in the North of England than in the south, according to a report published today.	Deaths from lung cancer, cancer of the throat and larynx are higher in the North of England than in the south, according to a report published today.
170	The Office for National Statistics has released a cancer atlas	The Office for National Statistics has released a cancer atlas
171	marking the geographical patterns of where certain types of cancer occur	marking the geographical patterns of whether certain types of cancer occur
172	and their mortality rates.	and their mortality rates.
173	Besides me is our correspondent Gill Higgins.	(Line) (…)
174	So, what does this cancer atlas tell us?	What does this (…) atlas tell us?
175	What they have done is to look at the rate of cancer over the last 10 years, from 1991 to 2000, for the 21 most common cancers	(Yellow) They've looked at the rate of cancer over the last 10 years, 1991 (…) 2002, for the 21 most common cancer(…)
176	to see if they can work out if some are more common in some parts of England – some parts of UK rather - than in others.	to see if they can work out if some are more common in some parts of (…) UK than in others.
177	By doing that, they then can see if there are any risk factors that are common to those parts where the rates are high,	By doing that, they (…) can see if there are any risk factors that are common to those parts where the rates are high,
178	to see if they can then perhaps target health education campaigns and try to get the rates down.	to see if they can then perhaps target health education campaigns and try to get the rates down.
179	What have we learnt?	(White) what have we learnt?
180	are there clear areas where where certain cancers are more prevalent?	are there clear areas where (…) certain cancer(…) are more prevalent?

181	There are.	(Yellow) There are.
182	And it's for some cancers and not others.	Approximate for some cancers and not others.
183	So they think that the message is very clear	(…) They think the message is (…) clear.
184	because it is the cancers that are related to smoking or to heavy drinking that are the ones where there are wide and distinct variations.	(…) It's the cancers (…) related to smoking or (…) heavy drinking (…) are the ones where there are wide and distinct variations.
185	So you've got cancer of the lung, the pharynx, the larynx, cancer of the mouth and the lip.	So you've got cancer of the lung, far inches, the larynx, cancer of the mouth and the lip.
186	There are very high rates in the central belt of Scotland, in the industrialised northern parts of England...	There are very high rates in the central belt of Scotland, in the industrialised northern parts of (…).
187	But do we know (...) where people smoke more?	(White) (…) Do we know that's where people smoke more?
188	That's linked with lower socio-economic standards	(Yellow) that's linked with lower socio-economic standards
189	and that is where people do still smoke more and drink more.	and (…) where people do (…) smoke and drink more.
190	And the rates are half of the the high rates in areas like the south-west of England and East Anglia.	(Line) (…) Approximate the rates are half of the (…) high rates (…).
191	There is quite a difference.	There is quite a difference.
192	You don't see that for cancers like breast cancer,	You don't see that for cancers like breast cancer,
193	where there are uniform rates across the whole country.	where there are uniform rates across the (…) country.
194	Because a huge number of cancers are preventable	(White) (…) A huge number of cancers are preventable
195	or there are factors that can be tackled.	or there are factors that can be tackled.
196	That's right.	(…)
197	And they, they, they actually see that's quite a positive message.	(Yellow) (…) They (…)(…) see this as quite a positive message.
198	And if in some parts of the country the rates are low,	(…)

199	they think they should be able to get the high rates down to that low(...) level.	They think they should be able to get the high rates down to a lower level.
200	And they actually say that if they concentrate on the smoking and the alcohol-related cancers(...)(…),	(…) If they concentrate on the smoking and drinking cancer(…),,(…)
201	they can reduce the cases of cancer of about 26.000 per year	(…)
202	just by tackling these	(…)
203	and they see that this gives them a good a good target to aim for.	(…) they have (…) a good target to aim for.
204	Ok, Gill, thanks very much indeed.	(…)
205	Now, let's have a weather forecast, with Jo Farrow.	(White) (…) let's have a weather forecast (…).
206	Hallo! We have got some wet weather in the north-east of Scotland.	(…) We have got (…) wet weather in the north-east of Scotland,
207	Wet weather coming in from the west, across Ireland.	west weather coming in from the west, (…)
208	Mainly in between some drier, brighter conditions.	but in between, (…) drier, brighter conditions.
209	And the rain will hold on across the Northern Ireland.	(…)
210	In the late morning we'll begin to see the clouds increasing across northern England	(…) We'll begin to see the crowd increasing across northern England.
211	as this wet weather moves in, moving across Cumbria by lunchtime.	(…)
212	We'll hold on to the drier, brighter weather for East Anglia,	We'll hold on to the drier, brighter weather for East Anglia,
213	but it looks that we'll begin to see a little bit more clouds creeping across towards the south-east,	but it looks like (…) we'll (…) see (…) more clouds creeping across (…) the south-east,
214	perhaps one or two showers by lunch-time.	perhaps one or two showers by lunch-time.
215	And those showers really getting going through the morning across the south-west of England.	(…) Those showers really getting going through the morning across the south-east of England.

216	We'll then begin to see some heavier rain across Wales,	We'll then begin to see (…) heavier rain across Wales,
217	some pretty heavy, persistent bursts through the day.	some pretty heavy, persistent bursts through the day.
218	And a pretty wet morning for Northern Ireland.	(…) A (…) wet day for Northern Ireland.
219	A lot of wet weather around, that will edge up towards southern Scotland.	(…) That will edge up towards southern Scotland.
220	So, wet weather in the north-east.	(…)
221	We see a few showers for southern Scotland.	We see a few showers for southern Scotland.
222	And then it then becomes drier and brighter for Northern Ireland.	(…) It then becomes drier and brighter for Northern Ireland.
223	And that rain is moving towards eastern England, by the late afternoon and the evening.	(…) That rain (…) moving towards eastern England (…).
224	By the end of the week, things become more settled and slightly warmer.	By the end of the week, things become more settled and slightly warmer.
225	Welcome (...)back.	Welcome comeback.
226	We're live in Singapore in the next minute or two.	We're live in Singapore in the next minute or two.
227	We're expecting a news conference from the British bid city, London, of course.	We're expecting a news conference from the British bid city, London, of course.
228	We're going to hear from Steve Redgrave, we're going to hear from Matthew Pinsent, Tanni Grey Thompson, Denise Lewis, and, oh, yes, David Beckham.	We're going to hear from Steve Redgrave, and Matthew Pinsent, Tanni Grey Thompson, Denise Lewis, and, off, yes, David beck ham.
229	We'll be back to this later.	(…)
230	Some news summary first.	A news summary first.
231	A big security operation is in place around Gleneagles	A big security operation is in place around Gleneagles
232	in preparation for the start of the summit of the G8 group of industrialised nations tomorrow.	(…) for the start of the summit of the G8 group of industrialised nations tomorrow.
234	Protesters are expected to start gathering in the nearby town of Auchterarder.	Protesters are expected to start gathering in the nearby town of Auchterarder.
235	Meanwhile, up to a 100 demonstrators arrested	Meanwhile, (…) 100 protesters arrested in the

