

Numero monografico sul

Respeaking

a cura di
Carlo Eugeni
Luigi Zambelli



INDICE

EDITORIALE: LE MONOGRAFIE DI SPECIALIZZAZIONE ON-LINE Carlo Rodriguez.....	6
PRESENTAZIONE DELLA MONOGRAFIA SUL RESPEAKING Carlo Eugeni e Luigi Zambelli.....	7
PARTE I: INTRODUZIONE AL RESPEAKING.....	10
1) INTRODUZIONE AL RESPEAKING Carlo Eugeni.....	11
1. Descrizione	11
2. Applicazioni e tecnologia.....	11
3. Storia e pratica.....	11
4. La ricerca.....	12
5. La formazione	12
2) FUNZIONAMENTO DI UN SOFTWARE DI RICONOSCIMENTO DEL PARLATO Antonio Ingrassia.....	13
Introduzione	13
Aspetti informatici	13
Aspetti tecnico-tecnologici	14
Funzionamento del software ASR	15
Articolazione del modello del parlato	17
3) I BENEFICI DEL RESPEAKING PER SORDI Tiziana Trapani.....	21
1. Breve introduzione alla sordità	21
2. Il respeaking nei media.....	21
3. Il respeaking nel resto della società	22
4. Conclusioni	23
4) LA STORIA DEL RESPEAKING – EVOLUZIONE DEL PROGETTO EUROPEO VOICE Giuliano Pirelli	25
1. La nascita dei sottotitoli	25
2. Difficoltà di comunicazione	26
3. Il riconoscimento della voce.....	26
4. Introduzione al Progetto VOICE.....	27
5. Un’azione di sensibilizzazione	28
6. Il Sito e il Forum VOICE	29
7. La sottotitolazione delle trasmissioni televisive.....	29
8. Il Gruppo Inter-Servizi sulla Disabilità	30
9. Il Respeaking.....	30
10. Il seguito del Progetto VOICE.....	31
11. La produzione dei sottotitoli televisivi tramite riconoscimento vocale	32
12. La produzione dei sottotitoli a distanza	33

5) LA TRASCRIZIONE TRAMITE RESPEAKING IN INTERSTENO Fabrizio G. Verruso	35
1. Introduzione	35
2. Le gare di ripresa del parlato da Amsterdam 1995 a Hannover 2001.....	35
3. Le gare di ripresa del parlato da Roma 2003 a Parigi 2012	36
4. Conclusioni	37
6) LA SOTTOTITOLAZIONE TRAMITE RESPEAKING NEL MONDO Pablo Romero-Fresco (traduzione a cura di Tiziana Trapani).....	39
Il respeaking in ambito televisivo	39
La selezione e la formazione dei respeaker.....	40
PARTE II: PROFESSIONI E FORMAZIONE.....	43
7) ASPETTI DIDATTICI Claudio Russello.....	44
1. La fase preparatoria: l'addestramento del software.....	44
2. Il lavoro in diretta: le competenze psico-cognitive	46
3. Aspetti didattici propri del respeaking e competenze complessive.....	49
4. Conclusioni	51
8) ASPETTI PROFESSIONALI NEL RESPEAKING IN PREREGISTRATO	53
8a. LA RESOCONTAZIONE DALLA STENOGRAFIA MANUALE AL RESPEAKING Luigi Zambelli.....	53
8b. PROPOSTA DEL FORUMTAL SUL VERBALE MULTIMEDIALE DI ATTI GIUDIZIARI AUDIOREGISTRATI Francesco Cutugno e Andrea Paoloni.....	61
9) ASPETTI PROFESSIONALI NEL RESPEAKING IN TEMPO REALE	64
9a. IL RESPEAKING TELEVISIVO Arianna Mori	64
9b. IL LIVE EDITING TELEVISIVO Silvia Velardi	68
PARTE III: RICERCA E SVILUPPO	73
10) ASPETTI TEORICI DEL RISPEAKERAGGIO TELEVISIVO Francesca Marchionne	74
1. Introduzione	74
2. "La mappa degli studi sul respeaking".....	75
3. I DRS.....	76
4. Conclusioni	78
11) METODI E TEMPI DI UNA VERBALIZZAZIONE – PROVE TECNICHE Rossella D'Arcangelo e Francesco Cellini	81
1. Introduzione e metodologia.....	81
2. Brano Lento — Durata 5 minuti	84
3. Brano Medio — Durata 5 minuti	86
4. Brano Veloce — Durata 5 minuti.....	90

5. Risultati.....	93
6. Conclusioni	94
12) SOTTOTITOLI DI CONFERENZA INTRALINGUISTICI E INTERLINGUISTICI – STRATEGIE TRADUTTIVE A CONFRONTO Silvia Bellucci.....	96
1. Introduzione	96
2. Metodologia	96
3. L’esperienza.....	97
4. Risultati della prima parte dell’esperienza – la sottotitolazione intralinguistica.....	98
5. Risultati della seconda parte dell’esperienza – la sottotitolazione interlinguistica.....	99
13) IL RESPEAKING INTERLINGUISTICO Francesca Marchionne	102
1. Introduzione	102
2. Perché e per chi.....	103
3. Sperimentazioni e difficoltà.....	107
4. Conclusioni	109
14) SVILUPPI FUTURI – PROGETTI DELL’ASSOCIAZIONE INTERNAZIONALE DI RESPEAKING onA.I.R. Tiziana Trapani.....	110
1. Introduzione	110
2. OltreSuoni – Dalla voce alla scrittura per una radio accessibile a tutti.....	110
3. RadioText – Leggendo anche noi ascoltiamo la radio	111
4. PALIOSS – Piattaforma per l’Accessibilità alla Lingua Orale per Sordi e <i>Stranieri</i>	112
5. A, E, I, OnAIR – Il prof. si sottotitola da solo.....	112
6. Lectio subtulit – Prendo appunti per te	112
7. TVonA.I.R. – L’informazione è uguale per tutti	113
8. Note Vibranti – Concerti accessibili per i sordi.....	113
15) NOVITA’ TECNOLOGICHE.....	115
15a) IL PROGETTO EUROPEO EU BRIDGE – UN PONTE PER COLMARE IL DIVARIO LINGUISTICO Margit Rödder (traduzione di Silvia Bellucci)	115
15b) IL PROGETTO EUROPEO SAVAS – RICONOSCIMENTO DEL PARLATO PER RESPEAKING E SOTTOTITOLAZIONE IN TEMPO REALE Carlo Aliprandi	117
15c) IL PROGETTO EUROPEO SUMAT – SOTTOTITOLAZIONE ASSISTITA DALLA TRADUZIONE AUTOMATICA Panayota Georgakopoulou (traduzione di Silvia Velardi)	120
1. Introduzione	120
2. Il progetto.....	120
3. Avanzamento del progetto.....	121
POSTFAZIONE: IL FUTURO DELLA TRASCRIZIONE Luigi Zambelli	123
BIBLIOGRAFIA.....	127

LE MONOGRAFIE DI SPECIALIZZAZIONE ON-LINE

Specializzazione on-line si è assunta il compito di essere voce dell'Accademia G. Aliprandi - F. Rodriguez. Oggi si arricchisce di una nuova iniziativa. Con questo numero speciale sul "Respeaking" prende corpo un'idea che ho coltivato a lungo, quella di realizzare una forma agile di comunicazione e di divulgazione delle discipline di cui l'Accademia G. Aliprandi - F. Rodriguez si occupa.

Il numero monografico pubblicato on-line permette revisioni, integrazioni ed aggiornamenti nel tempo. E anzitutto consente di comunicare storia, contenuti, strumenti, ricerche che non hanno in genere ampia diffusione né un palcoscenico idoneo attraverso il quale farsi conoscere, o almeno proporsi alle nuove generazioni.

Questo numero monografico dedicato al Respeaking è stato curato da Luigi Zambelli e Carlo Eugeni.

A loro ed a tutti gli Autori che hanno accolto l'invito a fornire il loro contributo al piano dell'opera ed alle sue intenzioni di fondo, va il mio ringraziamento e di tutto il Consiglio Direttivo dell'Accademia insieme all'augurio che questo sforzo possa essere imitato e portare ulteriori risultati.

Le finalità perseguite e l'interesse del lavoro realizzato possono essere un felice viatico per nuove iniziative a cui tutti sono invitati a concorrere con i loro lavori, idee e proposte.

Con questo auspicio ho il piacere di augurare una buona e produttiva navigazione a questa navicella varata nel grande mare delle attività e del lavoro dell'uomo.

Firenze, 31 maggio 2013

Carlo Rodriguez

Presidente Accademia G. Aliprandi-F. Rodriguez

[Torna all'indice](#)

PRESENTAZIONE DELLA MONOGRAFIA SUL RESPEAKING

Nel tentativo di divulgare il sapere scientifico negli ambiti di sua competenza, l'Accademia italiana "Giuseppe Aliprandi – Flaviano Rodriguez" ha ideato la pubblicazione online di numeri monografici sulle tecniche di scrittura, nei quali fornire la panoramica più ampia possibile di tutti gli ambiti di investigazione e applicazione della tecnica in questione. Questo primo volume è dedicato al respeaking. La miscellanea, curata dai professori Luigi Zambelli e Carlo Eugeni e resa possibile grazie al contributo di alcuni dei massimi esperti italiani e internazionali del settore che ringraziamo tutti vivamente, vuole offrire a tutto il bacino di potenziali lettori una visione a 360 gradi di questa recentissima tecnica di scrittura veloce applicata alla produzione di testo trascritto. A tal fine il volume è stato diviso in tre parti, tra esse propedeutiche e complementari.

La prima parte, intitolata "**Introduzione al respeaking**", è rivolta ai lettori neofiti, che potranno trovare nei sei articoli che la compongono una presentazione generale della disciplina in oggetto, a cominciare da una breve introduzione, che spiega che cos'è il respeaking, ne illustra brevemente le potenzialità e mette in contesto gran parte della terminologia che verrà adottata nel prosieguo dell'opera. Il secondo contributo verte invece sugli aspetti prettamente informatici che stanno alla base della tecnologia su cui si basa il respeaking. Segue una illustrazione delle ripercussioni positive dell'applicazione del respeaking su una delle categorie più linguisticamente svantaggiate della società, cioè le persone affette da sordità totale o parziale, con particolare riferimento a quelle che fanno fatica a percepire le frequenze del parlato. Il quarto articolo si concentra sulle origini del respeaking, ripercorrendo le tappe che hanno portato alla sua nascita. Il quinto e sesto capitolo chiudono la prima parte facendo il punto della pratica del respeaking nel mondo della trascrizione professionale e nel mondo della sottotitolazione in tempo reale.

La seconda parte del volume, intitolata "**Professione e formazione**", è incentrata sugli aspetti operativi del respeaking, sia in preregistrato sia in tempo reale e interessa in particolare i professionisti del settore. Il settimo capitolo propone una completa carrellata degli aspetti essenziali riguardanti la formazione alla più complessa, e per il momento poco nota al grande pubblico, professionalità della sottotitolazione in tempo reale, quella del respeaker. Nell'ottavo capitolo, suddiviso in due contributi, si analizzano in maniera approfondita e tramite l'occhio dei professionisti, le peculiarità del respeaking per produrre testi offline negli ambiti della trascrizione degli atti di un convegno, della verbalizzazione di un atto processuale e della resocontazione parlamentare. Allo stesso modo il capitolo nono, che consta anch'esso di due articoli, si

[Torna all'indice](#)

concentra sulle tre professionalità in atto nel respeaking in tempo reale usato per produrre sottotitoli televisivi di programmi in diretta: il respeaking, l'editing di sottotitoli prodotti tramite respeaking e il doppio editing della trascrizione automatica.

La terza parte, "**Ricerca e Sviluppo**", si rivolge principalmente a quanti hanno interessi di ricerca scientifica in ambito professionale e accademico sia in preregistrato sia in tempo reale. Più nel dettaglio il decimo capitolo fa il punto sullo stato dell'arte della ricerca in atto sul respeaking e ne illustra gli ambiti di investigazione. Seguono tre relazioni (numero 11- 12- 13) su altrettanti esperimenti sia nell'area della trascrizione offline sia in quello della sottotitolazione in tempo reale. La prima relazione propone una dettagliata analisi comparata tra respeaking, trascrizione automatica, stenotipia e dattilografia nella produzione di diversi tipi di testo; la seconda confronta le strategie messe in atto nel respeaking interlinguistico e in quello intralinguistico preceduto da interpretazione simultanea nell'ambito della sottotitolazione interlinguistica in tempo reale di convegni; la terza si concentra sulla fattibilità della sottotitolazione in tempo reale tramite respeaking interlinguistico mettendo a confronto il respeaking intralinguistico di determinati programmi in lingua straniera e quello interlinguistico degli stessi programmi.

Chiudono la terza parte quattro articoli su interessanti progetti internazionali che vedono il respeaking in particolare e il Trattamento Automatico del Linguaggio in generale come protagonisti, in vista di una evoluzione del respeaking sia in ambito applicativo (il primo articolo – numero 14 - passa in rassegna tutti i progetti realizzati dall'Associazione Internazionale di Respeaking onA.I.R. che vanno dalla sottotitolazione di lezioni universitarie e scolastiche all'accessibilità dei concerti, dalla sottotitolazione interlinguistica a distanza alla sottotitolazione radiofonica per sordi e stranieri passando per la telefonia accessibile per tutti), sia in ambito tecnologico dove la strada è spianata alla sottotitolazione automatica in tempo reale accompagnato da traduzione automatica plurilingue, oggetto di particolare interesse dell'Unione Europea che ha finanziato ben tre progetti in materia, oggetto degli ultimi tre contributi (raccolti nel punto n.15).

A corollario dell'opera, la postfazione offre una visione critica sul futuro della trascrizione. Da essa emerge chiaramente la necessità di un rinnovamento del mercato in modo da aprire una nuova domanda, che generi nuovi posti di lavoro per un'offerta altamente specializzata che sia in grado di fornire non solo un'importante quantità di lavoro, ma soprattutto un servizio che sia impeccabile dal punto di vista della corrispondenza alle esigenze del cliente che la qualità impone.

Tecnicamente, l'indice del numero monografico riporta il titolo di ogni Parte (con numerazione romana) e di ogni articolo/sezione (con numerazione araba). Nel caso di sezioni, che sono composte da diversi articoli, è indicata per ogni articolo una sotto numerazione alfanumerica (es. 15a). Cliccando su ognuno di essi si viene portati tramite link ipertestuale alla sezione, al capitolo o all'articolo corrispondente. Per ogni articolo è citato il nome dell'autore o degli autori. Cliccando sul nome si accede a una breve nota biografica a sostegno della validità del lavoro. La bibliografia di ogni articolo non segue l'articolo, ma è stata raggruppata in fondo al volume per garantire una maggiore completezza della stessa e per fornirne una più facile consultazione a chi volesse allargare i propri ambiti di interesse.

[Torna all'indice](#)

Vi auguriamo una buona lettura.

Carlo Eugeni e Luigi Zambelli

[Torna all'indice](#)

PARTE I: INTRODUZIONE AL RESPEAKING

1. [Introduzione al respeaking \(Carlo Eugeni\)](#)
2. [Funzionamento di un software di riconoscimento del parlato \(Antonio Ingrassia\)](#)
3. [I benefici del respeaking per i sordi \(Tiziana Trapani\)](#)
4. [La storia del respeaking \(Giuliano Pirelli\)](#)
5. [La trascrizione tramite respeaking in Intersteno \(Fabrizio G. Verruso\)](#)
6. [La sottotitolazione tramite respeaking nel mondo \(Pablo Romero Fresco\)](#)

1) INTRODUZIONE AL RESPEAKING

Carlo Eugeni

1. Descrizione

Il respeaking, o rispeakeraggio o ancora respeakeraggio, è una tecnica di trascrizione che si basa sulla dettatura al computer. Per dettatura al computer si intende la pratica per cui si parla a un microfono e un software – detto di riconoscimento del parlato o ASR (Automatic Speech Recognition) o ancora TAL (Trattamento Automatico del Linguaggio) o impropriamente di riconoscimento vocale – trascrive quanto è stato dettato. Se si detta al software ripetendo, riformulando o traducendo ciò che dice qualcun altro allora si sta facendo respeaking. Nel primo e nel secondo caso chi detta, cioè il respeaker o rispeaker, fa respeaking intralinguistico, nel terzo fa respeaking interlinguistico. Il respeaker può lavorare sia in tempo reale o online, cioè simultaneamente all'ascolto del testo orale da trascrivere o tradurre, oppure in preregistrato o offline, cioè fermando il testo registrato e dettandolo o traducendolo dopo averlo memorizzato, con l'ulteriore possibilità di riascoltare i pezzi che non ricorda.

2. Applicazioni e tecnologia

Le applicazioni del respeaking sono numerose. Si può usare per redigere verbali o resoconti, ma anche per produrre sottotitoli televisivi, di conferenze o di messe religiose. Il software commerciale più diffuso è Dragon NaturallySpeaking, della Nuance. Si tratta di un software *speaker dependent*, cioè che ha bisogno che chi lo usa lo addestri a riconoscere la propria voce. Il software crea così un profilo vocale personalizzato. Se il software viene usato da qualcuno che non ha provveduto a creare o esportare il proprio profilo vocale sul quel software allora il riconoscimento del parlato sarà molto poco accurato. Esistono anche software cosiddetti *speaker independent* che invece non hanno bisogno che si crei un profilo vocale e riconoscono il parlato di chiunque. Costano di più e sono meno diffusi, ma permettono di essere usati da chiunque senza che si crei il profilo vocale.

3. Storia e pratica

Il rispeakeraggio, in Europa e in una certa misura in tutto il mondo, è diventato prassi consolidata per la produzione di sottotitoli in diretta per sordi. Le sue origini sono da ricercare in parte negli studi informatici sul riconoscimento del parlato, che vanno ancora avanti per raggiungere risultati ancora non del tutto consolidati, come il riconoscimento automatico della punteggiatura (che ancora deve essere dettata perché sia scritta dal software) o addirittura la trascrizione automatica, cioè la produzione in tempo reale del testo da trascrivere senza la presenza del respeaker. Alcuni progetti in corso puntano addirittura alla traduzione automatica in tempo reale del parlato, grazie alla combinazione tra software ASR e software di traduzione automatica. Ma l'origine del respeaking, cioè la trascrizione o la traduzione da parte di un respeaker di un testo prodotto oralmente da qualcun altro, è da ricercare nel team del progetto

VOICE della Commissione Europea coordinato dall'ing. Giuliano Pirelli, che verso la fine degli anni 1990 giunse al primo prototipo di software per la sottotitolazione in tempo reale tramite respeaking. Il passo verso le applicazioni in TV non si è fatto attendere. Nel 2001 Chris Ales negli Stati Uniti, la BBC nel Regno Unito e la VRT nelle Fiandre lo introdussero per sottotitolare in tempo reale i programmi più semplici e adatti alla sperimentazione.

4. La ricerca

Perché anche la ricerca universitaria si interessasse al respeaking bisogna aspettare il 2006, quando il primo International Seminar on Intralingual Live Subtitling (www.respeaking.net) fu organizzato all'allora SSLMIT dell'Università di Bologna a Forlì, seguita dalla seconda edizione alla UAB di Barcellona (2009), dalla terza all'Artesis College di Anversa (2011) e dalla quarta sempre alla UAB (2013). Da allora *respeaking* e *live subtitling* (cioè la sottotitolazione in diretta) sono temi ricorrenti nelle maggiori conferenze internazionali sulla traduzione audiovisiva (TAV) come *Languages & The Media* (2006, 2008, 2010, 2012) e *Media for All* (2007, 2009, 2011, 2013). La prima pubblicazione accademica sul respeaking è la tesi di Master di Alison Marsh (2004), dell'allora squadra di sottotitolatori della BBC, seguita dalla pubblicazione del primo volume interamente dedicato al respeaking (a cura di Eugeni e Mack 2006) e da numerose altre pubblicazioni di articoli, tesi di laurea (tra gli altri Russello 2009, Vazquez 2009, Marchionne 2011, Bellucci 2012, Trapani 2012), tesi di dottorato (Eugeni 2009) e addirittura monografie (Eugeni 2008, Romero-Fresco 2011).

5. La formazione

Per quanto riguarda la formazione, i primi insegnamenti in materia di respeaking offline risalgono al 2003 presso l'ISIT di Trento, seguiti poi dai corsi extracurricolari sul respeaking online presso le SSLMIT di Forlì e di Trieste e infine da quelli curriculari sempre sul respeaking online della SSLMIT di Forlì, dell'Universitat Autònoma de Barcelona, dell'Artesis University College di Anversa, della University of Leeds, della Universidade de Vigo e di altri istituti ancora oltre a diversi corsi post-laurea, tra cui ricordiamo quello dell'ISIT di Trento che dal 2009 forma studenti al TAL. Prima del 2003, le società non avevano altra scelta se non formare esse stesse il proprio personale. Nessun corso (neanche codici di condotta o linee guida) erano disponibili a livello accademico, con l'eccezione di workshop occasionali durante le conferenze. Alcune aziende che utilizzano il respeaking per la produzione di trascrizioni o di sottotitoli live si occupano ancora di formare il proprio personale, ma molti altri ora assumono laureati provenienti da questi e altri corsi universitari. Così come avviene anche con la ricerca accademica, la formazione accademica, dopo aver esplorato a vista l'ambito del respeaking, sembra oggi essersi messa in carreggiata come disciplina scientifica a tutti gli effetti e soprattutto corre di pari passo con l'industria guardando avanti testando le innumerevoli possibilità che questa tecnica permette.