	during clashes with (...) police at the anti-capitalist protest in Edinburgh yesterday	clashes with the police (…) in Edinburgh yesterday
236	have started to arrive for court appearances today	have started to arrive for court appearances today.
237	and charged with various public order offences.	(…)
238	It's the final day of lobbying in Singapore for the five cities hoping to host the 2012 Olympics.	It's the final day of lobbying in Singapore for the five cities hoping to host the 2012 Olympics.
239	The International Olympic Committee will announce its final decision tomorrow.	The International Olympic Committee will announce its final decision tomorrow.
240	The Prime Minister has made a final passionate appeal for London to host the 2012 Olympics,	The Prime Minister has made a final passionate appeal for London to host the 2012 Olympics,
241	arguing that a London Games would embrace the spirit of the Olympic movement.	arguing that a London Games would embrace the spirit of the Olympic movement.
242	A love for sport, a belief in the ability of sport to bring people together to educate to enhance people's lives,	(Yellow) A love of sport, a belief in the ability of sport to bring people together to educate and enhance people's lives,
243	and a complete determination that if we are fortunate enough to host the Olympic Games	and a complete determination that if we are fortunate enough to host the Olympic Games
244	we build something that doesn't just last for the few weeks of the Games,	we build something that doesn't just last for the few weeks of the Games,
245	but last for a generation to come.	but last for a generation to come.
246	A man suspected of shooting a police officer at a house in Rawtenstall in Lancashire	(White) a man suspected of shooting a police officer at a house in Rawtenstall in Lancashire
247	has been found dead.	has been found dead.
248	Police entered the house after an 18-hour siege	Police entered the house after an 18-hour siege
249	and found a dead man and a dead dog in an upstairs bedroom.	and found a dead man and a dead dog in an upstairs bedroom.
250	No-one else was in the house at the time.	No-one else was in the house at the time.
251	34-year-old PC David Lomas is reported to be in a stable condition in hospital.	34-year-old PC David Lomas is reported to be in a stable condition in hospital.
252	A woman has been found dead in her house on Merseyside after an armed siege lasting seven hours.	A woman has been found dead in a house on Merseyside after an armed siege lasted seven hours.

253	Police were called to her house in Halewood yesterday	Police were called to a house in Halewood (…)
254	where a woman and a young boy were being held.	where a woman and a young boy were being held.
254	A 38 year-old man has been arrested.	(…)
255	In Iraq, a senior Bahrain… Bahraini diplomat has been shot and wounded	In Iraq, a senior (…) Bahraini diplomat has been shot and wounded
256	after gunmen opened fire on his car.	as gunmen opened fire on his car.
257	It took place when he was driving to work in the north-west district of Baghdad.	(…)
258	His injuries are not believed to be life-threatening.	His injuries are not believed to be life-threatening.
259	It's the second attack on an Arab envoy in Iraq in three days.	It's the second attack on an Arab envoy in Iraq in three days.
260	And in a separate instant, at least four people have been killed	And in a separate attack, several people have been killed
261	and several others wounded in an attack in the west of Baghdad.	and (…) others wounded in an attack on a minibus going to the airport.
262	They were travelling to work at Baghdad airport	(…)
263	when their minibus was ambushed.	(…)
264	Reports from India say armed militants have entered the disputed site of Ayodhya in the State of Uttar Pradesh.	(…) Armed militants have entered the disputed site of Ayodhya in (…) Uttar Pradesh.
265	Gunmen have settled a ferocious battle with police	Reports say that people have started to fight with the police.
266	and some reports say at least one attacker was killed.	(…)
267	In 1992, Hindu nationalists demolished a mosque at the site,	In 1992, Hindu's (…) demolished a mosque at the site,
268	sparking serious intercommunal riots.	sparking intercommunal conflicts
269	British Airways will, today, attempt to overturn an employment tribunal decision	British Airways will, today, try to overrule a (…) decision by the tribunal

270	which ruled in favour of a female pilot	that the rules in favour of a woman
271	who wanted to work part-time to look after her baby daughter.	who wanted (…)
272	She went to court for indirect sexual discrimination against the airline	(…)
273	after her request to cut her hours by half was turned down.	(…) to cut her hours by half (…).
274	Now the case will be heard at the employment appeal tribunal in central London.	(…) The appeal will be heard (…) in central London.
275	MPs will begin scrutinising the details of the Government's Identity Cards Bill today	(Line) MPs will begin scrutinising (…) details of the Government's Identity Cards Bill today
276	as the legislation enters its committee stage today.	as it enters the committee stage today.
277	They're expected to tackle a series of amendments	They're expected to tackle some amendments
278	and to call for a few personal details about card holders to be held on a central database.	and to call for fewer bits of information to be held on the (…) database.
279	You can keep up to date on any of the news headlines through our interactive service BBCI	(Line) You can keep up to date on any of the news headlines through (…) BBC1.
280	also if you want to go interactive personally through your own handset.	(…)
281	The London (...) bid team are about to hold a news conference in Singapore.	The London Olympic bid team are about to hold a news conference in Singapore.
282	There is the scene of our sports correspondent, James Munro.	There is (…) our sports correspondent, James Munro.
283	So, we're expecting a galaxy of talent?	So, we're expecting a huge amount of talent?
284	Yes, you've talked about David Beckham,	(Yellow) Yes, you've talked about David Beckham
285	you talked about Matthew Pinsent (...), Steven Redgrave.	and (…) Sir Steve(…) Redgrave and Matthew Pinsent,
286	I think there is a couple of areas in which London is hoping to score big	(…) (but there is a message here),
287	First of all you have the “wow” factor:	(…) apart from the “wow” factor, (…)

288	the fact that everyone in the world knows David Beckham	(…)
289	and now they know that David Beckham wants the Olympic Games in London, in the East End,	(…)
290	exactly in the area which he grew up in.	(…)
291	I think there is another more softer message here,	(but there is a message here)
292	which is that while parading so many Olympians,	is that (…)
293	London has realised that the make-up of the IOC with its 116 members (...) is changing.	(…) they know that the members of (…) the IOC board are changing
294	(who are voting tomorrow)	(…)
295	And it is changing in a particular way: a third of the members have been or are Olympians in some form,	(…) and about a third of the members have been (…) former athletes,
296	so they realise the needs that are actually wanted.	so they know what the athletes.
297	(…) That's why London has gone creating an athletes commission asking the athletes: what do you want out of a London bid?	So that's why the bid has asked (…) athletes (…) what they want out of a London bid.
298	It is not just an old form of politicians-oriented strategy.	(…)
299	So that's another area which the likes of Matthew Pinsent and Steven Redgrave are going to come across.	So that's another area which the likes of Matthew Pinsent and Steven Redgrave are going to put across.
300	And, John, what is the connection between this news conference and Singapore,	(White) And (…) what is the connection between this news conference and (…)
301	which will presumably largely be reported by the British media.	(…)
302	What's the connection between that and the delegates who are actually going to vote tomorrow?	(…) the delegates who are (…) voting tomorrow?
303	Well, they're going to hope that this press conference,	(Yellow) (…) They're going to hope that this press conference,

304	which is going out, live in BBC Britain,	which is going out, live in (…) Britain,
305	but will go out live around the world as well,	but (…) around the world as well,
306	they’re going to hope that the IOC members, (…) are going to watch this and coming across with the messages.	they’re going to hope that the delegates, will be watching it as well (…).
307	(tucked away in their hotel rooms,)	(tucked (…) in their hotel rooms)
308	I mean what London is trying to get across is these people are not here...	(…) What London is trying to get across is (…)
309	I think Sebastian Coe... Lord Coe has come up with an expression, he talked about big names being expensive calling cards.	- (…) Lord Coe has come up with an expression, he talked about big names being expensive calling cards.
310	And he's saying they are not calling cards.	And he's saying they are not (…).
311	These are people who are committed to the London (...) 2012 bid	They are people (…) committed to the London bid for 2012
312	and they've been in it for the long-term.	and they've been in it for the long-term.
313	They realised the advantages such a Games would have for London	They realised the advantages such a Games would have for London
314	and they want to get across Britain’s passion for sports.	and they want to get across Britain’s passion for sports.
315	And it is interesting... It is a very high-key approach,	(…) It is a very high-key approach,
316	and it also sort of resemble the New York approach today,	and (…) the New York approach was too,
317	Mohamed Ali was brought to parade in front of the cameras for a photo shot,	Mohamed Ali was paraded in front of the cameras for a photo shoot,
318	and New York has also gone down the same road,	and New York has (…) gone down the same road,
319	getting also big names to bang the drum for our bid.	getting (…) big names to bang the drum for their bid.
320	But it's very different to the bid... to the approach from Paris, for example.	But it's a different approach by Paris.
321	They had a very low-key press conference this morning.	(…)

322	If you have a look at some of the (...) names there, they're well-known in France	If you (…) look at some of the key names, (…) they're well-known in France, (…)
323	but (…) not really known anywhere else.	but many of them are not (…) know(…) outside France.
324	Laurent Blanc, who is a football... who was in the world cup in 1998.	(for instance, Laurent Blanc,)
325	He was there,	(…)
326	but apart from that there are very few people in the French bid team that the British people certainly would recognise	(…)
327	and it is a very different approach.	And that is their approach.
328	What the French is saying is: look IOC members are intelligent enough to read the text and read the report	(…)
329	and to make their minds.	(…)
330	They don't need to be persuaded.	(…)
331	And that is certainly something that Jacques Rogge has has been agreed on.	(…)
332	Ok James, thank you very much, (...)	(White) (…) thank you very much, James,
333	you've set the scene and whetted our appetite,	you've set the scene and whetted our appetite,
334	we'll await for the talent to arrive.	we'll await for the talent to arrive.
335	Thank you.	(…)
336	That's right.	(…)
337	We'll be back there as soon as they appear.	We'll be back there as soon as they arrive.
338	In the meantime, our sports news correspondent, James Pearce is in Singapore	In the meantime, our sports (…) correspondent, James Pearce, is in Singapore.
339	and he's been given an announcement and assessment from the British countdown.	(…)
340	Well, it was better four or five days ago.	(…)
341	It is getting unbearable now,	(Yellow) It is getting unbearable now,

342	because it's such a close-knit community now.	because it's such a close-knit community now.
343	Everywhere you go, everybody wants to ask you what you think is gonna happen	Everywhere you go, somebody asks you what you think is going to happen
344	and you ask them as well.	and you ask them as well.
345	And I'm taking you on a little walk here	I'm going on a nervous walk today,
346	which is going to be a very nervous one tomorrow for Lord Coe	because tomorrow this is where Lord Coe will walk,
347	because down here is where the voting is gonna take place,	(…)
348	where Jacques Rogge stands on a stage opening an envelope,	and Jacques Rogge will be holding an envelope,
349	which Lord Coe hopes will contain the simple word “London”.	which London hopes will really see the (…) word “London”.
350	Waiting for this big moment I will take you inside, with Tessa Jowell,	(…)
351	who is the London team secretary.	(…)
352	People are piling money on London and the odds have been slashed,	People are piling money on London and the odds have been slashed,
353	is it a wide investment(…)?	is it a wide investment, Tessa Jowell?
354	We’ll see, we’ll all know a bit later than this tomorrow.	(Cyan) We’ll see, we’ll all know a bit later than this tomorrow.
355	But I suppose that what I have to say is that I don't think we could have done more.	But (…) what I have to say is (…) I don't think we could have done more.
356	We are all here working as hard as we can,	We are all (…) working as hard as we can,
357	right up to the last minute to secure every last possible vote.	right up to the last minute to secure every last possible vote.
358	The Prime Minister has been here a couple of days before the French president,	(Yellow) The Prime Minister has been here a couple of days before the French president,
359	how much momentum has that given to the London bid?	how much momentum has that given (…) the (…) bid?
360	Oh, I mean his support has been absolutely critical for the bid.	(Cyan) Oh, he's given huge support.