2) FUNZIONAMENTO DI UN SOFTWARE DI RICONOSCIMENTO DEL PARLATO

Antonio Ingrassia

Introduzione

Il riconoscimento vocale è il processo per cui il linguaggio umano prodotto oralmente viene riconosciuto ed elaborato da un software detto “di riconoscimento vocale”. Le applicazioni sono numerose e vanno dalla telefonia (risponditori, call center automatici ecc.) ai comandi vocali per la macchina, dal controllo dei sistemi di navigazione satellitare alla sincronizzazione, dalla sicurezza al *media indexing*. Negli ultimi dieci anni sono state poi sviluppate applicazioni ormai del tutto affidabili anche per la dettatura e la sottotitolazione tramite respeaking che corrono parallelamente allo sviluppo di software con risultati sempre migliori per la trascrizione, la sottotitolazione e la traduzione automatiche.

Questi ultimi software, detti più propriamente di riconoscimento del parlato, riconoscono non tanto la voce, quanto appunto il parlato. Non operano cioè una trascrizione fonetica, suono per suono, della voce che sentono, ma la segmentano sulla base di un database, il vocabolario. In definitiva, della voce vengono trattenute solo le caratteristiche verbali (le parole) e non quelle extra- e paralinguistiche (fischi, pause piene, intonazione, prosodia, timbro di voce, difetti di pronuncia ecc.). Tecnicamente, tuttavia, queste variabili hanno un effetto sulla qualità del riconoscimento del parlato, in cui si devono tenere in considerazione molte variabili, legate alla natura del segnale (ampiezza, spettro, prosodia, interazioni acustiche con l'ambiente), alle unità linguistiche (frasi, parole, sillabe, morfemi, fonemi) e all'oratore stesso (genere, età, accento, dizione, velocità, stato emotivo, ecc.) (Baker et al., 1984).

Aspetti informatici

Lo scopo di questi riconoscitori automatici del parlato (ASR) è quello di creare un ‘modello percettivo’ del sistema uditivo umano (McGurk e McDonald, 1976), ossia un emulatore che cerca di gestire l'enorme variabilità del parlato spontaneo e di elaborarlo in maniera il più possibile simile al cervello umano e discriminando, laddove possibile, la voce dal rumore ambientale (Massaro, 1972). Altri ASR, invece, si indirizzano unicamente al parlato dettato, in un ambiente con un basso, o ben definito, livello di rumore. I problemi che un sistema del primo tipo si trova ad affrontare sono molto complessi e a tutt'oggi non si può parlare di soluzioni robuste neanche per i casi più semplici.

Tra gli sforzi operati nella creazione di un modello percettivo sono da evidenziare due fondamentali approcci: in alcuni casi una parola è riconosciuta a partire dai foni costituenti, in altri a partire dalle sillabe, ma entrambe le metodologie non hanno raggiunto le prestazioni di un essere umano. In questo

[Torna all'indice](#)

panorama ci sono pochi riferimenti, invece, a una visione differente del problema che focalizzi, per esempio, il riconoscimento di segmenti di segnale su più livelli linguistici e non su uno solo. Un sistema del genere potrebbe sfruttare le caratteristiche degli approcci descritti sopra per ottenere un ASR con prestazioni più alte. Questa idea ha radici in esperimenti percettivi su soggetti umani (Greenberg, 1996) e le poche implementazioni sembrano confermarla.

Quanto alla metodologia adoperata nella costruzione di un ASR, sebbene i dettagli delle implementazioni pratiche differiscano tra loro, essa è quasi sempre la stessa. Innanzitutto il segnale del parlato in ingresso viene suddiviso in porzioni proporzionali all'entità linguistica in esame (per esempio circa 250 ms nel caso delle sillabe), allo scopo di generare dei vettori numerici, o features (Martin, Nelson e Zadell, 1964) legati ad alcune caratteristiche fondamentali e discriminanti dell'entità. Questi vengono successivamente riconosciuti come appartenenti a una 'classe' di unità grammaticali da un sistema specifico (per esempio una rete neurale (Miller e Gorin, 1993) o un modello nascosto di Markov, o HMM (Wilpon, Rabiner, Lee e Goldman, 1990). Un ultimo blocco si occupa poi di ricostruire le parole a partire dalle classificazioni ottenute, servendosi di una grammatica e di un modello di pronuncia (Andreou, Kamm e Cohen, 1994).

Il risultato del riconoscimento, però, è variabile. Il divario tra prestazioni umane e artificiali ha un'enorme varietà di cause, alcune delle quali ignote, altre dipendenti da un'evidente incapacità, da parte degli ASR, di mantenere costanti le prestazioni anche in presenza di interferenze ambientali, al cambiare del parlante e della frequenza di campionamento del segnale in ingresso. È in questo ambito che si colloca l'importanza di avere una buona codifica del segnale, in quanto è intuibile che se l'informazione ricavata da quest'ultima prescindesse dalle alterazioni elencate, anche il riconoscimento risulterebbe più robusto (Nadas, Nahamoo e Picheny, 1988).

Aspetti tecnico-tecnologici

Parlare è un fenomeno complesso. Secondo gli studi sul linguaggio e la comunicazione, nel processo di riconoscimento del parlato sono coinvolti 4 stadi:

- *la ricerca della parte acustico-fonetica per formare parole;*
- *l'analisi sintattica e grammaticale;*
- *la determinazione semantica;*
- *il disambiguamento tramite l'inferenza pragmatica e l'interpretazione dell'atto illocutorio.*

In ambito ingegneristico, le tecnologie del riconoscimento parlato si sono notevolmente sviluppate raggiungendo l'obiettivo della conversione in singole parole del segnale orale (Juang, 1985). Molto più complessa e spinosa è però

[Torna all'indice](#)

l'inferenza e la generazione di intere frasi. In questa formulazione una sorgente di messaggi trasmette un messaggio M , che è reso come sequenza W di parole attraverso il canale linguistico, specificato da una probabilità $P(W|M)$ come sintetizzato in tabella 1.

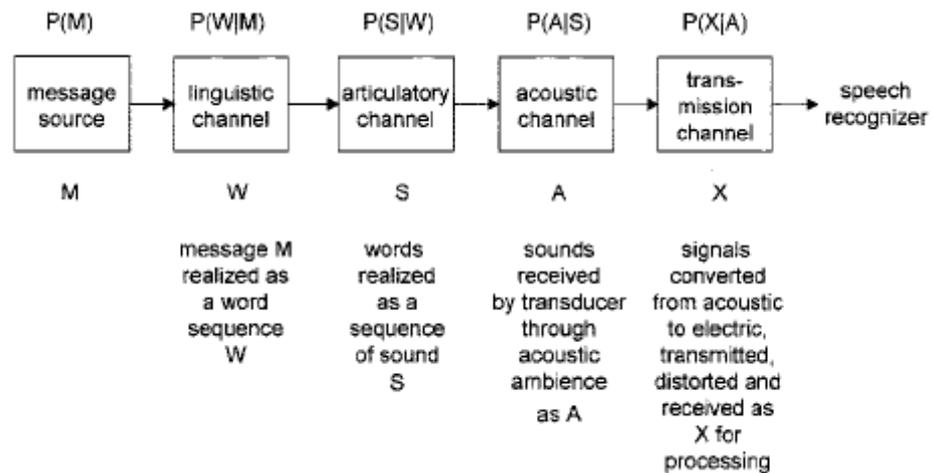


Tabella 1: Formulazione teorica della catena di produzione del parlato[23].

La sequenza W di parole diventa durante l'articolazione una sequenza S di suoni nel canale articolatorio specificato da una probabilità $P(S|W)$. Ciò perché ogni parlatore ha un suo apparato fonatorio con le sue proprie caratteristiche e non produce le medesime onde di un altro (Rabiner e Juang, 1993). La sequenza di suoni attraversa quindi l'aria, si mescola con il rumore dell'ambiente e raggiunge il microfono come input A . Questo passaggio è chiamato canale acustico A e la sua probabilità è $P(A|S)$.

In una seconda fase, il segnale acustico è convertito in elettrico e recepito dal sistema di riconoscimento. La risposta può variare molto da microfono a microfono, fino a 40 dB alla frequenza di 4 kHz, rendendo inaffidabile e fuorviante il confronto tra modelli del parlato. Questo canale di trasmissione X genera uno stadio di incertezza ed è caratterizzato da una probabilità $P(X|A)$. L'obiettivo è risalire ad M , il quale può avere una sua probabilità $P(M)$. Questa struttura è utilizzata nelle varie applicazioni per cui il riconoscitore viene utilizzato (Juang, 1996).

Funzionamento del software ASR

Lo stato dell'arte del riconoscimento può essere descritto in base alla dimensione del vocabolario riconoscibile (Moore, 1997), oppure in base allo stile del parlato, il più arduo dei quali è quello spontaneo non preparato. Esistono due approcci allo sviluppo di software di riconoscimento del parlato spontaneo, il primo è acustico-fonetico e l'altro è statistico. L'approccio acustico-fonetico del riconoscimento del parlato rappresenta la base di molti sforzi classici di ricerca.

[Torna all'indice](#)

Esso ha portato a buoni risultati, ma limitati. Il motivo sta nel fatto che si basa su una conoscenza scientifica, pur sempre limitata, del segnale acustico e sul confronto con modelli creati su suoni prodotti da professionisti della voce, spesso in ambienti protetti 'privi' di rumore. L'approccio statistico ha invece trovato largo impiego negli ultimi anni. Questo tipo di formulazione ha le sue radici nella teoria delle decisioni di Bayes.

La teoria delle decisioni di Bayes trasforma il problema di riconoscimento in un problema di stima della distribuzione e le conoscenze richieste possono essere apprese dai dati e non da deduzioni di esperti. L'approccio statistico al riconoscimento di modelli può essere quindi stabilito come segue:

- *raccolta e classificazione di un set di osservazioni;*
- *stima delle distribuzioni di probabilità a posteriori, per implementare la decisione sul valore massimo per avvicinarsi il più possibile al minimo del rischio di Bayes.*

Deve quindi essere trovato un buon metodo di stima dei parametri. L'approccio richiede infatti un set di prova di esempi noti sufficientemente grande e, statisticamente, maggiore è, migliore sarà la stima che si ottiene.

Operativamente, nell'analisi del segnale parlato, fonemi e parole non possono essere misurati direttamente, ma è necessario fare ipotesi e prendere decisioni. D'altro canto, il parlato è un segnale tempo-variante. In altre parole, quando parliamo, l'apparato fonatorio modula la pressione dell'aria, producendo una sequenza udibile di suoni. Nel breve intervallo in cui tale configurazione rimane costante – questa configurazione articolatoria non sopporta più di dieci forti cambiamenti al secondo – può essere osservata una regione di 'quasi-stazionarietà' nel segnale parlato prodotto, fondamentale per la distinzione linguistica. Certe variazioni temporali, invece, sono irrilevanti per la distinzione linguistica. Gli scopi delle rappresentazioni del parlato per il riconoscimento devono tenere questo in considerazione e includere l'abilità di conciliare variazioni irrilevanti con quelle rilevanti o eliminarle.

Quindi, il parlato non è un processo senza memoria visti i vincoli articolatori e fonotattici. In accordo con le regole fonologiche del linguaggio, infatti, c'è una certa dipendenza tra coppie di suoni che si presentano in una sequenza: alcune sono più frequenti di altre, altre addirittura vengono date per non esistenti. Il modello del parlato deve permettere la caratterizzazione di questa struttura sequenziale, idealmente in modo congruente con la quasi-stazionarietà del segnale parlato.

[Torna all'indice](#)

Articolazione del modello del parlato

Le caratteristiche

Le caratteristiche che il modello del parlato deve avere sono tre. In primo luogo, in un intervallo di 10 ms le misure nel tempo breve devono essere rilevate lungo le dimensioni pertinenti il parlato che al meglio trasportano l'informazione rilevante per la distinzione linguistica. Queste dimensioni determinano lo spazio di osservazione in cui la distribuzione è da definire. Questo è portato a termine nell'analisi del segnale e nella scelta della rappresentazione.

In secondo luogo, l'esistenza della regione quasi-stazionaria suggerisce che le misure nel tempo breve, cioè dell'ordine dei 100 ms, devono essere simultaneamente considerate come un gruppo di osservazioni distribuite indipendentemente e identicamente o come un segmento di un processo casuale forse non stazionario che copra due regioni quasi stazionarie.

In terzo luogo, è necessario un meccanismo che descriva il comportamento della variazione di suono lungo i segmenti sonori nell'enunciazione (D'Ydewalle, 1991). Questa caratterizzazione tiene in conto la struttura implicita dell'enunciazione nel senso di una distribuzione di probabilità.

I modelli ASR

L'uso di un modello nascosto di Markov, o HMM (Ferguson, 1980) è una scelta che porta a buoni risultati. Normalmente un modello o una distribuzione classe-dipendente è stimato in base ai dati della stessa classe in accordo con criteri statistici di stima ben noti: quello di massima verosimiglianza, o ML, ovvero 'maximum likelihood' (Juang, 1985), è uno dei più importanti. Il metodo ML non è tuttavia l'unico modo per risolvere il problema della stima. Esso non porta infatti solitamente al minimo tasso d'errore nel riconoscimento e ciò è dovuto alla limitatezza dei dati di addestramento e al disaccoppiamento tra la forma scelta di distribuzione e i dati correnti del parlato.

Un tentativo ulteriore è stato formulare il problema del progetto del classificatore come minimizzazione del tasso di errore di classificazione. Questo approccio è chiamato 'addestramento discriminante' e il suo obiettivo è discriminare correttamente le osservazioni per i risultati del migliore riconoscimento piuttosto che adattare le distribuzioni ai dati. L'obiettivo del classificatore è ancora raggiungere la minima probabilità di errore basata sulla funzione di perdita. In questo approccio, la difficoltà sta nel trovare una funzione obiettiva che sia consistente con il tasso di errore e anche adatta all'ottimizzazione (Sakoe e Chiba, 1978).

Unità del riconoscimento

In un parlato continuo occorrono molte volte fenomeni di coarticolazione. Quando viene pronunciata una sequenza di suoni, il nostro apparato fonatorio

[Torna all'indice](#)

normalmente non produce ogni singolo suono, ma aggiusta la propria configurazione anticipando i suoni successivi o modificandoli di fatto. Una frase pronunciata colloquialmente è spesso pronunciata in modo diverso da come indicato sul dizionario. Maggior complessità del sistema è richiesta al crescere della variabilità e dei parametri. La complessità è sostanzialmente proporzionale al numero di classi da riconoscere. Per il riconoscimento del parlato continuo il numero delle classi dipende dalla scelta dell'unità. Se essa è la parola singola allora, in un vocabolario di 20.000 parole, sono necessarie almeno 20.000 modelli per le parole e le prove durante l'addestramento possono essere milioni. Altre unità possono essere le sillabe (nell'ordine di decine di migliaia); i fonemi, cioè i suoni che compongono le parole (per l'italiano sono 30, visto che ogni lettera può essere letta in diversi modi a seconda delle altre lettere con cui va a formare una sillaba); o le emisillabe, cioè metà sillaba, indipendentemente dal fatto che essa sia composta da due o quattro fonemi (circa 2000) (Shire, 1997).

Una scelta diffusa consiste negli pseudofonemi che possono essere considerati fonemi con variazioni acustiche dovute ai vari contesti. Infatti, nonostante si sia detto che in italiano esistono 30 fonemi, gli pseudofonemi usati in ASR sono migliaia perché ogni fonema si pronuncia diversamente a seconda della sillaba in cui è inserito. Per esempio il fonema /p/ in "capire" non è pronunciato alla stessa maniera dello stesso fonema (detto anche allofono cioè suono diverso ma che per convenzione viene riconosciuto come variante dello stesso fonema) in "prato". Quando vengono usate unità dipendenti dal contesto, il modello di ogni pseudofonema nella rappresentazione necessiterà di avere la designazione contestuale. A questo punto si pone il problema di trovare un compromesso tra alta risoluzione acustica e basso rendimento della stima. Questo è un problema che necessita di particolare cura nel modellare un parlato ad ampio vocabolario. Una soluzione è considerare la dipendenza dal contesto al livello di 'quinfono', cioè uno pseudofonema con la dipendenza dai due fonemi precedenti e dai due seguenti. Con l'alto numero di potenziali quinfoni, la selezione di una unità diventa critica per evitare la perdita di affidabilità nei risultati stimati, controbilanciando il guadagno dell'accresciuta risoluzione acustica (Furui, 1997).

È usata spesso a tale scopo la verosimiglianza differenziale o la distanza Kullback-Liebler per HMM (Rabiner, 1989). Altra tecnica per guadagnare affidabilità nella stima dei parametri è l'uso dell'allacciamento. Esso forza i modelli di differenti unità a condividere i parametri e parte della topologia di rete, riducendo così il numero totale dei parametri di modello. Il confronto tra modelli si riferisce alla procedura computazionale di valutare la similarità tra modelli che sono da confrontare. La base del confronto di modelli è sottolineare la misura della distorsione che il progettista del sistema sceglie a priori. Una misura della distorsione è spesso associata con qualche nozione nelle distribuzioni statistiche e la quantità di distorsione può essere interpretata come un logaritmo negativo di verosimiglianza. Alternativamente si può anche

[Torna all'indice](#)

misurare la dissimilarità tra modelli attraverso la misura della distorsione oppure la misura della verosimiglianza di una sorgente rappresentata da un modello che ha prodotto il modello del parlato osservato, seguito da una ricerca e da un processo di decisione.

Scelta delle parole da riconoscere

Due tecniche base sono disponibili per il confronto di modelli del parlato. Uno è chiamato algoritmo di Viterbi e l'altro è l'algoritmo avanti-indietro. Queste due tecniche corrispondono alla misura di due quantità essenziali nella formulazione di un modello statistico per il parlato. Valutare la dissimilarità tra due modelli del parlato è più complesso della maggior parte della tecnica del confronto di modelli che opera in un semplice spazio vettoriale a dimensione fissata. Un fattore principale che contribuisce a questo è costituito dal modello del parlato come sequenza temporale che può essere compressa o allargata nel tempo a causa dei cambiamenti nella velocità del parlato. Un tale cambiamento risulterebbe usualmente in una durata non uniforme negli elementi fonemici dell'enunciato durante l'articolazione. Per ottenere un calcolo significativo del confronto di modelli tra due modelli del parlato di ineguale lunghezza, è necessaria qualche normalizzazione. Ciò dà luogo all'uso di una distorsione tempo-dinamica nel definire una misura della dissimilarità, oltre al calcolo della distanza tra due rappresentazioni vettoriali del parlato nel tempo breve.

Per il riconoscimento del parlato continuo ad ampio vocabolario, modelli composti comprendenti sequenze di modelli unitari sono usati per il confronto di modelli. Durante la decodifica, il sistema deve ipotizzare un numero di tali modelli composti per la valutazione della verosimiglianza e la decisione del riconoscimento. Poiché il numero di modelli unitari può essere enorme, una semplice elencazione basata su combinatoria esaustiva produrrà un numero proibitivo di sequenze ipotizzate. Anche con l'aiuto di un modello del linguaggio atto a mostrare le ipotesi di frase, la lista può essere consistente e comprendere in nuce il potenziale tranello rappresentato dall'esclusione delle ipotesi corrette. Una strategia di ricerca che combini tutti i livelli di verosimiglianza e li usi per guidare il processo di potatura delle ipotesi è critica in qualsiasi implementazione di un sistema di riconoscimento del parlato ad ampio vocabolario (Petrillo, 1999).

Per cercare di risolvere tali ostacoli è necessario utilizzare degli algoritmi di ricerca ad hoc. Gli algoritmi di ricerca possono essere categorizzati in due tipi essenziali: il primo migliore ed il primo largo. Un algoritmo chiamato ricerca della larghezza massima è stato ampiamente usato con buoni risultati. Questo cerca lo spazio di ipotesi solo attorno al percorso attuale migliore per ridurre l'entità del calcolo di verosimiglianza. Poiché il percorso attuale migliore può non essere il migliore in assoluto, altre strategie possono essere incorporate per ridurre il potenziale errore di ricerca.

[Torna all'indice](#)

Un altro algoritmo di ricerca usato è l'algoritmo dei migliori N, che produce non solo il migliore, ma gli N migliori candidati. L'algoritmo ad albero a trellis è particolarmente efficace. Con una lista delle N migliori ipotesi, si possono invocare il livello più alto oppure una conoscenza indipendente per eseguire una strategia multipasso per migliorare il risultato della ricerca.

3) I BENEFICI DEL RESPEAKING PER SORDI

Tiziana Trapani

1. Breve introduzione alla sordità

Quella che comunemente viene chiamata sordità, da un punto di vista strettamente clinico viene definita ipoacusia e indica, in linea generale, un deficit della funzione uditiva. A seconda del tipo di patologia, la sensibilità uditiva viene compromessa in maniera diversa: si passa, infatti, da casi in cui il deficit uditivo non è molto grave ed è in buona parte reversibile, a casi in cui l'alterazione della percezione sonora è superiore e può raggiungere livelli estremi di sordità. I rapporti sociali, affettivi e di relazione con l'ambiente sono sempre stati di fondamentale importanza per l'uomo che vuole integrarsi completamente nel mondo in cui vive: la sordità, da sempre, è vissuta come isolamento, emarginazione e handicap. Ecco perché la limitazione degli effetti emarginanti appare un imperativo categorico, specie nell'era moderna dove i progressi della tecnologia hanno cambiato profondamente la comunicazione tra individui, anche appartenenti a lingue e culture diverse.

2. Il respeaking nei media

Per comprendere appieno le potenzialità della tecnica del respeaking e conseguentemente le sue possibili applicazioni per favorire l'integrazione delle persone sorde, bisogna innanzitutto interrogarsi su quali siano le difficoltà quotidiane che si trovano ad affrontare i principali destinatari del prodotto finale, ossia il testo generato dal processo di rispeakeraggio.

All'interno di una società frenetica e basata sulla sostanziale spersonalizzazione dell'individuo, un'importanza fondamentale nella vita quotidiana è costituita dai mezzi di comunicazione di massa. Tra di essi, la televisione rappresenta senza ombra di dubbio quello più diffuso e apprezzato: la semplicità d'uso e l'attuale basso costo l'hanno portata ad affiancare sempre più efficacemente la stampa e la radio come fonte di informazione e, soprattutto, di svago, grazie alla vastità dell'offerta di trasmissioni. I problemi di natura sensoriale derivati dalla sordità non consentono di avere pieno accesso alla componente audio della programmazione televisiva: questo mina la capacità di comprensione e, di fatto, esclude le persone sorde da un buon numero di eventi. Proprio in virtù di questo, la sottotitolazione intralinguistica in tempo reale tramite respeaking risulta un mezzo efficace per superare le barriere imposte da un problema uditivo, consentendo non solo l'accessibilità a contenuti informativi o ludici preregistrati, ma anche a tutti quegli eventi in real-time che permettono alla persona sorda di non sentirsi esclusa da una parte importante della realtà quotidiana.

3. Il respeaking nel resto della società

Tuttavia la sottotitolazione tramite respeaking può essere applicata, fondamentalmente, a qualunque altra forma di comunicazione verbale che necessita di essere scritta per poter risultare fruibile anche a quelle parti della popolazione altrimenti escluse. Pensiamo, per esempio, ai luoghi pubblici e alla difficoltà di una persona sorda a comunicare con gli utenti che si trovano dietro allo sportello di una A.S.L., di un ufficio postale, di uno sportello bancario, dell'accettazione del pronto soccorso o anche, più semplicemente, con i commessi di un negozio. Si tratta di operazioni quotidiane che spesso rivestono un'importanza fondamentale ma che risultano estremamente complicate per chi non ha la piena facoltà di comunicare. Basterebbe introdurre, perlomeno all'interno degli uffici pubblici, un meccanismo che implementi il respeaking delle parole pronunciate dall'operatore all'automatizzazione delle risposte della persona sorda: si potrebbe, per esempio, prevedere uno sportello riservato a utenti con problematiche uditive, dotato di un piccolo schermo su cui proiettare le domande dell'operatore e di una tastiera attraverso la quale l'utente sordo possa dare le risposte.

Analogamente, rimanendo nell'ambito della sordità e dell'accessibilità a tutte quelle operazioni quotidiane che costituiscono la "normalità" della vita di un individuo, non si può non considerare l'ambito scolastico e le difficoltà di gestione a ogni livello di formazione. Infatti, che si parli di scuola elementare o di università, la difficoltà maggiore che una persona sorda si trova a fronteggiare è la comprensione di quello che viene spiegato dagli educatori. Introdurre all'interno del sistema scolastico una figura professionale qualificata che, tramite il respeaking, permetta allo studente di seguire in tempo reale la lezione, ne diminuirebbe senza ombra di dubbio il disagio, contribuendo al consolidamento della propria autostima e del senso di integrazione. Alcune università, infatti, per ovviare a questo problema propongono l'affiancamento di un tutor prendi-appunti: uno studente o un ex studente che si reca a lezione insieme allo studente sordo (o in alcuni casi in sostituzione dello stesso) e che prende appunti al posto suo, aiutandolo, se necessario, anche nella fase successiva di studio. Ma questo impedisce, fondamentalmente, allo studente sordo di essere autonomo e lo vincola alla presa di appunti di qualcun altro, oltre a non favorire affatto l'integrazione con gli altri studenti dato che ci si rapporta sempre a un'unica persona: vengono così a mancare l'interazione con gli altri colleghi di corso e la conseguente socializzazione.

Il medesimo discorso può essere applicato all'ambito lavorativo: il professionista sordo che lavora all'interno di un'azienda ha bisogno, per poter svolgere al meglio il proprio incarico, di avere pieno accesso alle informazioni. Questo si traduce, in primis, nella necessità di potere interagire con colleghi all'interno delle riunioni gestionali: l'introduzione di un sistema applicativo che preveda il respeaking dei dialoghi velocizza e migliora la comprensione da parte

[Torna all'indice](#)

del lavoratore sordo e facilita il conseguente scambio di opinioni. Il potenziale della pratica del *respeaking* appare inoltre chiaro anche quando ci si sofferma a pensare alla miriade di eventi organizzati all'interno dei centri urbani, partendo dai concorsi di ogni genere (letterari, artistici, di bellezza, musicali ecc.) fino ad arrivare ai seminari e alle conferenze di stampo accademico, passando attraverso le manifestazioni politiche. Oltre al vantaggio primario di rendere accessibili alle persone sorde delle attività culturali e/o ludiche che favoriscono il senso di appartenenza a una comunità, vi è il non trascurabile beneficio secondario di avere a disposizione la trascrizione dell'intero evento in tempi molto brevi, dopo una semplice revisione finale, con il conseguente risparmio in termini di tempo e denaro.

Ogni forma di discorso orale può essere trascritta in maniera praticamente istantanea tramite il *respeaking*, comprese, per esempio, le conversazioni spontanee tra amici. Pensiamo a una semplice situazione quotidiana in cui un gruppo di amici decide di trascorrere una serata in pizzeria. Normalmente la comunicazione tra sordi e udenti sarebbe vincolata ad alcuni fattori: innanzitutto la capacità degli udenti di discorrere tramite la Lingua dei Segni (se questa rappresenta il mezzo attraverso il quale si esprimono i sordi presenti all'interno del gruppo); il possesso da parte del sordo di un impianto cocleare o di una protesi acustica di qualsiasi tipo che ne faciliti la comprensione dei discorsi; il grado di riabilitazione logopedica cui il sordo è stato sottoposto e che gli permette di esprimersi in maniera più o meno comprensibile tramite la parola; la capacità di leggere i movimenti delle labbra e quindi comprendere quello che viene detto. È logico che questo crea, inevitabilmente, delle limitazioni. Se si potesse introdurre il *respeaking*, o anche semplicemente la trascrizione automatica, all'interno di situazioni di questo tipo si favorirebbero anche i rapporti sociali tra sordi e udenti, vincendo quel disagio e quella frustrazione che entrambe le parti provano di fronte alla difficoltà di comunicazione.

4. Conclusioni

Oggi l'approccio che ogni persona ha nei confronti della realtà che la circonda è strettamente collegato alla velocizzazione dell'acquisizione delle conoscenze: il lavoro, gli svaghi, le modalità con cui le scuole educano i giovani sono inseriti in un contesto di multimedialità che offre l'opportunità di conoscere notizie e immagazzinare dati a cui prima si accedeva molto più lentamente. La distanza fisica non è più un fattore fondamentale ai fini della comunicazione: si può lavorare, visitare una mostra dall'altra parte del mondo o ascoltare un concerto senza muoversi dalla propria abitazione. Appare abbastanza semplice, dunque, capire che, sebbene la sordità sia in qualche modo avvantaggiata dal mondo globale e multimediale, rimane per molti aspetti una condizione fortemente invalidante per chi la vive. Basti riflettere sul fatto che l'affermazione della famiglia nucleare su quella patriarcale ha notevolmente limitato

[Torna all'indice](#)

l'accoglienza e la condivisione dell'handicap in un ambito sociale più ristretto: l'individuo, in generale, è esposto al giudizio continuo e all'accettazione da parte di un mondo complessivamente composto da individui tra loro estranei. In questo contesto, che favorisce l'isolamento di ogni componente sociale, risulta quanto mai importante sfruttare pienamente il potenziale della multimedialità a favore di un'integrazione completa di persone che vivono la quotidianità di questo mondo frenetico con uno svantaggio iniziale, rappresentato dal loro deficit fisico. Sebbene siano tantissimi i campi su cui si dovrebbe intervenire, la fondamentale semplicità dell'intero processo di respeaking lo rende estremamente versatile e adattabile a ogni situazione ed esigenza.

4) LA STORIA DEL RESPEAKING – EVOLUZIONE DEL PROGETTO EUROPEO VOICE

Giuliano Pirelli

1. La nascita dei sottotitoli

In uno sguardo dei cicli storici si potrebbe dire che l'invenzione del telefono costituì un grande ausilio per la comunicazione delle persone con disabilità visiva, mentre le persone con disabilità uditiva non solo non ne hanno potuto fruire, ma ne sono stati indirettamente danneggiati, perché sono risultati maggiormente isolati rispetto agli altri. Al contrario, oggi, le conversazioni via chat sono molto utilizzate dalle persone con disabilità uditiva, ma sono meno accessibili alle persone con disabilità visiva, salvo il ricorso a dispositivi di sintesi vocale.

All'origine il cinema era muto, tutt'al più i film includevano alcune didascalie esplicative per tutti chiamate intertitoli (Figura 1) e nella sala del cinematografo un pianista suonava una musicchetta di accompagnamento per alcune scene. Con l'arrivo del cinema sonoro, in molti paesi i film in lingua straniera venivano doppiati, cioè la versione originale del parlato era sostituita da versioni in altre lingue, mentre in altri paesi essi venivano proiettati in versione originale con sottotitoli. Tutto ciò era fatto tenendo conto della massa degli spettatori, senza prendere in considerazione l'aspetto della sordità.

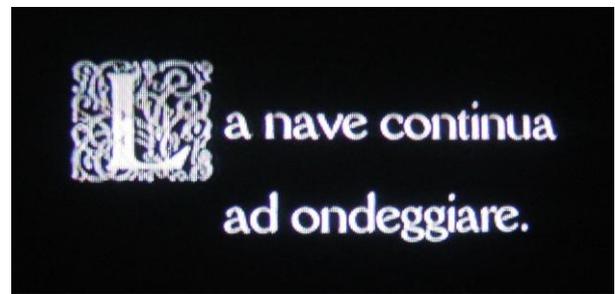


Figura 1: intertitoli cinematografici

Dietro questi approcci differenti vi erano molti elementi culturali e socio-politici relativi all'importanza attribuita alla lingua nazionale, all'apertura verso altre culture, alla dimensione del mercato e molte altre considerazioni pratiche. Tra queste tuttavia non era preponderante l'aspetto tecnico, nel senso che in fase di produzione o post-produzione di un film vi erano il tempo e la tecnologia per inserire i sottotitoli. Tali sottotitoli erano inseriti *in chiaro*, cioè *incrostat*i nell'immagine originale e dunque visibili da parte di tutti gli spettatori, più frequentemente in una sola lingua e talvolta in due lingue (per es. francese e olandese in Belgio).

2. Difficoltà di comunicazione

La disabilità uditiva comporta la mancanza, parziale o totale, di un canale di comunicazione, quello uditivo, che deve essere sostituito da quello visivo e da una marcata sensibilità *diffusa* e costante. Mentre una persona di buona volontà intuisce come aiutare una persona con disabilità motoria o visiva, che a sua volta sa indicare come può essere aiutata, viceversa pochi sanno come affrontare i problemi di comunicazione legati alla sordità. Ed inoltre, le persone sorde non sono tutte sorde allo stesso modo e sono cresciute in ambienti diversi e culture diverse. In particolare, dei diversi milioni di persone sorde presenti in Italia, la maggior parte è *oralista*, cioè riesce a comprendere quanto viene detto concentrandosi sulla lettura labiale e l'espressione del volto. I *segnanti*, cioè coloro che sono stati educati, per una serie di motivi, all'uso della *lingua dei segni*, sono un numero relativamente modesto.

Tra gli oralisti si possono includere i milioni di persone che, arrivando ad una certa età, cominciano a perdere parzialmente l'udito. Le persone di questo grande gruppo non riescono a seguire bene una trasmissione televisiva o una conferenza, ma d'altronde non si considerano nemmeno sordi e non penserebbero mai di studiare la lingua dei segni.

Il sordo non si vede. Se il numero di parti gemellari è 1 su 100, quindi se ogni 50 persone c'è un gemello, salendo su un autobus vi sono grandi probabilità d'incontrare un gemello, ma nessuno lo vede. I sordi sono più o meno lo stesso numero, orientativamente il 2 per cento o più, quindi in ogni autobus c'è una persona con sordità, più o meno grave, ma nessuno la vede, nessuno se ne rende conto, nessuno sa come aiutarla. Possiamo pensare che sia un loro problema, ma se ci mettiamo nei panni di un insegnante a scuola che deve fare lezione a un ragazzo sordo, il problema diventa dell'insegnante: il compito resta sulle sue spalle e risulta spesso molto gravoso.

3. Il riconoscimento della voce

La difficoltà delle persone con disabilità uditiva a seguire conferenze e trasmissioni televisive: spingeva i responsabili delle loro associazioni ad effettuare diversi tentativi di sottotitolarle, per es. tramite testi preparati su foglietti (Giuliano Pirelli, Natale 1984) o su rotoli trasparenti proiettati facendoli scorrere in modo continuato (Elio Parodi, giugno 1990).

Alla fine degli anni '80, nelle sedute di rieducazione ortofonica presso il Laboratorio del Policlinico di Milano si è utilizzato un sistema di riconoscimento delle intensità e delle frequenze dei suoni per visualizzare alcuni grafici sullo schermo del computer (Figura 2). Tali grafici erano d'aiuto agli specialisti, mentre dei loro adattamenti potevano generare delle immagini e giochi attraenti per stimolare la produzione di suoni e il rispetto di ritmi di pronuncia di parole da parte dell'allievo sordo.

[Torna all'indice](#)



Figura 2: sedute di rieducazione ortofonica presso il Policlinico di Milano

Nel '96, al momento della transizione dai primi prototipi del riconoscimento del parlato da parole *disgiunte* (pronunciate con una brevissima pausa fra le parole stesse) ai primi prodotti del riconoscimento del *parlato continuo*, la velocità di eloquio necessaria per essere ben riconosciuti era quella modesta permessa allora, con un certo ritardo nell'apparizione del testo. Qualche notaio e qualche avvocato iniziavano a sperimentare il sistema per dettare delle pratiche parzialmente ripetitive. Si poteva così intravedere la possibilità di un uso in tempo reale per generare dei sottotitoli, cosa che è stata realizzata grazie allo sviluppo di un'interfaccia apposita. Il testo dei sottotitoli generati era corretto, anche se richiedeva una lunga preparazione e un addestramento accurato.

Tutti coloro che si lanciavano nelle sperimentazioni in questo campo erano coscienti di quanto questa tecnologia fosse innovativa e condividevano entusiasmi e talvolta timori nell'uso di un computer relativamente grosso, non portatile ma trasportabile, sul quale caricare un programma (Dragon) tramite oltre 20 floppy disk, con la necessità di almeno una giornata di formazione dell'utente e di addestramento del software alla voce ed al vocabolario dell'utente.

4. Introduzione al Progetto VOICE

Il Progetto VOICE - "Per sostenere la VOCE dei sordi con un'azione di sensibilizzazione sulle potenzialità dei sistemi di riconoscimento della VOCE" - ha promosso lo studio e lo sviluppo di applicazioni dei sistemi di riconoscimento vocale alle conversazioni, conferenze, trasmissioni televisive e comunicazioni telefoniche. Nato nel 1996 all'interno del Centro Comune di Ricerca (CCR) della Commissione Europea, sito di Ispra (VA), negli anni 1998-2000 ha ricevuto il patrocinio ed il finanziamento della Direzione Generale Società dell'Informazione della CE.

Il Progetto ha sviluppato dei prototipi di interfacce che consentono un uso più facile di prodotti commerciali, per generare, sullo schermo di un personal

[Torna all'indice](#)

computer, sottotitoli in tempo reale di conferenze, lezioni scolastiche e universitarie, conversazioni telefoniche, trasmissioni televisive e spettacoli dal vivo (Figura 3). Questo è di notevole aiuto per gli audiolesi, poiché riduce la distanza che li separa dal mondo degli udenti.

Le prime applicazioni nella ricerca sulle tecnologie assistive e dei loro sviluppi presso il CCR hanno ottenuto risultati eccellenti e dimostrato la fattibilità di applicazioni di tecnologie a basso costo e alte prestazioni, accessibili a tutti i cittadini. Tali risultati sono stati presentati tramite una serie di conferenze sottotitolate iniziate nel 1996 presso il Laboratorio VOICE del CCR.



Figura 3: *prototipi di interfacce del progetto VOICE*

5. Un'azione di sensibilizzazione

I risultati così raggiunti in Italia dal 1996 sono stati presentati in una riunione internazionale di lancio del Progetto VOICE, organizzata a Vienna nel 1998 in collaborazione con l'Università di Linz. Successivamente sono stati organizzati oltre cento seminari per presentare a più di 6000 utenti il Progetto e il suo prototipo dimostrativo, che è stato usato per sottotitolare parte dei seminari stessi, come opera di sensibilizzazione, dimostrazione di fattibilità e prova di validazione sul campo. Il sistema di sottotitolazione è stato utilizzato in diverse scuole con l'obiettivo di trasmettere a tutti gli alunni di una classe le stesse informazioni, con le stesse parole, nello stesso momento.

In particolare, alcune attività sono state svolte in collaborazione con la Direzione Generale XXII (Educazione) della CE per i Progetti MOISE (International Organization Model for Persons with Special Educational Needs), Netdays Europe e NEMO (Non Excluding Models for Web-Based Education).

Il sistema è stato sperimentato presso diverse scuole con risultati eccellenti in alcune scuole pilota. In altri casi, tuttavia, molti insegnanti hanno dimostrato buona volontà, ma dopo brevi periodi l'interesse spesso decadeva di fronte a difficoltà pratico-organizzative, quali la necessità d'inserire frequentemente dei nuovi termini nel dizionario. Un aspetto positivo, che per assurdo si rivelava un ostacolo, era che una semplice rielaborazione dei

[Torna all'indice](#)

sottotitoli poteva consentire la produzione rapida e completa di dispense del corso, potenzialità che richiedeva un ulteriore impegno e controllo da parte degli insegnanti. In un asilo, il sistema è stato utilizzato a livello quasi giocoso, limitandolo a riconoscere automaticamente i nomi di alcuni animali o fiori da proiettare sullo schermo.

6. Il Sito e il Forum VOICE

Il Progetto VOICE ha promosso lo sviluppo di nuove tecnologie nel campo del riconoscimento del parlato e al tempo stesso ha stimolato e incentivato l'uso delle nuove tecnologie, in particolare Internet, da parte dei non udenti. Più precisamente, un altro obiettivo del Progetto è stato quello di unire, attraverso un Forum VOICE su Internet, associazioni, ditte, università, scuole, amministrazioni pubbliche e chiunque altro (a livello nazionale e internazionale) fosse interessato al riconoscimento vocale.

Il Forum ha contribuito alla definizione di specifiche tecniche relative ai bisogni degli utenti audiolesi nel campo del riconoscimento vocale e dei sottotitoli televisivi. Tutte le fasi del progetto hanno visto la partecipazione attiva delle Associazioni di audiolesi, alle quali è stato chiesto di contribuire con commenti e suggerimenti. Le informazioni raccolte sono state diffuse su Internet insieme ai risultati e agli esempi di buone prassi, anticipando, in qualche modo il trend dei social network. Ulteriori informazioni sono disponibili sul nuovo sito del progetto VOICE: <http://www.voiceproject.eu> in fase di transizione e aggiornamento dal precedente sito del CCR:

<http://ipsc.jrc.ec.europa.eu/index.php?id=681>

7. La sottotitolazione delle trasmissioni televisive

Dal '98 è iniziata una stretta collaborazione con il Servizio Televideo della RAI e i Servizi Teletext di altre reti televisive europee, per un'azione di sensibilizzazione e sperimentazione dei nuovi sistemi. Su questa base, nel 2001-2002 le attività si sono indirizzate verso l'armonizzazione degli stili e formati dei sottotitoli televisivi, in collaborazione con l'Unione Europea delle Televisioni (UER, EBU European Broadcasting Union) e l'organismo di standardizzazione CENELEC, con il supporto della Direzione Generale Imprese della CE. Il Progetto VOICE ha contribuito alla creazione del *TV Broadcasting for All Group*, che ha lavorato alla standardizzazione dei criteri di produzione dei sottotitoli televisivi, al potenziamento del servizio di sottotitolazione di varie emittenti europee e all'organizzazione della conferenza *TV Broadcasting for All* a Siviglia nel giugno 2003. *Nell'Anno Europeo delle Persone con Disabilità 2003* diverse attività del Progetto VOICE hanno mirato a diffondere ulteriormente i risultati raggiunti, aumentando l'informazione degli utenti e la sensibilizzazione dei fornitori di servizi. L'evento conclusivo dell'Anno Europeo è stata la Conferenza *Voce e Accessibilità*, organizzata ad Ispra (Varese) nel novembre 2003.