361	IOC members like him.	IOC members like him.
362	They trust him	They trust him
363	and respect him,	and believe him,
364	and his support and his passion about the bid is all the evidence they need to prove the strength of the Government's support (…).	(…) and his passion about the sport is all (…) they need to prove (…) the Government's support for the sport.
365	And now I'm gonna show you a little bit, where this is gonna happen tomorrow.	(…)
366	And what will you be thinking when you go into the hall for the announcement,	(Yellow) And what will you be thinking when you go into the hall for the announcement,
367	because I mean so much hard work and so different people has gone into where we are now.	because (…) so much hard work (…) has gone into where we are now.
368	Oh exactly, yes, and the responsibility.	(Cyan) (…) Yes, and the responsibility.
369	I mean, we'll all be very nervous (…).	We'll all be very nervous over our presentation.
370	We had a very good rehearsal this morning,	We had a very good presentation this morning,
371	and we're going to do more practice this afternoon,	(…)
372	we'll be up at the crack of dawn tomorrow to practice more, so....	and we'll be up at the crack of dawn (…) to practice more. (…)
373	This is a 45 minute presentation	(Yellow) This is a 45 minute presentation
374	which is going to be crucial isn't it?	(…).
375	A 45 minute presentation	(Cyan) Yes,
376	and then about 15 minutes of questions on the presentation from members of the IOC.	and then (…) 15 minutes of questions on the presentation from members of the IOC.
377	And the IOC, there seems to have a(…) wide-spread view that (...) actually gonna settle the contest?	(Yellow) (…) There seems to (…) an wide-spread view that that could settle the contest?
378	More and more people are saying that.	(Cyan) More and more people are saying that.
379	I mean, usually with these competitions the lobbying and the rest of it has more or less settled the result before the presentations actually happen.	(…) Usually with (…) the lobbying(…), the vote is settled before the presentations, (…)

380	(...) Nobody is saying that this time.	but nobody is say that this time.
381	Everybody is saying it's too close to call	Everybody is saying it's too close to call
382	and nobody is prepared to predict the result.	and anybody is prepared to predict the result.
383	OK, Tessa Jowell, thank you very much indeed.	(Yellow) Thank you very much.
384	I'm certainly not going to predict the result either.	I'm certainly not going to predict the result eerts. (corretto con either in un secondo momento)
385	(…) But I think it is right, we can stand here and say that everybody here has the right to say that London is a genuine contender.	But it is right we can stand here – either. But (…) it is right that we can stand here and say (…) that London is a genuine contender.
386	Fingers crossed.	(…)
387	Back to you.	(…)
388	Thank you James.	(…)
389	Now, they've arrived! the athletes have arrived at the press conference, live in Singapore as you can see.	(White) Now, (…) the athletes have arrived at the press conference, live in Singapore (…).
390	They have just trooped in.	They have just trooped in.
391	James Munro is with me so they are all there?	James Munro is with me (…).
392	They are looking very smart.	(Yellow) They are looking very smart.
393	As you can see, that's Colin Jackson sitting down, next to David Hemmery.	(…) That's Colin Jackson sitting down, next to David Hemmery.
394	And of course David Beckham sitting down.	(…)
395	And Sir Steve Redgrave, that great Olympian, at the centre.	An Sir Steve Redgrave (…) is there.
396	And in a few moments the communication's starting.	(…)
397	This is the last press conference from London 2012 prior to the presentation and the vote of the IOC tomorrow.	(White) This is the last press conference from London 2012 prior to the presentation and the vote of the IOC tomorrow.
398	I am delighted today to be joined by a number of sports Ambassadors,	I am delighted today to be joined by a number of sports Ambassadors,

399	who played a key role in the development and also the promotion of the London 2012 bid.	who played a key role in the development and also the promotion of the London 2012 bid.
400	Firstly, on my left, David Hemmery,	Firstly, on my left, David Hemmery come up
401	who famously won the 400 metre hurdle medal in Mexico in 1968,	with all famously won the 400 metre hurdle medal in Mexico in 1968,
402	in a race that inspired Seb Coe to take up running.	in a race that inspired Seb Coe to take up running.
403	Next to David, Colin Jackson,	Next (…), Colin Jackson,
404	110 metre hurdle,	(…)
405	world record holder,	world record holder
406	silver medallist in Seoul in 1988.	and silver medallist in Seoul in 1988.
407	And then Tanni Grey Thompson, Dame Tanni Grey Thompson,	And then (…) Dame Tanni Grey Thompson,
408	Britain's greatest Paralympian	Britain's greatest Paralympian
409	with 11 gold medals.	with 11 gold medals.
410	Then Mr David Beckham,	(…) Mr David Beckham,
411	who is the football captain of the England football team.	who is the football captain of the England football team.
412	Next to David, is Sir Steve Redgrave.	Next to David, is Sir Steve Redgrave.
413	Steve, I think you will all know is perhaps Britain's greatest Olympian,	Steve, I think you will all know is perhaps Britain's greatest Olympian,
414	five-times rowing gold medallist,	five-times rowing gold medallist;
415	Los Angeles, Seoul, Barcelona, Atlanta and Sidney;	(…)
416	next to Steve, Denise Lewis,	next to Steve, Denise Lewis,
417	gold medallist in Sydney in the heptathlon.	gold medallist in Sydney in the heptathlon.
418	And I'm delighted also to introduce Sir Matthew Pinsent,	And I'm delighted also to introduce Sir Matthew Pinsent,
419	four times rowing gold medallist,	four times rowing gold medallist,