[Torna all'indice](#)

8. Il Gruppo Inter-Servizi sulla Disabilità

Negli anni successivi il Progetto VOICE ha continuato la diffusione dei risultati del Progetto, incoraggiando la sperimentazione del sistema da parte delle associazioni di audiolesi, le scuole e gli atenei e il potenziamento del servizio di sottotitolazione da parte delle emittenti televisive. Sono state intraprese numerose attività per diffondere l'uso di tali applicazioni in vari settori della vita di tutti i giorni, come cinema, teatro, tempo libero e sport. Molte delle attività sono state svolte nell'ambito dell'Istituto per la Protezione del Cittadino (IPSC) del CCR, con particolare riferimento agli aspetti dell'informatica e in parte per i Progetti di *Ricerca Esplorativa* patrocinati dal Comitato Scientifico dello stesso CCR. Il progetto VOICE è stato un punto di riferimento per altri progetti nel campo della disabilità del CCR e il suo coordinatore ha partecipato alle attività del Gruppo Inter-Servizi sulla Disabilità (ISGD, InterService Group on Disability), creato a Bruxelles alla fine degli anni '80 con la finalità di tradurre in realtà gli sforzi della UE nella diffusione delle questioni legate alle disabilità, in linea con la Strategia Europea sulla Disabilità 2010-2020 e la Convenzione delle Nazioni Unite sui Diritti delle Persone con Disabilità.

9. Il Respeaking

L'autore della presente relazione viene da molti considerato il creatore del respeaking, essendo stato fra i primi a proporre tale soluzione. Ingegnere ricercatore presso il CCR, con l'accordo del CCR ha integrato i sistemi disponibili in quel momento, software molto elementari e delicati, con una videocamera per fare delle conferenze sottotitolate, coinvolgendo le reti televisive, in particolare la RAI e la BBC. Il Progetto VOICE, di cui era responsabile dal '96 come ricercatore e dal '98 come coordinatore europeo, ha prodotto risultati eccellenti, consentendo la sensibilizzazione di interlocutori di numerose associazioni di sordi di tutta l'Europa e del Canada, per discutere con i responsabili delle reti televisive e chiedere quanto serviva loro. Varie prove con i presentatori dei telegiornali ed altre trasmissioni televisive portavano alla definizione della figura del respeaker.

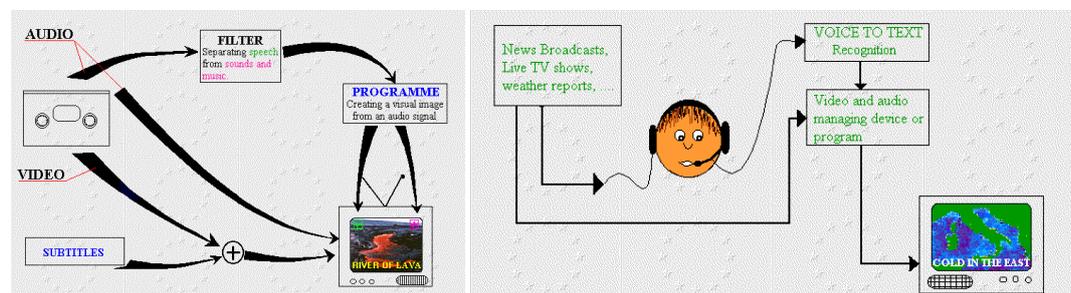


Figura 4: schema per la produzione di sottotitoli intralinguistici in televisione

In tale fase, l'azione di sensibilizzazione svolta dal Progetto VOICE si è rilevata significativa. Dopo varie considerazioni sulla possibilità di addestrare i presentatori dei telegiornali all'uso dei sistemi di riconoscimento vocale, si è preferito continuare le sperimentazioni ricorrendo ad un respeaker, cioè una persona che ascoltava il testo e lo ripeteva in modo leggermente abbreviato e

[Torna all'indice](#)

semplificato per facilitarne il riconoscimento da parte dei software di riconoscimento vocale in base al vocabolario addestrato in precedenza (Figura 4). Lo schema di questo approccio era parte integrante delle pagine pubblicate sul sito web del Progetto VOICE nel 1996. Degli ottimi risultati sono stati raggiunti sottotitolando tramite respeaking una parte di un telegiornale della BBC dedicata a un'intervista a Ronaldo. La sottotitolazione è stata effettuata dal vivo via respeaking durante la conferenza di presentazione del Progetto VOICE a Vienna nel 1998. Nel campo della ricerca il progetto VOICE ha collaborato con gli International Symposium on Live Subtitling (www.respeaking.net) e il Master on line di Traduzione Audiovisiva (METAV) dell'Università di Parma (www.metav.unipr.it) ed il Progetto SALES

Dai risultati del progetto VOICE si sono sviluppate numerose iniziative in collaborazione con diversi docenti universitari con l'obiettivo di sviluppare la figura professionale del respeaker, che usa i software di riconoscimento del parlato per produrre sottotitoli in tempo reale. In questo contesto, è particolarmente significativa la creazione dell'Associazione Internazionale di Respeaking (onA.I.R.) che ha l'obiettivo di diffondere le applicazioni della tecnologia di riconoscimento del parlato in tutti gli ambiti della vita della società, compresa la sottotitolazione di Web radio promossa all'interno del progetto OltreSuoni.

10. Il seguito del Progetto VOICE

Nel logo iniziale del Progetto VOICE del 1996, sulla parola Voice convergevano delle frecce che rappresentavano l'idea di raccogliere informazioni sul riconoscimento vocale, sulle videocamere, sui nuovi computer, sulla sordità e sui bisogni delle persone affette da sordità. Nel logo successivo, definito nell'anno 2003, Anno Europeo delle Persone con Disabilità, dalla parola Voice delle onde sonore si espandono nello spazio come raggi di sole, a testimonianza di un Voice adulto che comincia a diffondere informazioni (Figura 5).

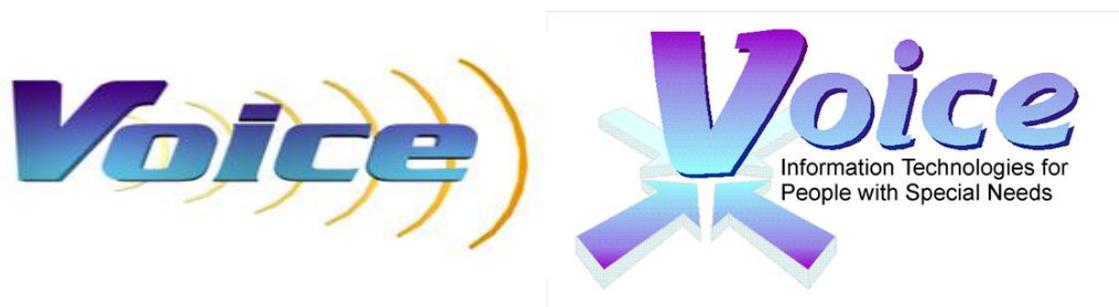


Figura 5: evoluzione del logo del progetto VOICE

Giuliano Pirelli, autore del presente articolo, coordinatore del Progetto come ricercatore e funzionario del Centro di Ricerca della Commissione Europea, ha continuato, anche recentemente come funzionario a riposo, le iniziative di sensibilizzazione e di diffusione di informazioni sul Progetto, anche come

[Torna all'indice](#)

Consigliere di ALFA (Associazione Lombarda Famiglie Audiolesi, Milano), Presidente Onorario di onA.I.R., membro onorario dell'Accademia Giuseppe Aliprandi - Flaviano Rodriguez (Ente Morale per il progresso scientifico didattico, tecnico e professionale delle scritture veloci), in stretto contatto con altre Associazioni, quali AFA (Associazione Famiglie Audiolesi, Cantù), APIC (Associazione Portatori Impianto Cocleare, Torino), TELECONTACT (Association pour la facilitation de la communication à distance entre les personnes sourdes, Bruxelles), EFHOH (European Federation of Hard of Hearing People) e IFHOHP (International Federation of Hard Of Hearing People).

11. La produzione dei sottotitoli televisivi tramite riconoscimento vocale

Con l'avvento della televisione e delle trasmissioni in diretta, la mancanza di tempo e di tecnologie appropriate si rivelavano ostacoli difficilmente sormontabili. La trasmissione di film con i sottotitoli *incrostatati*, cioè inseriti nell'immagine per tutti gli utenti era possibile ma utilizzata molto raramente. La trasmissione dei sottotitoli via Teletext iniziò in Italia nel 1986 per un numero limitato di programmi, consentendo agli spettatori di scegliere se fruirne o meno.

La sottotitolazione in diretta di eventi o di telegiornali richiedeva strumenti più rapidi, che via via sono stati sviluppati, quali l'interazione fra stenotipia e computer nei sistemi Teletype, Velotype o Michela. L'inserimento dei testi generati nel sistema televisivo e la loro trasmissione via Teletext richiedeva nuovi sviluppi, via via realizzati. La possibilità di fruire di testi generati da sistemi di riconoscimento vocale appariva interessante, ma doveva superare numerosi ostacoli.

Ai timori di molti che ritenevano complesso l'approccio, l'autore del presente articolo confermava di aver sviluppato un semplice sistema domestico, dal costo di due milioni di lire, mostrando il video della partita dei mondiali di calcio che stava andando in onda, sottotitolato tramite *respeaking* sul PC di casa da parte di suo figlio (udente) per l'altro figlio (sordo) e soprattutto una scena veramente molto toccante che evidenziava l'utilità del *respeaking*. All'ultimo minuto della partita, quando l'Italia segna ed il gol viene annullato, si vede la scena in cui il figlio che sottotitola, sentendo la trasmissione, è dispiaciuto e arrabbiato del fatto che il gol è stato annullato. L'altro, il ragazzo sordo, ha questa reazione solo qualche secondo dopo, quando il sottotitolo compare allo schermo: cioè manifesta rabbia e delusione alzando le mani oltre la testa solo quando capisce dai sottotitoli che il gol è stato annullato.

Questi video hanno contribuito al successo di una convincente opera di sensibilizzazione e possibile armonizzazione, svolta in collaborazione con numerose reti televisive europee, l'organismo di normalizzazione CENELEC e l'Unione Europea delle Televisioni (EBU: European Broadcasting Union). Come scrive Massimiliano Rubbi nella rivista *I TeleAspettatori* del Centro

[Torna all'indice](#)

Documentazione Handicap (CDH): “Va dato atto in particolare a Giuliano Pirelli, coordinatore di Voice e ricercatore presso il JRC di Ispra, di insistere da anni, presso direzioni televisive spesso sfuggenti, per un più serio coordinamento e una maggiore attenzione alle esigenze dei sordi”.

I tempi erano maturi, lo sviluppo dei software lo poteva consentire, delle interfacce specifiche sono state realizzate e via via diverse reti televisive hanno utilizzato tali approcci per la generazione dei sottotitoli on-line, trasmettendoli in diretta non solo dalla sede della rete televisiva stessa, ma anche da altre postazioni dislocate a distanza, in particolare dal domicilio degli operatori, consentendo maggiore flessibilità operativa.

12. La produzione dei sottotitoli a distanza

La produzione dei sottotitoli a distanza è un elemento particolarmente importante, non solo per ridurre i costi di viaggio dei sottotitolatori, ma addirittura per rendere possibile la sottotitolazione stessa nei casi di interventi brevi, o imprevedibili ma urgenti, quando non si potrebbe coprire i costi o nemmeno organizzare l'intervento: brevi dimostrazioni a responsabili di università o pubbliche amministrazioni, catastrofi naturali, emergenze (Torri Gemelle), dibattiti politici (elezioni del Presidente della Repubblica).

Una sperimentazione di particolare rilievo è stata recentemente effettuata nel corso delle sessioni Euroscola presso il Parlamento Europeo (Strasburgo, febbraio 2013). Il Parlamento Europeo organizza regolarmente delle visite per gli studenti delle scuole, mettendo a disposizione l'Emiciclo perché quasi mille studenti, provenienti ogni volta da diverse nazioni, possano presentare le loro attività e portare avanti la discussione tramite lo stesso sistema di microfoni, d'interpretazione e di votazione utilizzato dai Parlamentari Europei. Giacomo Pirelli, dipendente dell'Università di Torino, dove sta lavorando per i progetti multimediali per il sito web e per i servizi amministrativi per gli studenti dell'università, e responsabile dei contatti di onA.I.R. con le Associazioni, durante uno stage a Strasburgo per sviluppare il sito Web dell'Euroscola (<http://www.europarl.europa.eu/euroscola>) e di caricare alcune informazioni relative agli eventi su una pagina dedicata su Facebook (<https://www.facebook.com/euroscola>) aveva necessità di seguire le sessioni di Euroscola nonostante la sua sordità. Ha così ricevuto l'aiuto da parte di un'assistente alla comunicazione, che prendeva gli appunti per lui e trascriveva alcune frasi, proponendone l'inserimento come post su Facebook. onA.I.R. si è offerta di effettuare la sottotitolazione di due giornate Euroscola in diretta a distanza.

[Torna all'indice](#)

Il Parlamento Europeo aveva già predisposto il *Web Streaming* per le scuole (link a fondo articolo) e, su questa base, i respeaker di onA.I.R. hanno potuto generare i sottotitoli e trasmetterli sul sito di onA.I.R., consentendo allo stagista ed altri partecipanti di seguirli sullo schermo dei loro PC nella sala di Strasburgo (Figura 6).



Figura 6: sottotitolazione in tempo a reale a distanza tramite respeaking dell'Euroscola

Link video streaming Parlamento Europeo per le scuole:

<http://www.europarl.europa.eu/ep-live/it/other-events/video?event=20130214-1600-SPECIAL-EUROSCOLA>

5) LA TRASCRIZIONE TRAMITE RESPEAKING IN INTERSTENO

Fabrizio G. Verruso

1. Introduzione

Il presente contributo, sia pur brevemente e tralasciando volutamente alcune tecnicità di difficile lettura e comunque non indispensabili ai fini di una sua valutazione complessiva, intende illustrare l'aspetto "agonistico" relativo all'introduzione della tecnica del riconoscimento del parlato (SR – *Speech Recognition*) alle competizioni mondiali organizzate dall'Intersteno, la Federazione Internazionale di Trattamento dell'Informazione, Associazione Mondiale fondata a Londra nel 1887.

2. Le gare di ripresa del parlato da Amsterdam 1995 a Hannover 2001

Nell'ambito delle competizioni mondiali organizzate dall'Intersteno, avremo modo di soffermarci su quelle relative all'abilità di ripresa del parlato (SC – *Speech Capturing*) mediante le diverse tecniche di scrittura veloce quali la stenografia, la stenotipia, la dattilografia (anche nella sua "ultima versione" di scrittura alla tastiera di PC), il riconoscimento del parlato. Tradizionalmente, le gare di SC avevano visto la partecipazione di resocontisti stenografi, professionisti e praticanti indipendenti, per lo più provenienti da realtà parlamentari, i quali mediante l'uso della nobile celere scrittura – che caratterizzò l'*ars notaria* sin da tempi remotissimi – avevano finito per "contraddistinguere" lo stesso nomen di dette competizioni, definite per l'appunto "di stenografia", quando invero, scorrendo le relative classifiche, ci si rendeva conto della presenza di concorrenti – nel corso degli anni quantitativamente sempre più numerosi – che presenziavano altresì con altre tecniche di scrittura veloce, come per esempio la stenotipia.

In questo contributo ci si limiterà a un'analisi delle classifiche disponibili on line (www.intersteno.org), vale a dire al periodo che va dal Congresso di Amsterdam (luglio 1995, che incidentalmente coincide peraltro con la prima presenza del sottoscritto alle gare internazionali) all'edizione di Parigi (luglio 2011). Ebbene, un primo dato numerico sulle gare di Amsterdam potrà renderci conto di tale partecipazione: 92 concorrenti con stenografia manuale (25 professionisti, 56 praticanti e 11 *juniores*¹ di cui un "professionista" della velocità di ripresa) a fronte di 32 partecipanti con stenotipia (17 professionisti e 15 praticanti)².

¹ Con età inferiore a 20 anni compiuti nell'anno di gara.

² Va qui detto che la distinzione tra professionisti e praticanti indipendenti non rispecchia, per la verità, quella che comunemente è attribuibile in ambito professionale, piuttosto le corrispondenti categorie di velocità, con soglie di velocità differenziate e chiaramente più elevate nella categoria A, per l'appunto definita "professional". Diversi

[Torna all'indice](#)

Le competizioni di Losanna (luglio 1998) vedono in classifica 78 stenografi (22 professionisti, 47 praticanti e 8 *juniores* di cui uno con velocità professionale di categoria A) e 27 stenotipisti (12 professionisti, 13 praticanti e 2 *juniores*). Infine, rammentiamo le competizioni di Hannover 2001, le ultime del periodo che precede il “Rinascimento”³ dell’Intersteno, che non vedono ancora la possibilità di gareggiare mediante la tecnica dello SR: 82 stenografi (26 professionisti, 53 praticanti e 3 *juniores*) e 26 stenotipisti (10 professionisti, 14 praticanti e 2 *juniores*).

3. Le gare di ripresa del parlato da Roma 2003 a Parigi 2012

È nel 2003, quindi, che il Congresso di Roma sancisce l’introduzione della tecnica di SR nel novero di quelle ammesse per la gara di SC. La scelta – frutto di consolidati affinamenti (per quanto *primigeni*) nel settore – che ammette detta modalità (che ricordiamo era precedentemente sostanzialmente relegata a una dettatura in “discreto” con la necessità di piccoli intervalli tra una parola e la successiva) è tutta italiana: il nostro Gruppo nazionale ne caldeggia l’approvazione in sede di Board dell’Intersteno. Una precisazione va qui resa nota prima di passare alla rassegna dei risultati: le gare di Roma vedono per la prima volta anche un’ulteriore modifica regolamentare.

Sono, infatti, soppresse le summenzionate categorie “A” e “B” e la gara di SC è costituita da un’unica dettatura – della durata complessiva di 15 minuti – intervallata da 3 distinte sezioni che registrano incrementi di velocità per ciascun minuto di dettatura. Questi i risultati: 45 stenografi (di cui 2 *juniores*), 49 stenotipisti (7 dei quali *juniores*) e 3 *respeaker* (uno dei quali, lo scrivente, di “matrice stenografica”), che alle sue latitudini professionali, erano spregiativamente definiti e derisi come *pappagalli*, nonostante le precisazioni sull’uso invero professionale di tale modalità di ripresa del parlato, oggi finalmente strumento di lavoro in cui si riconosce una comunità di colleghi, senz’altro ampiamente “metabolizzata” dalle altre sensibilità di addetti ai lavori.

Va qui subito detto che chiaramente, malgrado i PC all’epoca dedicati avevano configurazioni hardware inferiori rispetto a quelle che verranno nelle edizioni successive dei Campionati, i risultati conseguiti con la tecnica di SR sono del tutto interessanti sin dai primordi: l’ultimo minuto di classifica, ottenuto con il ricorso alla tecnica di SR, registra una velocità pari a 380 sillabe (Verruso), corrispondente al quarto miglior risultato, in assoluto, di quanto validamente trascritto dal concorrente stenotipista e al secondo miglior risultato in classifica ottenuto in quell’edizione per la stenografia (414 sillabe).

concorrenti, ancorché stenografi di professione, preferivano saggiare minori velocità di scrittura, gareggiando nella categoria B, spuntando verosimilmente migliori posizioni in classifica.

³ Secondo un termine caro agli organizzatori italiani del Congresso di Roma 2003, da cui si vorrebbe per l’appunto un’era nuova per l’Intersteno.

[Torna all'indice](#)

Nei Campionati successivi di Vienna (luglio 2005), si classificano 50 stenografi (2 dei quali *juniores*), 34 stenotipisti (11 *juniores*) e nuovamente 3 *respeaker*. Miglior risultato in velocità è quello in SR, 393 sillabe/minuto, primato nuovamente tutto italiano (Verruso) che va rapportato alle 331 s/m in stenografia e alle 349 s/m per la stenotipia. Ai Campionati cechi (luglio 2007), si classificano 45 stenografi (uno dei quali nella categoria *juniores*), 46 stenotipisti (tra cui 16 *juniores* e uno in categoria *pupils*⁴) e un altro primato nuovamente italiano (Di Nepi), con unica presenza in classifica (367 sillabe all'ultimo minuto di ripresa, ancora quarto miglior risultato in assoluto rispetto alle velocità conseguite con altre modalità di scrittura). Passiamo all'edizione cinese (Pechino, 2009): i risultati risentono inevitabilmente, sul fronte numerico, della minor partecipazione di concorrenti del Vecchio Continente, a fronte di una massiccia partecipazione di colleghi del Sol Levante: 23 gli stenografi classificati (2 dei quali *juniores*), 113 gli stenotipisti (sono moltissimi, tra costoro, i colleghi orientali, 21 dei quali *juniores* e 5 *pupils*) e nuovamente due italiani (Covello e Scianna) con la tecnica di SR, raggiungendo la miglior velocità di 419 sillabe, sesto miglior risultato in assoluto.

L'ultima edizione disponibile, alla data del presente scritto, è quella di Parigi 2011, in cui nuovamente accorrono numerosi i colleghi d'Europa e, tra questi, moltissimi gli stenografi tedeschi: 50 i concorrenti con stenografia manuale (uno in categoria *juniores* e una in quella *pupils*), 39 con stenotipia (di cui 6 *juniores* e un *pupil*), ancora due italiani (Covello e La Greca) con SR, con la migliore prestazione di sempre, pari a 497 sillabe all'ultimo minuto (Covello). Si registrano, infine, anche le presenze dei colleghi "tastieristi" che hanno potuto "infrangere" il tabù che la gara di SC fosse appannaggio dei soli stenografi (nel tempo, come visto, arricchiti della presenza di stenotipisti e *respeaker*). Ben 18 sono i colleghi dattilografi a classificarsi nella gara di ripresa del parlato (di cui 5 *juniores* e 3 *pupils*).