420	Barcelona, Atlanta, Sydney and Athens.	Barcelona, Atlanta, Sydney and Athens.
421	Next to Sir Matthew is Shirley Robinson,	Next to Sir Matthew is Shirley Robinson,
422	double sailing gold medallist	double sailing gold medallist
423	in Sydney and also in Athens.	in Sydney and (…) Athens.
424	And finally, Jonathan Edwards,	And finally, Jonathan Edwards,
425	gold medallist in the triple jump in Sydney.	gold medallist in the triple jump in Sydney.
426	So, welcome to the London Olympic bid team press conference.	(…) Welcome to the (…) press conference.
427	A number of our ambassadors will say a few words	A number of (…) ambassadors will say a few words
428	and then we'll go to questions in the usual form.	and then we(…) go to questions in the usual form.
429	And in particular, because of the great work of the athletes commission in developing the London 2012 bid,	(…) In particular, because of the great work in the ats lease' commission in (…) the (…) 2012 bid,
430	I'd like to invite Sir Steve Redgrave to say a few words as chairman of the athletes' commission.	I'd like to invite Sir Steve Redgrave to say a few words as chairman (…).
431	Steve.	(…)
432	Thank you (…) Mark.	(Yellow) Thank you very much (…).
433	As chairman it's been a very easy group to manage.	As chairman it's been a very easy group to manage.
434	It is... Sport has always been at the heart of this bid	(…) Sport has always been at the heart of this bid
435	and it's always been about athletes trying to produce the best Games.	and it's always been about athletes trying to produce the best Games.
436	The (…) athletes' advisory group was a group of athletes, (...) to have their expertise come in.	The at lease' add individuality group was a grout – athletes' advisory group (…) was that it always had athletes on the board,
437	(mainly Olympians, there is a few non-Olympians involved as well on the board)	(…)
438	The elements of it are trying to get their experiences and views at (…) events they have	(…) to try and get their (…) views at (…) sporting events they have been to, (…)

	been to, the events they have been to, sporting events and other events as well	
439	to try and make sure that our candidate city (…) was in top class	to try and make sure that our bid (…) was (…) top class
440	(that went in last November)	(that went in last November)
441	and looked at all the issues right the way through	and looked at all the issues right the way through
442	from the experiences of the athletes right the way through.	from the experiences of the athletes right the way through.
443	The actual board of 2012 were quite surprised by the input they had from the athletes' advisory group.	The (…) board of 2012 were (…) surprised by the input they had from the athletes' (…) group.
444	They were expecting to have a little bit of information about the village, about the stadiums, about their particular... individual sports,	They were expecting to have a little bit of information about the village, (…) the stadiums and their (…) individual sports,
445	about every part of the Games.	but every part of the Games,
446	In fact, the evaluation committee came	in fact, the evaluation committee came
447	and valued 17 different areas,	and valued 17 different areas,
448	and the athletes had views and opinions and advice on all 17 parts of it.	and the athletes had opinions and views and advice on all 17 parts of it.
449	And a lot of that advice has been put into the candidate file that was put in November,	And a lot of that advice has been put into the candidate file that was put in November,
450	and has followed right the way through to the presentation tomorrow.	and has followed right the way through to the presentation tomorrow.
451	So I'm really pleased with everybody's elements that they've put in.	So I'm (…) very pleased with everybody's elements (…) they've put in.
452	With all the sporting ambassadors that we've had here over the last few days, they've all been playing a vital part to,	With all the sporting ambassadors (…) we've had here in the past few days, they've all played a vital part to,
453	hopefully (...) London winning the 2012 Games tomorrow.	hopefully, the London winning the 2012 Games tomorrow.
454	And it's not that they just have been wheeled out for the publicity here for the last few days,	And it's not just that they have been wheeled out for the publicity here in the past few days,

455	they've been involved for a number of months and years in the build-up to it.	they've been involved for a number of (…) years in the buildup to it.
456	So the athletes' advisory group has really been in the heart and forefront of the bid.	So the athletes' review body has (…) been right at the heart of it all.
457	So I'd like just to thank all of my colleagues here for their engagement in the last few years	So I'd like (…) to thank all of my colleagues' in put to the bid today.
458	and hopefully they are going to take in for the next few years as well.	(…)
459	So Thank you all very much indeed.	(White) (…) thank you (…) very much (…).
460	Thank you (…) Steve.	Thank you very much (…).
461	I'm gonna ask Denise Lewis to say a few words.	And I'll ask Denise Lewis to say a few words.
462	Tomorrow obviously is a very important moment for all the candidate cities	Tomorrow is obviously a very important day for all the candidate cities,
463	with their final presentations,	(…)
464	and I'm delighted that one of the presenters tomorrow will be Denise.	and I'm delighted that one of the presenters (…) will be Denise.
465	And I won't pass out, I don't think.	(Yellow) and I won't pass out, I don't think.
466	Yes, tomorrow is a big day.	Yes, tomorrow is a big day.
467	I'm speaking. I have the honour of speaking.	(…) I have the honour of speaking.
468	It gives me great pleasure to be able to perform this task on behalf of my colleagues here,	It gives me great pleasure to be able to perform this task on behalf of my colleagues here,
469	and, in fact, on behalf of Britain.	and, in fact, on behalf of Britain.
470	Just as to echo what Steve said, is that we're not here just as decoration,	(…)
471	we have been really involved from day one	We have been really involved from day one
472	and tried to convey really what sport means to us	and tried to convey really what sport means to us
473	and what the Olympics actually mean to us.	and what the Olympics actually mean to us.
474	It's the biggest event in the world,	It's the biggest event in the world,
475	and we just want to play a significant role in	and we just want to play a significant role in

	that.	that.
476	So tomorrow I'll be speaking,	So tomorrow I'll be speaking,
477	and hopefully conveying the passion that we have for sport.	and hopefully conveying the passion that we have for sport.
478	Sport has been - well, it's been my life so far,	Sport has been - well, it's been my life so far,
479	and I've enjoyed every minute of it.	and I've enjoyed every minute of it.
480	And hopefully I'll do my best tomorrow	And hopefully I'll do my best tomorrow
481	and bring the bid home.	and bring the bid home.
482	Thank you Denise.	(…)
483	Tanni?	(…)
484	Tanni will also be part of the podium group at the presentation.	(White) Tanni will also be part of the podium group at the presentation.
485	I feel very proud to be a British athlete and a Paralympian.	(Yellow) I feel very proud to be a British athlete and a para Olympian.
486	There is not another country in the world with so much attention to Paralympian athletes as the UK	(…)
487	both in terms of financial support, and in terms of the media coverage that we receive,	Both in terms of financial support, (…) the media coverage that we receive,
488	and also the general public wanting to be part of the Paralympics.	and also the general public wanting to be part of the Olympics.
489	For me, what's been a massive strength of the London 2012 bid has been about the Olympic Games and the Paralympics.	For me, what's been a massive strength of the London 2012 bid has been about the Olympic Games (…)
490	(…) Paralympics hasn't been something that has been tagged on at the end.	The paralympics hasn't been (…) tagged on at the end.
491	Every single part of the bid has been thought about with Paralympians in mind.	(…)
492	And that, for me, shows what a great country Britain is	(…) That for me shows what a great country Britain is
493	and how supportive it is of all the athletes	and how supportive it is of (…) the athletes
494	and all the athletes wanting to do well.	and all the athletes wanting to do well.