4. Conclusioni

Non è qui il caso di soffermarsi sulle ragioni che hanno impresso tali sorprendenti risultati con riferimento a una tecnica che, pionieristicamente, ha registrato sino a oggi primati tutti italiani nel corso delle varie edizioni dei Campionati Internazionali di Intersteno. Brevemente queste sono da ascrivere certamente al progresso intrinseco dei programmi di riconoscimento del parlato – dunque alla migliore accuratezza resa possibile da dizionari sempre più poderosi – e al maggior numero di "ore parlate" al *software* (un po' come le "ore volate" che, aumentando, rendono migliori i comandanti aerei!), alle migliori strumentazioni hardware con processori più veloci e performanti e alla maggior consapevolezza delle strategie competitive.

⁴ Con età inferiore a 17 anni compiuti nell'anno di gara.

[Torna all'indice](#)

Tutto questo non va, tuttavia, disgiunto da una qualità, peraltro comune a quella di ogni resocontista, che piace riassumere nelle parole riportate negli Atti del Congresso Intersteno di Firenze (1987)⁵, a proposito delle capacità intellettuali del professionista: “l’erudizione dello stenografo, le sue conoscenze nel campo concreto a cui si riferisce il testo da stenografare, la sua cultura linguistica e la padronanza sulle regole grammaticali”. Pare proprio in tali qualità racchiudersi il Rinascimento di una professione, una *ars notaria* che non può essere riassunta, come sfortunatamente accade, nell’idea che la semplicità nell’uso di un software racchiude *tout court* la complessità d’un antico mestiere: fare la cronaca del parlato. Rendendo alla Storia la sua Memoria.

⁵ in Atti del 37° Congresso Intersteno di Firenze (1987), pag. 105, disponibile su <http://www.intersteno.it/materiale/ComitScientifico/CollectionDocumentsLanguageComparabilityIntersteno.pdf>, p. 42.

6) LA SOTTOTITOLAZIONE TRAMITE RESPEAKING NEL MONDO⁶

Pablo Romero-Fresco (traduzione a cura di Tiziana Trapani)

Il respeaking in ambito televisivo⁷

Pur condividendo un obiettivo comune (la produzione di sottotitoli in tempo reale per sordi) e mezzi comuni (l'uso di software di riconoscimento del parlato), le aziende di respeaking di tutto il mondo si occupano di respeaking in modi diversi. Nel Regno Unito, Red Bee Media (RBM) sottotitola in diretta, tra gli altri canali, per la BBC, Channel 4 e Virgin Media Television. La maggior parte dei respeaker lavora in azienda, anche se alcuni respeaker della RBM lavorano a distanza. Le ore di lavoro dipendono dalla trasmissione di eventi in diretta e sono quindi irregolari e più simili a quelle di un interprete che a quelle di un traduttore o di un sottotitolatore. I respeaker di solito non lavorano su turni di 8 ore dalle 9 alle 17, ma piuttosto su turni di 10-12 ore con 2-3 giorni di riposo. In genere, il respeaking in tempo reale concreto copre solo la metà di una giornata di lavoro e idealmente non più di 20-30 minuti alla volta, anche se si può arrivare fino a 40 minuti. I respeaker dedicano l'altra metà della giornata alla preparazione dei successivi programmi da respeakere e alla produzione di sottotitoli pre-registrati.

Di solito lavorano in cabine insonorizzate o in stanze in cui hanno un televisore e un computer con un programma di riconoscimento del parlato (ViaVoice, anche se Dragon NaturallySpeaking è ora in fase di introduzione) e un'applicazione di sottotitolazione chiamata K-Live. Indossano le cuffie per ascoltare il programma originale e un microfono USB per (ri)parlare al computer. I respeaker della RBM utilizzano anche una piccola tastiera per assegnare i colori ai diversi locutori, che è una caratteristica indispensabile della sottotitolazione per sordi.

Nelle Fiandre, i respeaker della Vlaamse Televisie Maatschappij (VTM) lavorano in tandem. Ogni giorno hanno un incontro con i giornalisti incaricati di preparare e consegnare le notizie della sera e poi iniziano l'addestramento del software e ricevono gli script che verranno utilizzati nel telegiornale. I loro sottotitoli sono quindi in semi-diretta, in altre parole, preparati in anticipo e trasmessi in tempo reale. Tuttavia, per alcune notizie non ci sono script disponibili o, se ci sono, lo sono solo pochi minuti prima della trasmissione. Questi elementi devono essere respeakere in diretta da uno dei respeaker se è presente un solo locutore o da entrambi se c'è un dialogo. A differenza del Regno Unito, dove gli errori di sottotitolazione vengono corretti dai respeaker stessi appena li vedono apparire sullo schermo, nelle Fiandre i respeaker correggono i propri errori prima che il sottotitolo sia mandato in onda. Questo garantisce una grande precisione, ma aggiunge un ulteriore ritardo.

⁶ Questo articolo fa parte del progetto "Accessibilità linguistica e sensoriale: tecnologie per il voice-over e l'audiodescrizione", portato avanti dal *Ministerio de Economía y Competitividad* spagnolo (FFI2012-31023).

⁷ Una più approfondita descrizione del paesaggio del *respeaking* professionale in tutto il mondo si può trovare in Romero-Fresco 2011.

[Torna all'indice](#)

In questo compromesso tra la riduzione del ritardo e il miglioramento della precisione, la maggior parte delle aziende di respeaking in Francia ha chiaramente optato per la seconda. TF1, France 2 e France 3 sottotitolano telegiornali, programmi d'attualità e previsioni del tempo con la cosiddetta technique du perroquet, che coinvolge tre e spesso quattro persone. Il perroquet (letteralmente pappagallo) fa il respeaking nel "senso tradizionale", un souffleur (suggeritore), seduto accanto del respeaker, controlla le parole del respeaker e ciò che viene visualizzato sullo schermo e, se sono necessarie correzioni, avverte il correcteur (correttore), seduto accanto al suggeritore, che effettua le correzioni e manda in onda i sottotitoli. Le nuove linee guida francesi in realtà includono un ulteriore professionista nel team, un altro souffleur che avverte/sussurra al respeaker quando ha bisogno di aiuto e può subentrare quando il respeaker è stanco. Questo approccio fa pendere la bilancia verso il 100% di accuratezza dei sottotitoli a scapito del ritardo, che, in alcuni casi, può essere fino a 15 secondi, cioè 5-10 secondi in più rispetto a quanto avviene nel Regno Unito.

Altri due approcci inusuali al respeaking possono essere trovati in Canada e in Italia. Al fine di ridurre gli errori di riconoscimento derivanti dalla dettatura dei segni di punteggiatura, i respeaker della TVA in Canada utilizzano un joystick con cui si possono introdurre virgole, punti e altri segni di punteggiatura. In Italia, i respeaker che lavorano per Colby hanno un touch screen di fronte a loro, con tutti i segni di punteggiatura che potrebbero avere bisogno di usare e un vocabolario specializzato composto da parole che possono essere mal riconosciute dal software di riconoscimento del parlato. Toccando il tasto giusto nello schermo, i respeaker garantiscono il 100% di accuratezza, anche se aggiungono un po' di ritardo.

Infine, vale la pena segnalare che, sebbene il respeaking viene utilizzato principalmente a livello intralinguistico e in TV, non è sempre così. I respeaker alla RBM del Galles producono sottotitoli in diretta tramite respeaking dal gallese all'inglese e in società delle Fiandre come VTM viene fatto il respeaking dall'inglese all'olandese di eventi speciali in diretta come il discorso inaugurale di Obama. Per quanto riguarda il contesto, il *respeaking* è ora utilizzato anche in musei e luoghi d'arte, aule, conferenze, chiese, trasmissioni in diretta web e telefoni.

La selezione e la formazione dei respeaker

La maggior parte delle aziende si aspetta che i candidati abbiano una laurea, preferibilmente in lingue, ma non viene richiesto un corso di specializzazione in traduzione audiovisiva.

L'accento è posto sulle competenze richieste. I candidati sono tenuti ad avere ottime conoscenze linguistiche (con particolare riguardo all'ortografia), nonché competenze di comunicazione, correzione e capacità di lavorare in team.

Anche conoscenze di cultura generale e attualità, interesse nell'ambito televisivo e in questioni che riguardano i non udenti sono requisiti comuni, così come la capacità di multitasking, di essere a proprio agio in un ambiente altamente tecnologico e di superare la frustrazione causata dal verificarsi di

[Torna all'indice](#)

errori. Anche motivazione, capacità di recupero e flessibilità sono citate come competenze importanti.

Aziende come Mark Hall Associates, negli Stati Uniti, danno un'elevata importanza alla capacità di memoria e sono particolarmente ben disposte nei confronti di candidati che abbiano un background in ambiti come musica, teatro e interpretazione in lingua dei segni. In Svizzera, SWISS TXT accetta in particolar modo gli interpreti di simultanea, che sono particolarmente bravi nel *respeaking*, non da ultimo per le loro conoscenze linguistiche e la loro capacità di multitasking. In ogni caso, ciò che è interessante è che non c'è una sola di queste aziende di *respeaking* che si aspetta dai candidati delle conoscenze nell'ambito del *respeaking* e molte di loro sono sorprese di sentire che la formazione in questo campo esiste al di fuori dell'industria. In altre parole, la formazione e la ricerca all'università sul *respeaking* potrebbero andare di pari passo con quanto avviene nel mondo del lavoro, ma la maggior parte delle aziende di *respeaking* non ne sono ancora a conoscenza e la formazione aziendale comincia dall'inizio.

Prima di ottenere il posto di lavoro, per i candidati *respeaker* potrebbe essere richiesto un colloquio telefonico, un test di conoscenza sullo sport e le notizie in genere o, come nel caso della RBM, un breve esercizio su come intenderebbero sottotitolare un programma TV per sordi. Potrebbe essere richiesto loro di avere conoscenze di sottotitolazione, ma non di *respeaking*. Normalmente i candidati *respeaker* ricevono una spiegazione di ciò che è il *respeaking* e di come funziona (ascoltare e ripetere la colonna sonora originale, compresi i segni di punteggiatura, seguire il testo di partenza il più possibile, ma modificarlo per tenere il passo con il ritmo, ecc). Il test di *respeaking* può durare da 1 a 10 minuti e spesso diventa progressivamente più lungo e complesso. Se passano il colloquio, i *respeaker* neo assunti di solito sono sottoposti a un periodo di formazione che può durare da 3 settimane a 3 mesi.

Nella prima fase, entrano in confidenza con il software di riconoscimento del parlato e fanno pratica di *respeaking* con generi "facili" come lo sport; poi passano a telegiornali, dibattiti e interviste, considerati pezzi più impegnativi.

Il problema è che, data la pressione a sono sottoposte oggi le imprese, spesso non hanno altra scelta se non quella di tagliare i periodi di formazione dei candidati *respeaker*. Questo si traduce in *respeaker* mandati in onda in diretta prima di essere pienamente preparati e, soprattutto, in telespettatori non udenti che ricevono sottotitoli in tempo reale di qualità discutibile. Inutile dire che questo non accadrebbe se le imprese di *respeaking* potessero assumere candidati con una formazione precedente in *respeaking*, il che significherebbe che il periodo di formazione in un'azienda potrebbe essere speso aiutando i candidati a entrare in confidenza con le convenzioni aziendali su come fare *respeaking*.

[Torna all'indice](#)

Come detto in precedenza, a oggi sembra che la formazione in *respeaking* all'università si sia finalmente messa alla pari con l'industria, come dimostra il fatto che alcune aziende stanno impiegando studenti, alcuni dei quali sono già formatori in *respeaking*. Il compito dei ricercatori e dei formatori è ora di guardare avanti e investigare nuove strade che potrebbero non necessariamente essere state completamente esplorate dall'industria, come il *respeaking* interlinguistico, la sottotitolazione in tempo reale *speaker-independent* (in cui i *respeaker* diventano correttori dei sottotitoli creati automaticamente dal software di riconoscimento del parlato), il *respeaking* in 3D o la trascrizione basata sul riconoscimento del parlato automatico combinata con la traduzione automatica.

[Torna all'indice](#)

PARTE II: PROFESSIONI E FORMAZIONE

7. [Aspetti didattici \(Claudio Russello\)](#)

8. [Aspetti professionali nel respeaking in preregistrato](#)

[8a. La resocontazione dalla stenografia manuale al respeaking \(Luigi Zambelli\)](#)

[8b. Proposta del forumTAL sul verbale multimediale di atti giudiziari audio registrati \(Andrea Paoloni e Francesco Cutugno\)](#)

9. [Aspetti professionali nel respeaking in tempo reale](#)

[9a. – Il respeaking televisivo \(Arianna Mori\)](#)

[9b. – Il live editing televisivo \(Silvia Velardi\)](#)

7) ASPETTI DIDATTICI

Claudio Russello

1. La fase preparatoria: l'addestramento del software.

La figura del respeaker ha molte caratteristiche in comune a quella dell'interprete di simultanea e pertanto alcuni degli esercizi mirati allo sviluppo delle competenze psico-cognitive che sono richieste a un interprete di simultanea possono essere adottati anche per la formazione dei respeaker. Tuttavia, la complessità dell'insieme di operazioni da svolgere durante la sottotitolazione in diretta richiede alla figura professionale di un respeaker ulteriori competenze. Al fine di individuare l'insieme delle conoscenze che un respeaker deve possedere, divideremo l'attività di rispeaking in due fasi, la fase preparatoria e la fase del lavoro in diretta. Per quanto riguarda la fase preparatoria il respeaker condivide con il simultaneista le competenze relative alla creazione di glossari e alla ricerca della terminologia specifica del settore trattato, nonché la conoscenza del codice deontologico che sarà tenuto a rispettare durante le fasi lavorative.

Relativamente alle competenze proprie del respeaking per la fase preparatoria, sono di primaria importanza le conoscenze specifiche relative al software di riconoscimento del parlato, sia per ciò che riguarda gli aspetti tecnici (la tecnologia alla base del funzionamento, potenzialità e possibili limitazioni, caratteristiche del proprio profilo utente, ecc.), sia per ciò che riguarda le possibilità di addestrare il software stesso per migliorarne i livelli di riconoscimento. Parlando di addestramento del software ci si riferisce ovviamente ai software speaker-dependent, cioè a quei sistemi di riconoscimento del parlato che devono essere addestrati per riconoscere un determinato utente. La prima fase è la creazione di un profilo personale che avviene facendo leggere all'utente un certo numero di frasi, in tal modo il software è in grado di identificare l'impronta vocale dei singoli fonemi e di creare un'immagine degli stessi su un archivio. L'accuratezza del riconoscimento aumenta in modo proporzionale alle sedute di addestramento eseguite poiché il modello vocale acquisterà ogni volta maggiore precisione.

La fase di addestramento del software è di fondamentale importanza per il buon esito dell'attività del respeaker, in quanto il riconoscimento degli enunciati dettati al software può essere perfezionato grazie a un'adeguata, costante e ampia dettatura di nuovi testi. Arumí-Ribas e Romero-Fresco (2008) così evidenziano l'importanza dell'addestramento del software:

the amount of work carried out by this software in the respeaking process is directly proportional to the amount of work that must be carried out by the respeaker to constantly train it and improve it beforehand.

[Torna all'indice](#)

L'obiettivo di questa fase preparatoria è di far sì che il software ASR si discosti il minimo possibile da ciò che il respeaker andrà a dettare durante il lavoro in diretta e un'indicazione largamente condivisa dalle imprese del settore e dal mondo accademico suggerisce il raggiungimento da parte dei futuri respeaker di un livello di accuratezza del riconoscimento pari almeno al 97% prima di poter avanzare verso la sottotitolazione in diretta vera e propria. Per la valutazione complessiva del testo prodotto al termine della dettatura si rimanda al modello NERD in Romero Fresco (2011: 152).

Contestualmente al miglioramento del modello vocale tramite dettatura, un primo periodo di formazione deve essere dedicato all'arricchimento del dizionario del profilo utente tramite l'inserimento di un numero il più possibile ampio di testi selezionati e pertinenti al genere di argomenti che si dovranno respeakere. Tale operazione può essere fatta anche individualmente dagli studenti inserendo interi testi tramite un'apposita funzione del software o dettando al microfono quanto leggono e non è indispensabile in questa fase rispettare dei limiti temporali aumentando la velocità di dettatura o simulare delle fasi di lavoro in cui la ricezione del TP è simultanea alla produzione del TA, mentre è indispensabile rileggere e correggere il testo prodotto dal software.

Una volta che gli studenti abbiano creato e addestrato il proprio profilo utente, si può procedere a ulteriori esercizi specifici selezionati del docente. Come insegnamento per la fase preparatoria al respeaking, il docente può selezionare degli articoli che appartengono a diverse tipologie testuali e chiedere agli studenti di individuare all'interno di questi testi i termini che potrebbero non essere presenti nel proprio profilo utente. Una volta individuate le parole "sconosciute" al software, lo studente procederà alla creazione di una o più liste terminologiche e tematiche e le detterà al software decidendo se inserirle nel proprio profilo utente generale o in determinati vocabolari specialistici. Durante la dettatura si raccomanderà agli studenti di far attenzione alla pronuncia di particolari termini poco comuni, di nomi propri o di vocaboli stranieri, in quanto il software sarà in grado di riconoscerli durante la fase di lavoro solo se pronunciati esattamente nello stesso modo in cui sono stati dettati in fase di addestramento.

In una seconda batteria di esercizi si potrà dare agli studenti un'indicazione generica sull'argomento che si dovrà sottotitolare (per esempio un titolo giornalistico) sulla base della quale si dovranno effettuare le ricerche terminologiche più appropriate ed allenare il software con l'inserimento di nuove liste, con la creazione di macro e house-style, di vocabolari specialistici, ecc.

Una volta inseriti i termini nel vocabolario e addestrato il software, i candidati respeaker potranno dettare il testo presentato dal docente e in seguito si procederà all'individuazione dei termini non riconosciuti e si valuterà se gli errori di riconoscimento potevano essere evitati attraverso un ulteriore/migliore

[Torna all'indice](#)

addestramento. Una volta esaminati gli errori, è consigliabile procedere alla correzione degli stessi fornendo al software una nuova forma vocale per le parole non riconosciute, in tal modo si ridurranno le probabilità che il software commetta una seconda volta uno stesso errore di riconoscimento. L'esercizio andrebbe ripetuto anche a distanza di tempo per monitorare i probabili miglioramenti e risolvere eventuali problemi persistenti.

Si può a questo punto far risaltare l'importanza di questo primo spunto didattico caratteristico del *respeaking*. Questi insegnamenti specifici dovranno trovar posto all'interno di un curriculum apposito per la formazione dei *respeaker*, in quanto non contemplati nei percorsi accademici per interpreti di conferenza o per sottotitolatori.

Dedicare una prima parte del percorso formativo degli studenti a familiarizzare con il programma di riconoscimento è di fondamentale importanza in quanto il *respeaker* deve immediatamente prender confidenza con le potenzialità e le limitazioni del software che potremmo definire, parafrasando Ribas e Fresco, un suo compagno di lavoro costante che, se non vengono fatte correzioni, avrà l'ultima parola su quello che apparirà sullo schermo.

2. Il lavoro in diretta: le competenze psico-cognitive

Una volta completati gli esercizi preparatori e di addestramento del software, o ancor meglio in contemporanea a questa fase, il candidato *respeaker* passerà alle esercitazioni che mirano al perfezionamento delle capacità psico-cognitive che derivano in larga parte dalla didattica dell'interpretazione simultanea.

Molti esercizi propedeutici adoperati in maniera diffusa nei corsi accademici di Interpretariato di conferenza si adattano perfettamente anche alla formazione al *respeaking* e si tratta soprattutto di quegli insegnamenti che, più che concentrarsi sull'aspetto traduttivo (un *respeaker* solitamente lavora nella sua lingua madre), mirano a dotare gli studenti di quelle abilità cognitive che rivestono una parte fondamentale delle tecniche interpretative.

Per individuare tali insegnamenti può risultare opportuno rifarsi al modulo accademico di Sylvie Lambert (1989), docente presso l'università di Ottawa e ideatrice di un corso di introduzione all'interpretazione simultanea basato su 12 esercizi propedeutici d'apprendimento volti a sviluppare le competenze interpretative e la cui validità per un corso di simultanea è pienamente riconosciuta dalla prassi e da diffuse opinioni accademiche. Noi analizzeremo alcuni di tali esercizi soffermandoci su quelli che riteniamo più validi per la formazione di futuri *respeaker*.

Gli esercizi proposti da Lambert sono i seguenti: ascolto e memorizzazione, *shadowing*, esercizi di *dual-tasking*, parafrasi, sintesi, riempimento dal contesto, traduzione e interpretazione a vista, gestione di cifre,

[Torna all'indice](#)

nomi propri e acronimi, controllo del decalage, anticipazione e lateralizzazione cerebrale.

Partendo dal secondo esercizio della lista, uno studente di simultanea che abbia già acquisito familiarità con lo shadowing sarà molto avvantaggiato nella tecnica del respeaking che, infatti, viene a volte confuso da studenti alle prime armi o da estranei al settore come una pura e semplice ripetizione parola per parola e nella stessa lingua di un discorso. Pur non essendo così, una buona predisposizione verso lo shadowing aiuterà ogni futuro respeaker. Bisogna infatti notare che a differenza dello shadowing intralinguistico, per fare il respeaking bisogna prestare la massima attenzione a numerosi accorgimenti aggiuntivi, come, ad esempio: enunciare ogni singola parola in modo chiaro e scandito, aggiungere la punteggiatura, limitare e mantenere costante la velocità di eloquio, controllare e prevedere cosa verrà riconosciuto dal software e cosa bisognerà parafrasare o aggiungere manualmente dalla tastiera, ecc.