495	Thank you.	(…)
496	And David, as the boy from the East End of London,	(White) (…) David, as the boy from the East End of London,
497	I'd like to ask David to say also a few words.	I'd like (…) David to say (…) a few words.
498	I think I, as a sportsman, I am very proud and very honoured to be part of this team,	(Yellow) I think I, as a sportsman, I am very proud and (…) honoured to be part of this team,
499	and I'm sure if you asked individually everyone of these people that have sat here at this table,	and I'm sure if you asked individually everyone of these people that have sat here at this table,
500	they would tell you how much they believe and how much hard work has gone into this bid.	they would tell you how much they believe and how much hard work has gone into this bid.
501	I'm obviously from the East End of London,	I'm obviously from the East End of London,
502	so to see the Olympics being in that area would be incredible for me personally	so to see the Olympics being in that area would be incredible for me personally
503	and for everyone involved in the East End.	and for everyone involved in the East End.
504	I think if we are given the chance to host this Olympics,	I think if we are given the chance to host this Olympics,
505	then people will see that it was the right decision,	then people will see that it was the right decision,
506	because I think that it could be - it could go down in history as being one of the best Olympics that has ever been.	because I think that it could be - it could go down in history as being one of the best Olympics that has ever been.
507	So I'm here supporting	So I'm here supporting
508	and I'm here as one of the team,	and I'm here as one of the team,
509	and we're hoping for the best result tomorrow night.	and we're hoping for the best result tomorrow night.
510	Thank you, David.	(White) Thank you, David.
511	I will take questions in the usual form.	(Line) We'll take questions in the usual form.
512	If you could indicate,	If you could indicate,
513	wait for the microphone,	wait for the microphone,
514	indicate your name and your organisation	indicate your name and your organisation

515	and we'll try and take as many questions as possible.	and we'll try and take as many questions as possible.
516	We'll start there with Vicky.	We'll start there with Vicky.
517	Ian.	Ian.
518	Hi ?????? the Sun.	(…)
519	Can I ask David:	(Yellow) Can I ask David:
520	Everyone here has very clear Olympic memories,	Everyone here has (…) clear Olympic memories,
521	and I just wondered, growing up, if there's anything that stands out, watching on TV or anything that inspired you.	and I just wondered, growing up, if there's anything that stands out, watching on TV or anything that inspired you.
522	I think a lot of things inspired me,	(Cyan) I think a lot of things inspire me,
523	but I think when I was watching the Olympics,	but I think when it was watching the Olympics,
524	I think people like Seb Coe, running in barefoot,	I think people like Seb co-running in barefoot,
525	that inspired a lot of people.	that inspired a lot of people.
526 527	I think watching Dailey Thompson and so many of these, you know, athletes that you (…) look up to and you watch and, you know,	I think watching Daley Thompson. So many of these (…) athletes that you to look up to and you watch (…)
528	that inspire different children all around the world.	and inspire different children all around the world.
529	So, as I said, to be part of this team, you know,	So, as I said, to be part of this team (…)
530	and to have memories from Olympics that have gone by,	and to have memories from Olympics that have gone by,
531	as I said, from Daley Thompson,	as I said, from Daley Thompson
532	Sebastian Coe	and Seb co-
533	and different athletes,	and different athletes,
534	you know, their my memories from different Olympics.	you know, their my memories from different Olympics.
535	Thank you.	(White) Thank you.

536	Ben, just behind you.	Ben, just behind you.
537	Ben Brown, BBC News.	Ben Brown, BBC News.
538	David, can I ask you:	(Yellow) David, can I ask you:
539	What do you think it would mean for the whole of British sports to win this bid is tomorrow?	What do you think it would mean for the whole of British sport to win this bid is tomorrow?
540	I think, as everyone... as Sir Steve, Sir Steve has... has said, you know, there's a lot of passion that is in our country with sports,	(Cyan) I think, (…) as Sir Steve, (…) has said, (…) there's a lot of passion that is in our country with sports,
541	with football,	with football.
542	I realise more than anything.	I realise more than anything,
543	During big competitions (…) our country comes together better than any other country I've seen before.	(…) big competition as our country comes together better than any (…) country I've seen before.
544	I think that's what our country can give, and especially the East End of London can give to this bid, you know.	I think that's what our country can give, and especially the East End of London can bring to this bid (…).
545	It could be it could be an incredible thing,	It could be (…) an incredible thing.
546	because that's one thing about our country... that's one thing that our country has got.	(…) That's one thing that our country has got.
547	You know, when there is a big competition, it comes together like no other country has ever come together before.	(…) When there is a big competition, it comes together like no other country has ever come together before.
548	Take Victoria at the back, there.	(White) Take victoria at the back (…).
549	Hi Victoria ????? BBC Radio five Live.	(…)
550	I'd like to ask this of David Beckham and Colin Jackson, if I may.	(Yellow) I'd like to ask this of David Beckham and Colin Jackson (…).
551	There isn't unanimous support for this bid from the British public.	There isn't nam support (…) from the British public.
552	What do you say to those back home who don't want London to win?	What do you say to those bach home who don't want London to win?
553	You answer that.	(Cyan) You answer that.
554	That's for Colin, I think.	(White) That's for Colin, I think.