Per la formazione dei respeaker, i contenuti audiovisivi utilizzati per lo shadowing si potrebbero arricchire di ulteriori criticità. Si potrebbe, ad esempio, chiedere agli studenti di effettuare lo shadowing di discorsi pronunciati con marcati accenti, cadenze o rumori di fondo, con una velocità di eloquio in graduale aumento o di testi in cui si verifichi la sovrapposizione momentanea di più voci (per ricreare una caratteristica abbastanza comune alle reali condizioni di sottotitolaggio televisivo).

Passando alla tecnica del multi-tasking, la capacità di compiere più operazioni simultaneamente è, ovviamente, di fondamentale importanza sia per l'interpretariato che per il respeaking. La differenza tra le due discipline in questo caso consiste nel fatto che in un corso per interpreti si esercita la capacità di ascoltare in una lingua e riformulare in un'altra, mentre in un corso per respeaker si porrà l'accento sulla capacità di ascoltare e riformulare nella stessa lingua, ma con l'aggravio di aggiungere la punteggiatura e gestire i comandi per interagire con il software di riconoscimento del parlato (alternare colori, ricorrere a macro di dettatura, inserire il testo manualmente, ecc). Imparare ad automatizzare queste operazioni permetterà di evitare un sovraccarico cognitivo e un conseguente calo della concentrazione.

Un'altra abilità di sicura rilevanza per un respeaker è la capacità di sintesi, riduzione e parafrasi. Le limitazioni spazio-temporali del sottotitolo impongono spesso un adattamento del TP che può avvenire tramite riduzioni, omissioni o semplificazioni. Gli esercizi in questa fase mireranno a rendere automatici quei processi mentali che fin dall'ascolto di un enunciato suggeriscono una rielaborazione più semplice e snella: le espressioni ridondanti devono essere evitate (Buonasera e benvenuti ancora una volta...> Benvenuti...), si possono omettere le forme colloquiali (Mi lasci dire che...) e ove possibile è consigliabile semplificare le espressioni idiomatiche ricorrendo a soluzioni dal significato

[Torna all'indice](#)

equivalente (Per quanto riguarda il tema dello situazione...> Sullo situazione...). Anche a livello lessicale è spesso opportuno operare una semplificazione delle frasi subordinate e relative (La casa che era stata acquistata > La casa acquistata). Per rafforzare le abilità di parafrasi si può far esercitare lo studente a riformulare un messaggio originale, senza utilizzare gli stessi termini (parafrasi lessicale) e modificando l'ordine sintattico delle frasi (parafrasi sintattica).

Un altro esercizio che viene comunemente adottato per la formazione alla simultanea e che può risultare di sicura utilità per i futuri respeaker è il cosiddetto cloze test (Lambert, 1992b). Consiste nell'eliminare alcune parole da un testo e chiedere allo studente di riempire gli spazi vuoti senza pause o esitazioni. La capacità di creare una continuità di significato riempiendo gli spazi vuoti di un testo scritto, o i brevi silenzi di una traccia audio, sulla base delle informazioni fornite dal contesto, permette al respeaker di affinare quelle strategie che risulteranno particolarmente utili in delle situazioni lavorative, non certo poco frequenti, caratterizzate da una scarsa qualità dell'audio, da un eloquio non accurato dell'oratore, dalla presenza di frasi troncate o false partenze, da problemi tecnici o locali non insonorizzati, ecc.

Queste tecniche di riformulazione, semplificazione e riduzione del testo garantiscono una segmentazione più efficace, dei sottotitoli più leggibili e la possibilità di limitare il ritardo sul TP.

Parlando di riformulazione del TP, può essere interessante far notare che questo è uno dei temi più discussi tra gli addetti ai lavori della sottotitolazione. Alcune associazioni in difesa dei diritti dei sordi avanzano la richiesta di poter accedere ai contenuti audiovisivi senza alcuna distinzione rispetto agli altri fruitori. Da qui il loro appello: "Don't patronise us, please" (Eugeni 2009), per richiedere alle emittenti televisive di non semplificare o modificare i dialoghi originali del contenuto audiovisivo. Tuttavia, bisogna far notare che si tratta di una richiesta che quando è rivolta agli operatori che si occupano di sottotitolazione in diretta (come i respeaker) è a volte impossibile da esaudire. La ripetizione di ogni singola parola del messaggio di partenza li obbligherebbe a compiere esclusivamente e in ogni circostanza il respeaking verbatim e non terrebbe conto del fatto che dettando la punteggiatura aumenterebbe il numero delle parole pronunciate dal respeaker e di conseguenza il ritardo di comparsa del testo.

Anche il tema del ritardo di apparizione dei sottotitoli è un aspetto critico. Nel campo della sottotitolazione, le linee guida più comunemente adottate (OFCOM, ESIST, ecc.) impongono che la comparsa dei sottotitoli sia sincronizzata all'immagine. Questo aiuta i non udenti alla comprensione della situazione che circonda l'evento comunicativo attraverso le immagini e in alcuni casi permette loro di integrare la lettura del sottotitolo con la lettura labiale. Inoltre, nel caso della sottotitolazione di programmi televisivi caratterizzati dall'alternarsi delle

[Torna all'indice](#)

immagini a un'elevata velocità (tg, talk-show, ...) il ritardo di comparsa non potrà essere superiore a pochi secondi, poiché la totalità del messaggio è data dalla sua caratteristica multimodale e un ritardo del sottotitolo anche di soli 5 secondi potrebbe rendere complicata l'attribuzione di una frase a un determinato oratore o a una particolare scena visiva.

Per la sottotitolazione in diretta un *décalage* minimo è tuttavia inevitabile. Indipendentemente dalla professionalità o dall'abilità del respeaker di ridurre al minimo l'intervallo tra la ricezione del messaggio sonoro e la riformulazione tramite dettatura, esistono delle limitazioni tecniche dei software utilizzati che richiedono dei tempi, seppur minimi, per processare l'input audio, produrre e trasmettere il testo scritto. Nella formazione al respeaking è pertanto consigliato dedicare particolare attenzione all'esercizio che abitui lo studente a ridurre al minimo il *décalage*.

Essendo la fase dello sviluppo delle competenze psico-cognitive la parte probabilmente più consistente nell'insegnamento della tecnica del respeaking e poiché in molti corsi per interpreti di simultanea tali insegnamenti fanno parte del curriculum di base, si può ribadire nuovamente che gli studenti di Interpretariato avranno dei proficui vantaggi al momento di doversi avvicinare alla tecnica del respeaking perché avranno già acquisito le conoscenze necessarie per svolgere più operazioni in multi-tasking, per ovviare al sovraccarico cognitivo, saranno in grado di ricorrere alle strategie di compensazione, riformulazione e riduzione, sapranno regolare il *décalage*, saranno abili nella gestione dello stress e sapranno adattarsi alle caratteristiche specifiche di ogni evento comunicativo.

3. Aspetti didattici propri del respeaking e competenze complessive

Nel delineare l'insieme delle competenze per il lavoro in diretta richieste ad un respeaker, oltre a quelle già esposte provenienti dal campo dell'interpretazione di simultanea, bisogna individuare ulteriori competenze che sono specifiche del respeaking vero e proprio o che in parte provengono dalla sottotitolazione. Tra queste, è opportuno indicare tre macrocategorie.

1 - L'abilità e la velocità nel cambiare il colore e la posizione del sottotitolo, di aggiungere delle etichette, la capacità di operare delle scelte di pre-editing immediate (ad es. nel caso in cui si capisce che un termine che si sta per dettare al software non potrà essere riconosciuto e quindi cambiare espressione o decidere di aggiungerlo manualmente), la capacità di gestire il feedback istantaneo offerto dal testo scritto che compare sullo schermo e di gestire possibili imprevisti tecnici o 'capricci' della tecnologia impiegata. Tali competenze potranno essere acquisite grazie a un costante utilizzo dei software di riconoscimento e di visualizzazione su schermo affiancato dalla correzione e valutazione dei testi prodotti tramite respeaking.

[Torna all'indice](#)

In questa fase del training si può introdurre il vincolo di controllare simultaneamente alla dettatura quanto viene prodotto sullo schermo per abituare gli studenti a dedicare una parte del carico cognitivo alla lettura del sottotitolo, al monitoraggio della sua forma grafica (struttura, allineamento, colori, ecc.) e a gestire le possibili sensazioni di sconforto nel vedere i propri errori sullo schermo. Allo scopo di sviluppare le capacità di modificare il TP in funzione delle limitazioni del software, il docente potrà apportare delle modifiche o aggiungere nuovi termini, possibilmente non presenti nel profilo utente dello studente, a un testo che gli studenti avevano preparato precedentemente, addestrando il software. In questo modo lo studente dovrà trovare delle soluzioni immediate alternative per aggirare questi potenziali errori di comprensione del programma di riconoscimento anche attraverso l'aggiunta dei nuovi termini tramite tastiera. Un altro esercizio potrebbe essere quello di richiedere agli studenti di evitare di utilizzare alcune parole specifiche per far sì che trovino delle soluzioni per mantenere intatto il senso generale del messaggio senza restare legati a determinate espressioni e mantenendo inalterata la lunghezza complessiva del TP.

2 - Un'altra serie di competenze che il respeaker non condivide con un simultaneista riguarda la particolare accuratezza nella produzione e articolazione degli enunciati da dettare al software di riconoscimento. In questa fase gli studenti dovranno acquisire le capacità di pronunciare frasi di lunghezza limitata e pressoché costante, di ricorrere a una pronuncia neutra e chiara, di prestare attenzione nel marcare chiaramente la fine e l'inizio di una parola, di non emettere suoni, allungamenti di vocali e pause piene tra una parola e l'altra, di non introdurre pause non necessarie tra le parole (potrebbero essere intese dal software come la fine di un enunciato) e la capacità di mantenere una velocità di dettatura elevata.

Questa seconda serie di competenze fonetiche richieste ad un respeaker sono evidentemente non condivise da un interprete di simultanea, o almeno solo in parte, e dovranno pertanto essere sviluppate attraverso insegnamenti specifici. Sembra conveniente ricorrere a una serie di esercizi mirati al miglioramento della gestione della voce proposti da Eugeni (2009) e raggruppati in quattro categorie: "esercizi di riscaldamento del corpo della voce, esercizi di respirazione, esercizi di modulazione della voce ed esercizi di articolazione." Si tratta di una parte molto importante per la formazione di un respeaker in quanto il risultato finale del suo lavoro sarà fortemente influenzato dall'accuratezza della dettatura e dall'automatismo delle proprie capacità fonetiche. Eugeni (2009) suggerisce che questa automaticità può essere raggiunta solo attraverso un esercizio costante e pertanto è consigliabile che gli insegnanti dedichino una parte iniziale di ogni lezione al perfezionamento delle competenze fonetiche e all'aggiornamento del profilo vocale del software.

[Torna all'indice](#)

3 - Infine, sono molto importanti anche le competenze che provengono dal campo della sottotitolazione interlinguistica e dalla sottotitolazione per non udenti (SDH) in maniera specifica. In questa categoria rientrano da un lato le conoscenze tecniche relative alle tecnologie usate per la sottotitolazione e dall'altro le competenze teoriche relative alla capacità di individuare le caratteristiche precipue di ogni genere di evento comunicativo da sottotitolare (differenti tipi di programmi TV e non solo), di saper come gestire le situazioni di accavallamento di voci o di veloci scambi tra più oratori, di presentare le componenti para ed extra-linguistiche del TP e di impaginazione e visualizzazione dei sottotitoli (Scroll, Roll o Block Mode). È fondamentale inoltre che lo studente conosca le indicazioni delle linee guida sulla leggibilità e, a tal proposito, saranno necessarie anche le conoscenze teoriche relative ai principi dell'accessibilità e della sordità, affinché il respeaker sappia con precisione quali sono le caratteristiche del pubblico a cui è destinato il sottotitolo e le sue necessità (velocità di lettura, comprensione linguistica,...).

Una volta effettuata la disamina delle competenze richieste alla figura del respeaker e i rispettivi insegnamenti necessari per lo sviluppo di tali competenze sembra possibile proporre un modulo accademico per la formazione di professionisti del respeaking che si inserisca all'interno di un completo corso accademico per la formazione di traduttori ed interpreti.

Un corso di respeaking potrebbe essere inserito sia in una prima fase di un corso accademico in Interpretariato di conferenza, sia in una più avanzata. Nel primo caso gli insegnamenti che rappresentano il nodo fondamentale della preparazione al respeaking, quelli mirati allo sviluppo delle capacità psico-cognitive, sarebbero impartiti in concomitanza con la normale linea didattica del corso di laurea in Interpretariato. Nel secondo caso, invece, le competenze psico-cognitive fornirebbero il supporto precedentemente acquisito al quale aggiungere gli insegnamenti aggiuntivi tipici del respeaking e della sottotitolazione individuati in precedenza.

4. Conclusioni

Si può a questo punto procedere a un riepilogo dell'insieme delle competenze che un corso per respeaker dovrebbe comprendere:

- Competenze fonetiche e linguistiche.
- Competenze tecniche e pratiche.
- Competenze teoriche e di contenuto.
- Competenze psico-cognitive.

1) Per l'analisi delle competenze fonetiche e per la delineazione delle esercitazioni utili al miglioramento delle stesse si è ricorso alle indicazioni di Eugeni (2009) e si ribadisce l'utilità dello sviluppo di queste abilità per tutta la

[Torna all'indice](#)

durata del corso. Per quanto riguarda le competenze linguistiche sembra opportuno sottolineare che un respeaker dovrà avere delle conoscenze della lingua materna estremamente elevate in tutti i suoi aspetti: morfologia, ortografia, grammatica, sintassi. I sottotitoli, infatti, oltre ad avere la funzione di permettere l'accesso all'informazione e alla cultura, rappresentano uno strumento per gli utenti sordi, o per un altro tipo di pubblico come stranieri e bambini, attraverso il quale apprendere e migliorare le proprie conoscenze della lingua italiana. Da qui deriva l'estrema importanza di presentare il testo privo di errori grammaticali o sintattici, seppur a volte dovuti al software. Inoltre il respeaker dovrà essere dotato di capacità creative e di sensibilità linguistica per riuscire a cogliere le sfumature prosodiche del linguaggio parlato e trovare la forma adeguata per trasmetterle, riconoscere il registro, il grado di formalità, l'accento di un oratore, ecc. e decidere se e come renderli.

2) Le competenze tecniche e pratiche riguardano il profondo conoscimento del software di riconoscimento del parlato con il quale si lavora e di altri strumenti tecnologici che possono rivelarsi essenziali per la fase di preparazione (ricerche terminologiche, tecniche di documentazione, ecc). Inoltre sembra consigliabile una generale predisposizione verso le evoluzioni della tecnologia per restare al passo e saper cogliere i progressi raggiunti dalla tecnologia nel momento in cui si verificano. È opportuno far presente, infatti, che sono ogni giorno più promettenti le prospettive legate alle future tecnologie che mirano all'evoluzione dei programmi di riconoscimento del parlato speech-independent. A tal proposito riportiamo un'arguta espressione di Arumì Ribas e Romero Fresco che sottolineano l'importanza del ruolo del software nel lavoro del respeaker affermando che: "in the same way that speech recognition software is often described as speaker-dependent [...] the respeaker can be said to be software-dependent" (2008).

3) Le competenze teoriche devono fornire al respeaker una perfetta conoscenza delle caratteristiche delle disabilità uditive, con particolare riferimento ai differenti gradi di sordità (perlinguale, peri-linguale e post-linguale; lieve, media, grave e profonda) e le relative competenze linguistiche di ogni gruppo di disabili (velocità di lettura, comprensione dell'italiano come prima o seconda lingua...). Solo su questa base il respeaker sarà in grado di effettuare delle scelte strategiche relative alla produzione del TA (trascrizione più o meno fedele, o addirittura verbatim, trasposizione, semplificazione, ecc.). Inoltre sembra opportuno possedere una visione generale sul tema dell'accessibilità e una conoscenza dettagliata delle linee guida per la produzione di sottotitoli. Per competenze di contenuto si fa riferimento invece alla conoscenza delle caratteristiche dei contenuti audio-visuali e alla componenti semiotiche che li caratterizzano.

4) Infine, per quanto riguarda le competenze psico-cognitive si può fare riferimento alla serie di attività ed esercizi proposti prendendo lo spunto dal

[Torna all'indice](#)

modulo per la formazione di interpreti di simultanea. Un'ultima considerazione potrebbe riguardare l'indicazione a rafforzare la tecnica dello shadowing (con tutte le varianti precedentemente esposte) dato che questa tecnica è alla base di molti esercizi simili, ma da adottare in differenti fasi della formazione del respeaker.

8) ASPETTI PROFESSIONALI NEL RESPEAKING IN PREREGISTRATO

8a. LA RESOCONTAZIONE DALLA STENOGRAFIA MANUALE AL RESPEAKING

Luigi Zambelli

1. Introduzione

L'office automation, che specialmente dagli anni '80 in poi ha interessato ogni settore produttivo presentando sul mercato prodotti fortemente innovativi e nel contempo user friendly, ha inciso positivamente anche nell'area del word processing e, in particolare, nei processi di ripresa del parlato e nella sua trasposizione nella forma scritta. Per quanto riguarda in particolare la situazione italiana – e qui mi limito a estendere la mia analisi alla resocontazione parlamentare – le tecniche di ripresa del parlato, consolidate nel corso di decenni, sono state essenzialmente due, ovvero la stenografia manuale e la stenografia meccanica, poi divenuta elettronica. Le Aule parlamentari, Camera e Senato, sono storicamente un ambito privilegiato in cui la documentazione scritta ha un'importanza strategica e la stesura dei resoconti/atti parlamentari è regolamentata da precise norme legislative e procedurali, consolidate nel corso degli anni. Da sottolineare che tali disposizioni hanno sancito l'uso della stenografia manuale alla Camera e di quella meccanica (sistema Michela) al Senato. Solo in epoca più recente – ovvero verso la metà degli anni '90 – l'informatizzazione ha interessato anche il Parlamento italiano. Grazie alla lungimiranza di alcuni funzionari e a un'incisiva attività di lobbying da parte di aziende del settore, sono iniziate le prime sperimentazioni che hanno portato all'introduzione della stenotipia elettronica e ai sistemi di riconoscimento del parlato (*speaker dependent*), prevedendo addirittura la possibilità di affidare ad aziende esterne la stesura dei resoconti.

Oggi, le sperimentazioni in corso da parte di primarie aziende nel settore, già in essere sia in ambito parlamentare che giudiziario, sono orientate all'introduzione dei software *speaker independent*, con l'obiettivo di giungere alla realizzazione del resoconto multimediale delle sedute assembleari. Ho voluto fare questo breve accenno alla situazione operativa in ambito parlamentare perché profondamente diversa, sia per le tecnologie in uso, sia per la frammentarietà e eterogeneità dei processi produttivi e la conseguente professionalità degli operatori, da quella che si riscontra nelle aziende che si

[Torna all'indice](#)

occupano di resocontazione non parlamentare e di verbalizzazione giudiziaria. Prima di addentrarci nello specifico del tema degli aspetti operativi dell'uso del respeaking quale tecnica di ripresa del parlato a livello professionale, devo ricordare che, fino a una ventina di anni fa, in Italia il mercato della verbalizzazione/resocontazione in real time era quasi inesistente. Nella maggior parte dei casi si procedeva alla registrazione analogica degli eventi affidando poi la trascrizione (o, meglio, "la sbobinatura") a personale impiegatizio e con una professionalità diversa da quella richiesta oggi a un resocontista. Per la stesura degli atti di eventi particolarmente significativi (congressi, seminari di alto livello, consigli di amministrazione di enti pubblici e di importanti aziende nazionali), gli organizzatori degli stessi si rivolgevano agli stenografi manuali, professionisti di altissimo livello, provenienti dalle aule parlamentari o dalle redazioni giornalistiche. Oggi esiste invece una nuova figura professionale, quella del resocontista/verbalizzatore, solitamente un free lance o un collaboratore di realtà aziendali ben strutturate. Questo professionista possiede degli skill che vanno al di là delle abilità che si esprimono nella capacità di ripresa del parlato anche alle velocità oratorie: deve possedere infatti una solida cultura generale, una forte motivazione, predisposizione a lavorare in team, la capacità di controllare situazioni di forte stress ecc.

2. L'avvento della Stenograph: dalla stenotipia meccanica a quella elettronica

Con l'entrata in vigore delle norme previste dal nuovo Codice di procedura penale (24.10.1989), che ha sancito la possibilità di utilizzare, accanto alla dattilografia, la stenotipia "o altro strumento" per la ripresa e la stesura dei verbali delle udienze penali, si sono aperti nuovi orizzonti e a quanti potevano fornire dei servizi di stenotipia si aprivano concrete opportunità lavorative. La macchina per stenografare Michela continuava ad essere utilizzata al Senato ma, attraverso sofisticati interventi di ingegnerizzazione, venne informatizzata: la macchina forniva il testo ripreso in aula in caratteri corsivi, di immediata lettura e pronto per i successivi interventi di correzione da parte dei funzionari dell'Ufficio Resoconti. Da qualche anno era tuttavia iniziata l'importazione in Italia della Stenograph una macchina meccanica per stenotipia da decenni utilizzata nelle aule giudiziarie americane. Nella versione meccanica e, soprattutto, dopo qualche anno nella sua versione elettronica, riscosse molto interesse. Grazie soprattutto alla lungimiranza del prof. Marcello Melani, autore del metodo di stenotipia italiano applicato a tale macchina, dopo anni di positiva sperimentazione didattica la stenotipia ottenne significativi risultati in occasione dei Campionati internazionali organizzati dall'Intersteno.

3. Da un mercato di nicchia a una presenza capillare sul territorio nazionale

A dimostrazione dell'interesse del mercato verso la verbalizzazione giudiziaria, credo debba essere ricordato un fatto molto significativo e che ha sicuramente contribuito a catalizzare l'attenzione di molti soggetti imprenditoriali nei confronti della trascrizione elettronica del parlato: nel 1998 –

[Torna all'indice](#)

proprio in concomitanza con l'entrata in vigore del nuovo Codice di procedura penale – tre importanti aziende diedero vita a un consorzio nazionale, “Endike”, con sede a Roma, che si propose al Ministero di Grazia e Giustizia per la *“gestione del servizio di fonoregistrazione e trascrizione delle udienze penali”* su tutto il territorio nazionale. La società Philips, che già aveva vinto l'appalto per la fornitura degli apparecchi di registrazione analogica, aveva installato gli impianti per la registrazione video delle udienze presso importanti sedi giudiziarie. Data Management, Olivetti e Stenotype Italia si sarebbero occupate del servizio di trascrizione degli atti processuali.

Le prime due aziende erano fortemente interessate a quanto stava accadendo nel mondo giudiziario perché all'epoca avevano iniziato la produzione e la commercializzazione di due prototipi di macchina per stenografare. In particolare, Data Management aveva presentato sul mercato lo “stenoterminale Michela”, ossia una versione informatizzata del modello del prof. Michela Zucco ancora in uso al Senato. La società Olivetti si presentava sul mercato con un modello di macchina per stenotipia prodotto dalla Mael, gruppo Olivetti. Si trattava della versione informatizzata di un prototipo di macchina per stenografare meccanica inventata dal prof. Gornati di Varese, che tuttavia non ha avuto grande successo. Infine Stenotype Italia, che poteva e può essere considerato il soggetto imprenditoriale più forte. In realtà le aziende aderenti al consorzio Endike non erano interessate alla vendita al Ministero dell'hardware, ma l'obiettivo era unicamente quello del “servizio di fonoregistrazione” (Philips) e verbalizzazione (Data Management, Olivetti/Mael, Stenotype Italia). Il progetto era alquanto ambizioso: si trattava di dare vita a una serie di piccole realtà imprenditoriali operanti sul territorio nazionale e aderenti al consorzio Endike, che avrebbero messo a disposizione le risorse umane (gli stenotipisti) in grado di garantire al Ministero il servizio di fonoregistrazione – e in alcuni Tribunali anche di videoregistrazione - nonché quello di trascrizione delle udienze penali in un consistente numero di sedi giudiziarie.

La nascita del Consorzio Endike si collocava tra le iniziative imprenditoriali coinvolte nelle vicende di “tangentopoli” e quindi, nonostante gli ingenti investimenti iniziali, a seguito di vicende non ebbe modo di operare. Il progetto di dar vita ad un organismo in grado di fornire il servizio di fonoregistrazione e verbalizzazione delle udienze penali a livello nazionale aveva sicuramente una validità al punto. A distanza di quasi vent'anni, infatti, è stato riproposto al Ministero dal Consorzio “Astrea”, ovviamente con un'opportuna rivisitazione non solo dei servizi offerti ma anche delle condizioni economiche. Una rete di ditte individuali, cooperative, piccole imprese consente oggi ad Astrea di essere presente in un consistente numero di sedi giudiziarie.

4. Il riconoscimento del parlato

Parallelamente alla diffusione della stenotipia si consolida una spinta propulsiva alla ricerca e alle sperimentazioni in atto sul versante dei software di

[Torna all'indice](#)

riconoscimento del parlato e del trattamento automatico del linguaggio. Ecco quindi che, accanto agli studi che da anni si stavano compiendo in particolare negli Stati Uniti, dobbiamo ricordare quelli svolti dall'ing. Pirelli presso il Centro Comune di Ricerca della Commissione Europea, con sede a Ispra (VA), e le ricerche che parallelamente si stavano compiendo in altri centri di riconosciuta eccellenza, tra i quali va sicuramente annoverato l'Istituto Trentino di Ricerca (ora Fondazione Bruno Kessler) di Trento, diretto dall'ing. Lazzari. Ai primi software di riconoscimento del parlato – *speaker dependent* – capaci di riconoscere un numero di parole alquanto limitato e comunque circoscritto a un determinato settore, si sono aggiunti programmi sempre più evoluti e in grado di dare risposte positive alle richieste di un mercato allo stato nascente. I primi software applicativi riguardavano in particolare il riconoscimento della terminologia usata per la refertazione radiologica mentre oggi noi possiamo disporre di software di riconoscimento di voce *speaker dependent* in grado di riconoscere il parlato e fornire la trasposizione scritta delle parole con un margine di errore ridottissimo.

5. La resocontazione presso lo Studio Acta – aspetti organizzativi

Sebbene l'attività di resocontazione – core business dello Studio Acta – inizi nel 1998, l'utilizzo del software di riconoscimento *speaker dependent* Dragon NaturallySpeaking è alquanto recente. Dopo un primo periodo di sperimentazione e valutazione delle diverse opzioni presenti sul mercato, nel 2008 si avvia l'attività sul campo. Parallelamente all'uso dei tradizionali metodi di ripresa del parlato (stenografia manuale e trascrizioni al PC), venne strutturato un percorso formativo teorico/pratico con un periodo di training on the job, per permettere ai collaboratori un passaggio al riconoscimento del parlato che fosse il meno traumatico e nel contempo il più professionale possibile.

La formazione partiva da un assunto a nostro avviso determinante, ovvero il fatto che nell'attività di resocontazione l'elemento più qualificante e al quale si deve porre costante attenzione non è lo strumento (la stenotipia, la scrittura al PC, i software di trattamento automatico del linguaggio) ma quel contorno variegato di competenze che, solo se ben coordinate e potenziate, portano a un positivo risultato finale, ovvero una professionalità degna di tale nome. L'attenzione alle competenze trasversali (capacità di analisi, sintesi, disponibilità a lavorare in team ecc.) deve essere potenziata giorno dopo giorno proprio perché gli ambiti in cui i resocontisti operano sono estremamente ampi e in dinamica espansione: può essere richiesta la resocontazione di un convegno a carattere giuridico/economico; di un meeting aziendale con un linguaggio settoriale, tecnico, molto sofisticato; di una *lectio magistralis* di un eminente uomo di cultura; la trascrizione di una conferenza con l'intervento di oratori stranieri; ecc. Tutta questa varietà di situazioni e linguaggi comporta, come vedremo in seguito, la necessità di arricchire il proprio vocabolario e di "istruire" adeguatamente il software di riconoscimento del parlato. Ebbene, sia nella fase

[Torna all'indice](#)

di preparazione del progetto, sia soprattutto durante l'iter formativo, la risposta dei collaboratori è stata molto positiva e addirittura propositiva in quanto sono state studiate congiuntamente nuove soluzioni e sono state valutate concrete proposte operative per salvaguardare – o meglio valorizzare – la loro pregressa esperienza nella produzione di atti in differita, acquisita in tanti anni di attività.

6. L'attività di promozione e la conquista di nuovi mercati

Parallelamente alla formazione sul campo dei resocontisti, lo Studio Acta ha dovuto affrontare – come ogni altra azienda *customer oriented* – le problematiche connesse alla “commercializzazione” del prodotto “resoconto”. Poiché, tuttavia, lo Studio Acta poteva contare su un “pacchetto clienti” acquisito grazie a una presenza quasi ventennale sul mercato, dal 2008 in poi si è trattato soprattutto di consolidare la posizione e di trovare nuove nicchie di mercato in cui potersi posizionare per garantire ai collaboratori una stabilità in termini occupazionali. Quest'azione di “esplorazione” di nuovi mercati non è stata facile. Nelle azioni di sensibilizzazione del mercato e di promozione dei nostri servizi si incontrano non poche difficoltà in quanto si tratta di far superare al potenziale cliente pregiudizi nei confronti dei servizi di verbalizzazione e di resocontazione ormai consolidati e che, purtroppo, sono in parte giustificati. Con l'avvento massiccio sul mercato della stenotipia (a partire dalla metà degli anni '90) e la presenza di molti operatori improvvisati che, finalizzando il loro lavoro unicamente al profitto a danno della qualità, hanno deturpato l'immagine del nostro lavoro, la resocontazione ha subito dei pesanti contraccolpi.

Attività promozionali fuorvianti, che ancora oggi fanno passare la semplice trascrizione *verbatim* per resoconto, hanno disorientato gli organizzatori congressuali (PCO, Professional Congress Organizer), i *meeting planner*, i responsabili della comunicazione /documentazione di enti pubblici e di organismi privati, che si trovano ora a dover scegliere tra un servizio professionale (con i relativi costi) e una “sbobinatura” a costi irrisori. Da tutto questo consegue che il posizionamento in una nicchia di mercato ben precisa e, soprattutto, l'attività di fidelizzazione dei clienti richiedono un quotidiano impegno di tutti i nostri collaboratori, da chi si occupa delle relazioni esterne e della vendita ai resocontisti, ai revisori, all'addetto al controllo qualità. La resocontazione non è, infatti, il prodotto di un singolo professionista, ma è il risultato dell'impegno di un'équipe motivata, capace di dominare e superare gli inevitabili momenti di stress e i problemi di natura tecnica che inevitabilmente si presentano nella quotidianità. Nella resocontazione – soprattutto in quella in real time – ognuno deve esprimere il meglio di sé e non solo in termini capacità di ascolto e di velocità di immediata ri-dettatura al computer. Sotto il profilo operativo il team “tipo” dello Studio Acta è composto da tre professionisti che si alternano nella ripresa del parlato con turni di 20/25 minuti.

Ognuno corregge il proprio lavoro e il file Word viene poi inviato al responsabile che coordina tutte le fasi della ripresa intervenendo negli eventuali

[Torna all'indice](#)

momenti di criticità. Al termine dell'evento può quindi essere consegnata al cliente la trascrizione in prima stesura e, entro due ore, il testo definitivo con duplice controllo ed eventuale correzione di termini stranieri, citazioni ecc. con ricerca/verifica su Internet. La durata della ripresa in respeaking di ogni singolo professionista è limitata a 20/25 minuti perché, soprattutto allorché un oratore parla velocemente, la ridettatura rischia di diventare imprecisa, i suoni emessi dal respeaker perdono cioè di nitidezza con il conseguente aumento degli errori da parte del software di riconoscimento del parlato. Tra l'altro l'esperienza ci ha dimostrato che, di norma, l'intervento di un relatore in sede congressuale non supera questa durata.

Questo facilita il lavoro perché il resocontista può seguire, anche dal punto di vista concettuale, un intero intervento con una concentrazione ai massimi livelli. Una disattenzione, anche in un passaggio brevissimo di un intervento, può portare a errori con pesanti riflessi sulla qualità di tutta la prestazione. Questa "parcellizzazione" del lavoro potrebbe sembrare un elemento che penalizza l'utilizzo del respeaking rispetto alla ripresa con la stenotipia. Una valutazione di questo tipo, a nostro parere, è non solo affrettata e fuorviante in quanto non tiene conto della concomitanza di una molteplicità di fattori che giocano un ruolo determinante nella valutazione complessiva del rapporto "costo-benefici". Infatti, l'esperienza ci ha dimostrato che una stenotipista, pur se in grado di seguire interventi anche ad alte velocità e, addirittura, più interventi in successione e per tempi oggettivamente più lunghi rispetto a quelli in cui si misura la ripresa di un nostro respeaker, deve essere comunque affiancata da uno o due colleghi che si dovranno occupare della revisione del testo. Questo perché anche la stenotipista più capace non sfugge allo stress derivante dalla ripresa alle alte velocità. Con la stanchezza inoltre aumenta il coefficiente di errore e, di conseguenza, la qualità dell'intero lavoro. Se quindi vi è un "risparmio" in termini di utilizzo del personale direttamente addetto alla ripresa, vi è comunque la necessità di coinvolgere nel processo produttivo più revisori per poter garantire l'assoluta aderenza del testo resocontato al parlato che, tra l'altro, deve anche essere sottoposto a un'accurata revisione dal punto di vista linguistico.

7. Gli errori

Mi sia permesso, a questo punto, una considerazione relativamente agli errori che, oltre a inficiare la qualità del lavoro, appesantiscono – in termini di costi – l'intero processo produttivo. L'analisi degli stessi rilevati in fase di revisione è estremamente importante perché ci consente di porre in essere delle strategie per di migliorare la dettatura e ci dà inoltre delle indicazioni sull'opportunità di arricchire il dizionario inserendo termini stranieri, nomi propri ecc. Per quanto riguarda gli "errori" che, impropriamente siamo soliti definire "errori di Dragon" possiamo affermare con una certa sicurezza che, a partire dalle ultime release, Dragon NaturallySpeaking ha un'affidabilità altissima e certificata.

[Torna all'indice](#)

Gli errori che vengono rilevati in fase di correzione sono essenzialmente imputabili a una dizione poco nitida da parte del respeaker. Nella dettatura di un testo in velocità una particella pronominale, una desinenza, un monosillabo o una parola breve (due sillabe) possono essere pronunciate in maniera poco chiara “inducendo” il software a commettere un errore nella trascrizione (es. “non ci può aiutare” – “non si può aiutare”; “trattasi di mero errore procedurale” – “trattaci di vero errore procedurale”). La fase di addestramento non si può quindi considerare conclusa con l’installazione del programma sul PC, ma è da quel momento che il programma entra in simbiosi con il respeaker con il quale si instaura una relazione biunivoca: quanto viene “dato” viene poi “restituito” in termini di precisione. Una considerazione a parte può essere fatta relativamente ai termini stranieri. Poter arricchire il vocabolario di parole straniere, siano esse sostantivi, verbi, locuzioni, nomi propri ecc. è sicuramente un punto di forza del software in quanto garantisce l’esatta trascrizione. Per quanto possibile, soprattutto allorquando si dispone di un minimo di documentazione relativamente all’evento che deve essere resocontato, l’inserimento di nuovi termini è un’operazione quanto mai opportuna perché ci evita di cadere in possibili errori soprattutto nella trascrizione di nomi propri affini tra loro (esempio: Kurt Schmitt (economista tedesco) – Adam Smith (economista/filosofo scozzese) – Brian Schmidt – premio Nobel per la fisica 2011); e più semplicemente: Mc Donald’s – Mac Donald). Si tratta spesso di grafie molto simili ma che identificano personaggi molto diversi tra loro, non sempre noti al vasto pubblico ma – ovviamente – familiari ai relatori che li citano, non raramente in modo concitato e impreciso.

8. Considerazioni

Onde evitare possibili strumentalizzazioni o accuse di “partigianeria”, ovvero di privilegiare una tecnica di ripresa rispetto a un’altra, sottolineo ancora che l’esperienza dello Studio Acta si basa essenzialmente su attività nell’area della resocontazione assembleare e congressuale, non occupandosi di verbalizzazione giudiziaria che ha delle connotazioni ben precise, comunque diverse da quelle che contraddistinguono il resoconto congressuale. In quest’ambito, infatti, al respeaker viene chiesto di operare una prima “pulizia” del testo da ridettare al PC, ovviamente nei limiti imposti dalla fedeltà concettuale del resoconto. Questa fase – peraltro molto importante nella resocontazione quanto sconosciuta ai non addetti ai lavori – non è invece ammessa nella verbalizzazione giudiziaria. In un’aula giudiziaria il professionista (indipendentemente dallo strumento o dal software utilizzato) deve limitarsi a riprendere ciò che viene effettivamente detto, senza alcun intervento sotto il profilo editoriale. Non deve cioè correggere errori o disambiguare concetti poco chiari. Questa discriminante può essere definita la linea di demarcazione che, tra gli altri elementi, contraddistingue il processo attraverso il quale si perviene alla produzione di un verbale da quello più articolato che porta alla stesura degli atti congressuali o di un verbale di un organismo elettivo/rappresentativo. Nella

[Torna all'indice](#)

pratica quotidiana e laddove questo sia possibile, la scelta dello strumento è funzionale al prodotto che viene richiesto (verbale o resoconto) valutando anche presente gli ambiti operativi e i vincoli imposti dal committente (tempi di consegna, location dell'evento, supporti audio disponibili ecc.), le risorse disponibili, i margini operativi sotto il profilo economico. Per tale motivo non escludiamo che si possano utilizzare, per il medesimo evento, sia la stenotipia che il respeaking, anche a distanza.

9. Conclusioni

In ogni caso, indipendentemente dalla tecnologia usata, voglio ribadire che è una nostra precisa scelta aziendale – che si è rivelata nel tempo strategicamente vincente – quella di non consegnare mai un testo così come viene ripreso dal resocontista e ciò indipendentemente dal software (Dragon NaturallySpeaking) o dall'hardware (Stenograph) utilizzato. Tutto il materiale viene accuratamente controllato e corretto. Proprio per fornire al cliente un prodotto ineccepibile, l'impegno dei nostri operatori è massimo anche nella fase successiva alla ripresa del parlato. In alcuni casi, soprattutto quando si tratta di linguaggi non usuali, ovvero allorquando il registro linguistico dei relatori è alquanto "personale", o si è in presenza di un linguaggio estremamente tecnico o di un'esposizione che caratterizza per la presenza di periodi involuti, ridondanti ecc., non si può prescindere da una revisione accurata del testo ripreso sotto il profilo linguistico.

Talvolta, sia pur raramente, è opportuno procedere addirittura al riascolto del file audio perché il rischio di possibili ambiguità e di errori è sempre in agguato. In altre parole, poiché noi intendiamo sempre privilegiare l'accuratezza del testo rispetto ad altri elementi, ogni considerazione sui costi conseguenti assume un valore relativo. Un testo, sia pure ripreso dal più veloce professionista e quindi consegnato al cliente contestualmente con il termine dell'intervento, non ha alcun valore se è impreciso o, peggio, contiene qualche termine che, in fase di revisione, si rivela palesemente errato e soprattutto compromettente per quanto riguarda la fedeltà concettuale al testo effettivamente pronunciato dall'oratore. La *customer satisfaction* – che si conquista fornendo elaborati ineccepibili – è un obiettivo per noi irrinunciabile perché ha una duplice valenza: si riflette positivamente su ogni collaboratore, ne valorizza la professionalità stimolandone al contempo il miglioramento delle prestazioni. È, infine, lo strumento più efficace per un'effettiva fidelizzazione del cliente e che permette di guardare con fiducia al futuro.

8b. PROPOSTA DEL FORUMTAL SUL VERBALE MULTIMEDIALE DI ATTI GIUDIZIARI AUDIOREGISTRATI

Francesco Cutugno e Andrea Paoloni

1. Introduzione al verbale multimediale

Anche nel settore della Giustizia, come in molti altri settori della pubblica amministrazione, è giunto il momento di fare ricorso a tecnologie moderne nel campo del trattamento automatico del linguaggio e della multimedialità. I tempi sono ormai maturi, e gli operatori nel campo della Giustizia possono oramai agevolmente disporre di trascrizioni degli atti giudiziari in forma multimodale, vale a dire che è possibile consultare gli atti processuali audioregistrati in aula e successivamente trascritti, accedendo ad un unico documento digitale in cui visualizzare il testo trascritto e ascoltare la corrispondente porzione dell'audio. Su documenti di questo tipo possono essere utilizzate delle funzioni di ricerca grazie alle quali raggiungere velocemente il punto di interesse. Le tecnologie per realizzare tutto questo sono ormai disponibili e il mercato offre più di una soluzione; in una recente sperimentazione condotta da alcuni soggetti per conto del Ministero della Giustizia se ne sono utilizzate un paio.

All'interno della iniziativa sulla agenda digitale per la pubblica amministrazione, in relazione alla messa a punto di tali sistemi, si pongono però dei vincoli cruciali: la soluzione tecnologica che si adotterà per il verbale multimediale dovrà essere basata su un formato aperto e condiviso. In questo ambito la parola "aperto" indica il ricorso a tecnologie *open source* in linea con scelte analoghe che si conducono in altri settori della pubblica amministrazione mentre la parola "condiviso" implica un processo di standardizzazione o comunque di uniformazione atto a garantire che la scelta tecnologica sia unica e indipendente dalle sedi e dagli uffici che affidano a terzi il processo di trascrizione e allineamento audio dei verbali. In sintesi il ForumTal vorrebbe avviare i lavori per predisporre uno standard per il formato dei documenti multimediali che soddisfi i requisiti posti dall'utilizzatore finale, vale a dire l'operatore della Giustizia, e che al contempo sia accettato da tutti i produttori e fornitori di servizi di trascrizione (basata, quest'ultima, sia su tecniche di riconoscimento automatico del parlato, sia su trascrizione manuale).

2. Controllo della qualità delle trascrizioni

Un secondo tema che al momento si pone maggiormente per coloro che intendono avvalersi di sistemi di riconoscimento automatico del parlato per curare le trascrizioni di audio "pubblico" (in questo caso oltre all'audio nelle aule di Giustizia il discorso può valere per altri casi quali ad esempio i discorsi parlamentari, atti di commissione, eccetera), ma che è di interesse anche nei casi di trascrizioni condotte da operatori umani, è quello del controllo della qualità delle trascrizioni. In altre parole, il secondo problema è la costruzione di un

[Torna all'indice](#)

sistema di verifica oggettiva per la valutazione quantitativa e qualitativa della bontà delle trascrizioni. Le aziende che ricorrono a trascrittori automatici voce/testo, per indicare l'affidabilità dei sistemi proposti, utilizzano come parametro di riferimento il numero di parole correttamente trascritte rispetto al totale delle parole trascritte (il cosiddetto *Word Error Rate – WER in sigla*). I sistemi commerciali allo stato dell'arte, indipendenti dal parlante, per fini generici di dettatura e trascrizione di testi, riportano un WER intorno al 2% se le condizioni acustiche in cui la voce è registrata sono buone.