555	You're definitely going to have that.	(Green) you're definitely going to have that.
556	I think every single bidding city has people that have a negative attitude towards the Olympic Games for me.	I think every single bidding city has people that have a negative attitude towards the Olympic Gamesment for me,
557	I don't live in London.	I don't live in London.
558	I live out in one of the regions.	I live out in one of the regions.
559	So it might seem very distant to the people of the United Kingdom who live in the regions	So it might seem very distant to the people of the United Kingdom who live in the regions
560	and can't quite understand what is going on,	and can't quite understand what is going on,
561	what the hoo-ha is.	what the hoo-ha is.
562	But my job has been quite clear.	But my job has been quite clear.
563	I have been educated them	I have been educated them
564	and putting them through their paces	and putting them through their paces
565	so they can understand what the Olympic Games will mean, not just to London but to Great Britain itself as a whole nation.	so they can understand what the Olympic Games will mean, not just to London but to Great Britain itself as a whole nation.
566	Sport unifies nations.	Sport unifies nations.
567	Sport unifies the world.	Sport unifies the world.
568	It's been documented in history.	It's been documented in history.
569	So I think once we get the bid announced that we shall win,	So I think once we get the bid announced that we shall win,
570	what will happen is, I think, you'll get a huge group of unity there.	(…) you'll get a huge group of unity there.
571	??? And people that won't get more than understanding will get behind the wall ????	(…)
572	It won't be difficult once the announcement is made to get the people to disbelieve all of a sudden to believe.	It won't be difficult once the announcement is made to get the people to disbelieve all of a sudden to believe.
573	Can I just say the last opinion poll put public support at 80%.	(White) Can I just say the last opinion poll put public support at 80%.
574	And I think for a country like the United Kingdom to have 80% support for anything is	I think for a country like the United Kingdom to have 80% support for anything is pretty

	pretty remarkable.	remarkable.
575	Any other questions?	Any other questions?
576	We'll take the colleague there. You're next.	We'll take the colleague there. (Line) You're next.
577	David, you're an international star,	(Yellow) David, you're an international star,
578	but you are an East London boy at heart.	but you are in east London boy at heart.
579	What I want to know from that East London boy is how much does it mean to have the Olympics in your hood?	What I want to know from that East London boy is how much does it mean to have the Olympics in your hood,
580	And how much will it change the lives of that local people in the area who tonight will be holding their breath?	and how much will if change the lives of (…) local people in the area who tonight will be holding their breath?
581	I think to have the Olympics in my manor would be... I think... as I said, you know, I've grown up... I grew up in the East End of London.	(Cyan) I think to have the Olympics in my manner would be - I think, (…) I said, (…) (…) I grew up in the East End of London.
582	I've got friends that have got children that have grown up in the East End of London,	I've got friends that have got children that have grown up in the East End of London,
583	and they're all saying to me, to have the Olympics, as I said, in their... in our manor would be a special thing for kids, you know,	and they're (…) saying to me, to have the Olympics (…) (…) in our manner would be a special thing for kids (…)
584	to have inspiration from different athletes from all around the world, not just from our country but from all around the world,	to have inspiration from different athletes from all around the world (…)
585	that they can inspire to, that they can have their dreams and their goals,	that they can inspire to, that they can have their dreams and (…) goals,
586	and they can go on and watch them,	and they can begun watch them,
587	which is just going to be down the road from them.	which is just going to be down the road from them.
588	That's going to be the biggest thing.	That's going to be the biggest thing.
589	It's going to regenerate so many things in London, in our country.	It's going to regenerate so many things in London, in our country.
590	You know, it's not... it's not just obviously about London, you know, it's about the whole country.	(…) It's not (…) just obviously about London. (…) It's about the whole country.

591	And that's inspiration enough.	And that's inspiration enough.
592	The colleague next to you.	(White) The colleague next to you.
593	?????? from Japan ???? station.	(…)
594	I have a question to David Beckham.	(Yellow) I have a question to David Beckham.
595	I'm sorry too much for this kind of question:	I'm sorry (…)
596	do you have confidence to win tomorrow?	do you have confidence to win tomorrow?
597	And could you tell me the reason why?	And could you tell me the reason why?
598	I think the reason why we have the confidence, of course we need confidence, we need to have confidence	(Cyan) I think the reason why we have the confidence, (…) we need to have confidence
599	because there's so many people, you know, and so many people behind the scenes that work very hard for this bid to get it to, you know,	because there's so many people (…) and so many people behind the scenes that have worked very hard for this bid to get it to
600	for me, and I'm sure I'm speaking for the rest of the team, that, you know, we've succeeded already in getting to this stage.	- for me, and I'm sure I'm speaking for the rest of the team, that (…) we've succeeded already in getting to the stage.
601	I think the hard work that has gone into this bid has been incredible.	I think the hard work that has gone into this bid has been incredible.
602	A lot of people don't see the hard work behind the scenes that go into it.	A lot of people don't see the hard work behind the scenes that go into it.
603	But for me, you know, this is the easy part of turning up, meeting people, and speaking to the press.	But for me, (…) this is the easy part of turning up, meeting people, and speaking to the press.
604	You know, that's the easy part.	(…) That's the easy part.
605	The hard work has been done for a long time now.	The hard work has been done for a long time now.
606	So, you know, this is the easy part.	So (…) this is the easy part.
607	But of course we have confidence in the result,	But of course we have confidence in the result,
608	we'll have to wait and see tomorrow night.	we'll have to wait and see tomorrow night.
609	Ok, The colleague just behind you and there.	(White) (…) The colleague just behind you and there.

610	I'm ???? from the ????	(…)
611	This question is for David Beckham.	(Yellow) This question is for David Beckham.
612	You do not have a speaking part in London's presentation,	You do not have a speaking part in London's presentation,
613	and you play for Real Madrid.	and you play for Real Madrid.
614	So by being here, how much value do you think you'll add to London's bid?	So by being here, how much value do you think you'll add to London's bid?
615	I think my value is being part of a team that I've been,	(Cyan) I think my value is being part of a team that I've been,
616	as I said, I am very honoured to be part of this team, you know.	as I said, I am very honoured to be part of this team (…).
617	I'm not here because of my celebrity profile.	I'm not here because of my celebrity profile.
618	I'm here because I'm a team member.	I'm here because I'm a team member.
619	And I've been asked, and as I said I'm very honoured to be part of this team, you know.	And I've been asked, and (…) I'm very honoured to be part of this team (…).
620	So my value is being part of this team	So my value is being part of this team
621	and showing the strength and you know, the inspiration that can be shown to having the Olympics in our country.	and showing the strength and (…) the inspiration that can be shown to having the Olympics in our country.
622	Thank you.	(…)
623	I think I can add that David has been involved in over a number of months on some of the promotional videos that we've done	(Green) (…) (…) David has been involved in (…) some of the promotional videos (…)
624	promoting London and the promoting of the bid,	promoting London and (…) the bid,
625	and he's been right at the heart of it, right the way through.	and he's been right at the heart of it, right the way through.
626	It's not like he's just turned up and been part to here.	It's not like he's just turned up (…).
627	He's been an ambassador for a long time.	He's been an ambassador for a long time.
628	Before the England football team left for the European championships,	Before the England football team left for the (…) championship(…),

629	that they had a photo call and pledging their support as a team towards the London 2012.	(…) they had a photo call (…) pledging their support as a team towards (…) London 2012.
630	Simon?	(…)
631	Simon ??????? ITV news.	(…)
632	Matthew, you've been here a few days doing the rounds.	(Yellow) Matthew, you've been here a few days doing the rounds.
633	Do you get any sense of how it's going?	Do you get any sense of how it's going?
634	I think you tend to hear the same stories going round and round again.	(Cyan) I think you tend to hear the same stories going round and round again.
635	It's absolutely impossible to accurately predict what's going to happen tomorrow.	It's absolutely impossible to accurately predict what's going to happen tomorrow.
636	That makes it very exciting, nerve-wracking as well.	That makes it very exciting, nerve-wracking as well.
637	But we have to stay confident in what we believe.	But we have to stay confident in what we believe.
638	We believe London has the best bid.	We believe London has the best bid.
639	We have always believed that.	We have always believed that.
640	It's a huge amount of preparation and a huge amount of work that goes into one day,	It's a huge amount of preparation and a huge amount of work that goes into one day,
641	but we're used to that.	but we're used to that.
642	We're (...) athletes, or ex-athletes.	We're other athletes, or ex-athletes.
643	So as hard as it's going to be, we have to stay focused and concentrated right the way through tomorrow.	So as hard as it's going to be, we have to stay focused and concentrated right the way through tomorrow.
644	But can we actually predict the outcome of any context? No.	But can we read contractly predict the outcome of any context? No.
645	Does that make it easy? No.	Does that make it easy? No.
646	Do we want to be anywhere else? No.	Do we want to be anywhere else? No
647	Thank you.	(White) Thank you.
648	John.	John.
649	John ???? from ITV news.	(…)

650	A question for David.	(Yellow) A question for David.
652	David, you've become wealthy through sport.	David, you've become wealthy through sport.
653	You've won many an honour and we know that through sport,	You've won many an honour through sport,
654	because you never and can never win an Olympic gold medal.	(…) but you (…) can never win an Olympic (…) medal.
655	Are you envious of the people around you there,	Are you envious (…),
656	and what's special, do you think, about the Olympics as a sporting event?	and what's special (…) about the Olympics as a sporting event?
657	Yes, that's very true.	(Cyan) (…) That's very true.
658	I've never won a gold medal	I've never won a gold medal
659	and probably never will win a gold medal, I mean in the Olympics.	and probably never will win a gold medal (…) in the Olympics.
660	But I have achieved it in other ways, which is obviously my football.	But I have achieved in other ways, which is obviously my football.
661	I've won certain thinks with Manchester United	I've won certain thinks with Manchester United
662	and, nothing with Real Madrid yet.	and, nothing with Real Madrid yet.
663	I've hoping to win something with the national team, you know,	I've hoping to win something with the national team (…).
664	that's what we all aspire to, we all aspire to winning things, winning gold medals (...) and doing the best in your sport possible.	(…) We all aspire to winning things, winning gold medals than and doing the best in your sport possible.
665	That's what I'm looking to do.	That's what I'm looking to do.
666	Thank you David.	(…)
667	The colleague in front of him.	(…)
668	You need a microphone.	(…)
669	Give him a microphone.	(…)
670	So, here you are.	(…)
671	????????? from the United Nations	(White) REPORTER: (Yellow) I'm (…) from

	environmental programme in Nairobi.	the United Nations environment programme in Nairobi.
672	How does the environment play a role in sport and what..., in the Olympics, and what changes do... is planned if London wins the bid?	How does the environment play a role in sport, (…) in the Olympics, and what changes are planned if London wins the bid?
673	I don't have one specific question to one.	I don't have one specific question to one.
674	Whoever is the most appropriate person.	Whoever is the most appropriate person.
675	Steve or Jonathan perhaps.	(White) Steve or Jonathan perhaps.
676	I guess David has answered enough questions already.	(Yellow) I guess David has answered enough questions already.
677	Mainly from journalists	(…)
678	Mainly from journalists, yes.	(White) Mainly from journalists, yes.
679	The environment is a very big issue.	(Cyan) The environment is a very big issue.
680	Over the last four or five Games, it's become very big news	Over the last four or five Games, it's become very big news
681	and Sydney was really the first to make very big changes.	and Sydney was really the first to make very big changes.
682	The London bid is not different to that,	The London bid is not different to that,
683	and also trying to move things on from where we've been before.	and also trying to move things on from where we've been before.
684	Where… the Olympic Park will be just outside Stratford, in the new East End of London,	Where (…) Olympic Park will be just outside Stratford (…)
685	is a half (...) derelict site	is a half dare derelict site
686	and that is being cleared as we speak.	and that is being cleared as we speak.
687	There will be an aquatic centre,	There will be an aquatic centre,
688	there will be two hockey pitches	there will be two hockey pitches
689	and even if we're unsuccessful tomorrow, that will still happen	(…) even if we're unsuccessful tomorrow, that will still happen
690	and that site will be cleared.	and that site will be cleared.
691	But for the Games itself –	But for the Games itself –

692	Sorry about the flash photography that disturbs you	(White) Sorry about the flash photography (…).
693	but there we have Britain's best.	(…) There we have Britain's best.
694	We've heard from David Beckham, saying that he wants the Olympics in his manor.	We've heard from David Beckham, saying that he wants the Olympics in his manner.
695	A very engaging performance by all of them.	A very engaging performance by all of them.
696	Yes, it seemed very confident, and very relaxed.	(Yellow) (…) it seemed very confident, (…) relaxed.
697	What Matthew Pinsent was saying there, we have got to stay focused.	(…) We've got to stay focused.
698	We don't know if we gonna win or not.	(…)
699	All we can do is to believe that we have the best bid.	All we can do is to believe that we have the best bid.
700	I think time and again you heard from the likes of Sir Steve Redgrave that sport is at the heart of this bid.	I think time and again you heard from (…) the Red grave that (…)
701	You have David Beckham: there is a lot of passion in the sport.	(…) there is a lot of passion in the sport.
702	I think that's what we are trying to get across here.	(…)
703	Thank's very much James.	(…)
704	We'll be back in Singapore in (...) a few moments.	(White) We'll be back in Singapore in just a few moments.