Questo tipo di parametro di valutazione richiede dei correttivi quando si considera lo specifico dominio semantico in cui rientra il testo da trascrivere. Ad esempio, per la trascrizione degli atti giuridici esiste un lessico complesso, fatto di termini gergali, di evocazione di oggetti ed eventi, di nomi propri e di altri elementi, la cui errata trascrizione genera sicuramente maggiore sconcerto che non l'eventuale errore su un ausiliario o su una generica parola funzionale. Sarebbe di conseguenza opportuno individuare un metodo per dare diverso peso agli errori commessi sulle parole che sono cruciali per la comprensione e la definizione del testo specifico, rispetto alle parole meno "importanti". Anche in questo caso la ricerca scientifica ha a disposizione soluzioni efficaci che devono solo essere propriamente configurate ricorrendo a scelte condivise fra tutti gli attori del sistema delle trascrizioni.

3. Proposta del ForumTAL

Il ForumTAL che ha da sempre mostrato interesse a questi problemi, propone la costituzione di una commissione i cui compiti dovrebbero essere prevalentemente due:

1) fornire indicazioni (se non addirittura proporre uno standard) per il formato dei dati multimediali trascritti e condivisi nelle differenti procure;

2) fornire indicazioni (se non addirittura strumenti off-the-shelf) per la valutazione dei sistemi di trascrizione automatica utili sia a fini di controllo della qualità, sia allo scopo di indurre un circolo virtuoso nel processo di messa a punto a cui i sistemi automatici sono tradizionalmente sottoposti.

La commissione potrebbe essere formata da membri provenienti dai settori della ricerca pubblica in funzione di esperti qualificati e super-partes e da esperti indicati dalle aziende interessate.

La presenza dei rappresentanti delle aziende avrebbe il duplice scopo di favorire l'accettazione delle misure proposte da parte di un numero più ampio possibile di attori sulle scene considerate, oltre a finanziare direttamente l'iniziativa con contributi economici che potrebbero coprire le spese della commissione stessa. Questo tipo di procedura costitutiva prende spunto dalle operazioni messe in atto con successo dal World Wide Web Consortium (W3C).

[Torna all'indice](#)

Sulla base di tutte queste premesse, il 19 febbraio il ForumTal ha riunito in un unico tavolo tecnico le principali realtà pubbliche e private interessate a queste iniziative. I temi sopra esposti sono stati illustrati e tutti i convenuti hanno riconosciuto l'importanza di dare luogo a specifiche commissioni incaricate di portare in tempi ragionevoli a delle proposte.

Inoltre è stato proposto di istituire una ulteriore commissione sulla qualità del segnale audio acquisito, anche perché questo ha fondamentale impatto sulle prestazioni dei sistemi e quindi sulle metodologie e sulle metriche da utilizzare e non ultimo sulle fondamentali definizioni necessarie a definire i SLA (Service Level Agreement) dei sistemi.

9) ASPETTI PROFESSIONALI NEL RESPEAKING IN TEMPO REALE

9a. IL RESPEAKING TELEVISIVO

Arianna Mori

1. Introduzione

Accade a chi si è avvicinato per professione al mondo della sottotitolazione in diretta, di discutere con alcuni conoscenti e di descrivere il proprio lavoro come “respeaking con finalità di sottotitolazione in tempo reale per non udenti”. L'immediato feedback che si ottiene dagli interlocutori, ovverosia telespettatori di tutta Italia, è un gesto delle mani che simulano la battitura a tastiera, seguito dalla classica domanda “li scrivete in tempo reale?”, risposta: “no, respeakeriamo”. Ed ecco che compare un punto interrogativo sul volto dell'interlocutore. Per dissipare quindi talune perplessità in merito al respeaking, nel presente caso svolto presso Colby, azienda milanese che si occupa di sottotitolazione mediante svariate modalità, di seguito verranno descritte le molteplici fasi iniziando dalla metodologia e la collaborazione tra gli operatori, passando per la combinazione competenza personale-tecnologia touch screen, per arrivare alla messa in onda.

2. Il respeaking presso Colby

Occorre porre l'accento su esempi concreti che costituiscono le tre modalità standard di respeaking in diretta adottate da Colby: *Annozero*, conosciuto programma di dibattito politico, che riporta una durata molto elevata rispetto ai casi precedenti, un ritmo intermedio e una maggiore presenza di parlanti; il *Tg3* delle 12, programma di durata ridotta, a velocità elevata e interamente in tempo reale; il *Tg5* delle 20, ugualmente di breve durata, con una velocità intermedia, in parte live, in parte semi live. In genere, la modalità è la seguente: il coordinatore è colui che segnala ai due colleghi quando è il momento di andare in onda e decide chi deve iniziare a respeakerare il lancio del programma, seleziona il pulsante “On Air” per prendere il collegamento con la sede della trasmissione e inizia la messa in onda. I respeaker si alternano al microfono con un intervallo di cinque minuti, approssimativamente, stabiliti da un segnale manuale del coordinatore, seduto nella postazione centrale. Il processo di sottotitolazione vede l'alternarsi della voce dei respeaker, all'inserimento di punteggiatura, colorazione e terminologia sconosciuta al software utilizzato. Il respeaker indossa le cuffie corredate di microfono, ascolta l'eloquio e, riassumendone il messaggio, ripete quanto ha sentito. La voce viene intercettata da un software di riconoscimento vocale, in questo caso Dragon Naturally Speaking, arricchito da altri software aziendali che insieme costituiscono il *Colby Live Subtitling* che manda in sovrapposizione il sottotitolo alla pagina dedicata del teletext.

3. Aspetti tecnici

All'interno dei diversi programmi sono presenti il conduttore, l'inviato, gli ospiti. In che modo, quindi, il non udente può riconoscere chi sta parlando in un dato momento, se il respeaker parla e non scrive? Se nel parlato le pause sono fondamentali per rendere il senso del discorso, come vengono riprodotte, se il sottotitolo viene emesso vocalmente? Ecco che sopraggiunge in ausilio la tecnologia mediante un sistema studiato e brevettato dall'azienda. Per rendere più concreta la spiegazione, cataloghiamo le funzionalità del touch screen citando quelli che dagli operatori del settore sono chiamati "bottoni" e li distinguiamo in bottoni per la **colorazione**, per la **punteggiatura** e per i **vocaboli**.

La **colorazione**: il respeaker vede il parlante, ne sente l'esposizione e prima di riprodurlo digita il colore che desidera attribuirgli. Solitamente il discorso del conduttore è di colore bianco, mentre il secondo conduttore o il giornalista inviato, nel caso dei telegiornali, per esempio, è giallo. I restanti colori a disposizione vengono assegnati via via che si presentano gli ospiti.

La **punteggiatura**: per rendere ottimale il senso del discorso, il respeaker alterna la propria voce alla digitazione su schermo tattile della punteggiatura. Il suo discorso deve quindi interrompersi per un istante per cliccare il bottone corrispondente, bottone che a sua volta è collegato a un comando vocale preregistrato che dà l'impulso al programma, e riprendere il discorso. Il tutto può sembrare complesso, ma con un po' di pratica questa operazione diventa così automatica da non rendersi conto che si sta digitando la punteggiatura.

I **vocaboli**: esistono, all'interno del discorso dei parlanti, alcuni termini che il software potrebbe non riconoscere, essendo pur sempre una macchina e non un essere umano. Il programma, infatti, senza l'intervento dell'operatore potrebbe trascrivere errori di diversa natura, trasposti poi in sovrimpressioni. Per ovviare a questo potenziale ostacolo, è possibile preinserire termini stranieri o nomi propri di persona, di località e così via. Colui che si occupa della trascrizione di tali vocaboli è il coordinatore, seduto tra i due respeaker, che, ascoltando l'eloquio, sentendo un potenziale termine sconosciuto al software lo scrive immediatamente. La parola compare sullo schermo del collega che potrà usufruire nell'immediatezza del bottone corrispondente. È da sottolineare che il coordinatore, dopo alcuni mesi di avviamento alla professione, intuisce quali termini potrebbero inficiare la buona riuscita del sottotitolo e quali, al contrario, possono essere pronunciati senza alcun timore. Permangono altri fattori che contribuiscono a un eventuale errore nel sottotitolo, riconducibili a rumori che il microfono, molto sensibile, potrebbe percepire: l'indecisione dell'operatore nel riprodurre la frase o la variazione improvvisa della tonalità di voce, elementi da evitare per un prodotto ottimale. Non mancano casi di sovrapposizione della voce degli ospiti del programma o dei giornalisti, così come la presenza di canzoni o frasi emesse in coro, e ancora interferenze nell'audio che non permettono un ascolto adeguato. Per risolvere situazioni di questo tipo, si è

[Torna all'indice](#)

dovuto ricorrere all'intervento dei cosiddetti "cartelli". Il respeaker in tali circostanze provvede all'inserimento tramite tastiera del PC di appositi cartelli che segnalano la situazione corrispondente. Un esempio è il cartello SOVRAPPOSIZIONE DI VOCI, così come PROBLEMI AUDIO oppure CORI DI PROTESTA. L'elenco potrebbe proseguire. Menzione a parte meritano le canzoni che, se inserite nella scaletta della trasmissione, vengono precedentemente preparate dai respeaker e dal coordinatore nell'apposito formato mediante uno specifico programma per la divisione dei sottotitoli che poi vengono mandati in onda in corrispondenza del testo cantato. La grafica e i tempi di permanenza del sottotitolo vengono gestiti automaticamente senza intervento degli operatori.

4. Annozero

Per spiegare in modo più comprensibile l'intero processo in cui si intersecano le fasi sopra menzionate, accenniamo alla preparazione di un programma standard di durata considerevole come *Annozero*. È richiesta la presenza di tre operatori, due dei quali respeaker e uno con funzione di coordinatore. Presentandosi in sede di lavoro circa mezz'ora prima dell'inizio della trasmissione, prendono visione della scaletta in cui si rendono noti gli ospiti presenti in studio e le tematiche che verranno trattate. Per trasmissioni a scopo divulgativo, per i talk show e, in generale, per le trasmissioni con la presenza di un cospicuo numero di ospiti, è stato adottato il sistema "legenda". Nel caso di *Annozero* ogni pulsante è stato assegnato a un politico, in modo che compaia in sovrimpressione, unitamente al sottotitolo colorato, il nome e il cognome corrispondente all'ospite.

5. Il Tg3

Passando ai programmi di durata più contenuta, per esempio i telegiornali, è sufficiente la presenza di due respeaker che svolgeranno, alternandosi, sia la funzione di respeaker, sia quella di coordinatore, dandosi il cambio rispettivamente alla voce e alla tastiera del PC. Nel caso del *Tg3*, gli operatori preparano un certo numero di vocaboli prendendo spunto dalle notizie ANSA, sapendo indicativamente quali termini potrebbero essere utili. Un esempio per il telegiornale è il nome di qualche località in cui sono avvenuti fatti di cronaca oppure i nomi dei Ministri che potrebbero essere menzionati nel corso delle notizie. Nella maggior parte dei casi non è disponibile alcun servizio prima del telegiornale, ciò comporta da parte dei respeaker un continuo aggiornamento in modo da conoscere le principali notizie e saper gestire l'eloquio. L'operatore, inoltre, prima della messa in onda inserisce sullo schermo un certo numero di termini che vengono classificati per pagina e per argomento: cronaca, politica, sport, economia e così via, rapidamente selezionabili con un semplice tocco nel corso della trasmissione; programmi di questo tipo prevedono un'elevata rapidità di esposizione. Il respeaker deve essere dotato di notevole concentrazione, predisposizione all'ascolto, riflessi e velocità, tanto nella

[Torna all'indice](#)

riformulazione dell'eloquio, quanto nella battitura, laddove il turno preveda il coordinamento.

6. Il Tg5

Per completare la classificazione delle modalità adottate da Colby per la sottotitolazione in real time tramite respeaking, prendiamo in considerazione il *Tg5*, in cui si inquadra la medesima procedura del *Tg3*, con una differenza: viene adottata la modalità semi live. Il meteo, alcuni servizi e lanci vengono mandati dalla redazione ai respeaker con un certo anticipo rispetto all'inizio del telegiornale, in modo da poterli preparare e mandare in onda già suddivisi, revisionati, colorati e con l'adeguata punteggiatura e di conseguenza senza la necessità di respeakerare. I restanti servizi, che non sono stati programmati dalla redazione o dal giornalista, sono sconosciuti per il respeaker, il quale dovrà pertanto riprendere immediatamente a trasporre simultaneamente al microfono come spiegato precedentemente, tornando quindi al sistema live.

7. Conclusioni

Si fa presente che nella vasta gamma di trasmissioni sono stati presi ad esempio solo alcuni casi, considerando, tuttavia, che ogni programma potrebbe prevedere delle eccezioni e delle particolarità che andrebbero, in altra sede, trattati nello specifico. Nel prodotto finale sul teletext osservato dal divano di casa, ci si accorge della presenza del sottotitolo, delle frasi colorate che scorrono, del discorso a volte più succinto rispetto a quello che si sente, e si pensa alla tecnologia di una tastiera del PC. La tecnologia esiste, è colei che la fa da protagonista sugli schermi dei telespettatori e i suoi errori colgono di sorpresa i respeaker che risolvono prontamente la situazione. Il respeaker **C'È** ma non si vede.

9b. IL LIVE EDITING TELEVISIVO

Silvia Velardi

1. Introduzione

I sistemi di riconoscimento del parlato hanno reso possibile, dal 2001 (Lambourne 2006), lo sviluppo di sistemi di trasmissione di sottotitoli di programmi televisivi in diretta attraverso la tecnica del respeaking.

Per quanto idonei alla tecnica del sottotitolaggio, questi software non possono essere intesi come 'perfetti': la voce umana presenta, infatti, variazioni e livelli di adattamento non del tutto indifferenti ai sistemi. È principalmente per questa ragione che, nel processo di sottotitolazione in diretta del discorso orale da parte del respeaker, il supporto e la coordinazione con altre figure professionali, come il live editor, si rivela fondamentale. Il presente contributo vuole offrire una panoramica sulle caratteristiche funzionali del servizio di live editing, il processo di revisione del parlato in tempo reale, in accostamento al respeaking.

2. L'editor in aiuto al respeaker

Numerosi sono gli sforzi psico-cognitivi a cui il respeaker è soggetto durante la fase di dettatura dei sottotitoli in tempo reale (Marsh 2004). Tra questi:

- vincoli temporali dettati dalle emittenti dei programmi televisivi in onda;
- la necessità di una pronuncia chiara e attenta delle parole: la dettatura deve avvenire nel modo più chiaro possibile, prestando particolare attenzione alla pronuncia di lessico specifico, in quanto il software sarà in grado di riconoscere l'esatta pronuncia solo se si avvicinerà a quella dettata durante la fase di addestramento del sistema (Eugeni 2009);
- un buon livello di gestione dello stress;
- la difficoltà del passaggio da una lingua all'altra in modalità simultanea (respeaking interlinguistico);
- l'abilità di manipolare e strutturare il discorso di partenza in maniera comprensibile per i destinatari dei sottotitoli, con l'obiettivo di garantire una lettura e comprensione più agevole del discorso di arrivo. Tutto questo deve avvenire dividendo l'attenzione tra fasi diverse e concomitanti, vale a dire ascolto analitico del testo di partenza e dettatura e controllo del testo in uscita, rispettando un giusto grado di décalage anche quando il testo originale si rivela essere pronunciato velocemente.

Risulta quindi chiaro, da un punto di vista cognitivo, che il respeaker è impegnato e coinvolto in una serie di processi molteplici e articolati, al punto

[Torna all'indice](#)

che, se dovesse aggiungere ulteriori skills prima della messa in onda del testo di arrivo finale, incorrerebbe in una saturazione delle abilità cognitive in atto, diventando lo sforzo mentale enorme e di difficile gestione. Per questa principale ragione, la figura professionale spesso in affiancamento al respeaker è l'editor che supporta il lavoro del respeaker in tempo reale, attraverso la correzione, la revisione, e la messa in onda dei sottotitoli in modalità live (live editing).

3. L'editor e la postazione di sottotitolazione

Diversi sono gli approcci e le interfacce utilizzate nel live editing, a seconda delle necessità, delle abilità e del livello di addestramento dell'editor o degli editor. In genere, la postazione è costituita da diversi moduli: di acquisizione del segnale audio-video; del doppio riconoscitore del segnale audio per le modalità respeaking e trascrizione automatica; dell'editing; della gestione dell'invio al sistema di sottotitolazione in tempo reale.

Uno dei grandi vantaggi, senza dubbio, deriva dalle funzionalità della tastiera del PC, programmata sia per navigare sullo schermo, sia per correggere, così da non rallentare il processo di correzione con l'ausilio di un mouse che, in questo caso, impedirebbe movimenti diretti e veloci sullo schermo, che sfrutta invece una tecnologia "touch". L'interfaccia del programma permette quindi di lavorare e accedere a varie funzionalità in modo pratico e semplice, realizzando così modifiche sul testo in uscita. Nel live editing, la postazione di sottotitolaggio in diretta (tabella 1) è utilizzata in maniera simultanea da due operatori. Le funzioni dei due operatori variano a seconda della modalità di sottotitolazione che si decide di adottare: respeaking o trascrizione automatica.



Tabella 1: interfaccia di un software di sottotitolazione in tempo reale (vedi § 4)

[Torna all'indice](#)

Oltre all'ausilio dello schermo, la postazione di sottotitolazione può prevedere anche l'ausilio di una doppia pedaliera, alla quale si può abbinare una o più funzioni, associando così a ogni pedale appositi comandi creati, come la cancellazione di una parola e l'invio dei sottotitoli sullo schermo. In questa maniera, si permette all'editor di verificare in maniera ulteriore i sottotitoli prima del loro invio definitivo all'emittente.

4. L'editor nella sottotitolazione con respeaking

Nella sottotitolazione con respeaking, il respeaker ascolta il discorso dell'oratore di una conferenza o il programma televisivo da sottotitolare e lo ripete dettandolo al software di riconoscimento del parlato. L'editor, con l'ausilio di una piattaforma che permette la gestione del testo in uscita dal server, è impegnato nel processo di revisione e correzione, nonché della successiva (quasi immediata, con uno scarto di pochi decimi di secondo) messa in onda dei sottotitoli.

L'operatività dell'editor, che deve essere ottimale e la più automatizzata possibile, è schematizzabile nelle seguenti operazioni (vedi tabella 1):

1. *ascoltare e memorizzare non solo del testo di partenza, ma anche delle parole pronunciate dal respeaker, che potrebbe riformulare il testo pronunciato dagli oratori;*
2. *identificare gli errori nel testo trascritto prossimo a essere mandato in onda;*
3. *correggere eventuali errori nel testo trascritto;*
4. *inserire eventuali segni di interpunzione;*
5. *rendere maiuscolo o minuscolo un termine;*
6. *abbinare colori diversi a ogni oratore diverso;*
7. *inserire i verbi ausiliari, le parole più utilizzate, gli articoli, le preposizioni, le congiunzioni e gli elementi più ricorrenti nel parlato dalle colonne al di sopra, a destra e sinistra dell'area centrale di editing;*
8. *selezionare le "parole a distanza uno" che vengono proposte automaticamente dal software ogni volta che evidenzia un termine;*
9. *assicurarsi che il sottotitolo abbia un senso logico con quello precedentemente mandato in onda.*

5. L'editor nella sottotitolazione con trascrizione automatica

Oltre a sistemi di sottotitolazione con respeaking, che necessitano di un addestramento vocale specifico della macchina alla voce del respeaker, il live editing può prevedere ulteriori modalità di lavoro, come l'attivazione della

[Torna all'indice](#)

sottotitolazione con trascrizione automatica, diversa da quella tradizionale, perché basata su un sistema di riconoscimento vocale speaker independent in grado di trascrivere la voce live dell'oratore indipendentemente dal parlante, senza perciò bisogno di addestramento vocale.

Nello specifico, si tratta di un expertise esclusivo dei sistemi di riconoscimento vocale che permette di rendere fruibile e intelligibile la sequenza multimediale del parlato per mezzo di una decodifica automatica dell'audio che garantisce, a sua volta, la trascrizione dei contenuti senza necessità che l'operazione di dettatura e restituzione del testo di partenza sia compiuta propriamente dal respeaker.

Esistono comunque dei limiti. Non sempre, infatti, il testo di partenza può essere riconosciuto in toto nella modalità di trascrizione automatica; questo si verifica quando, ad esempio, vengono pronunciate parti del discorso che non possono essere riconosciute automaticamente dal sistema, come parole in forma dialettale, discussioni di più oratori che si sovrappongono, nomi propri o in lingua straniera non segnalati prima della trascrizione. È per questa ragione che, quando viene utilizzata la modalità di trascrizione automatica di sottotitolazione in diretta, l'editor è comunque chiamato a svolgere la sua funzione essenziale di revisore live.

La trascrizione automatica rappresenta una tecnologia estremamente interessante, ma indubbiamente non immune da un'accurata messa a punto del sistema in funzione dell'ottimizzazione del testo di arrivo. Ragion per cui l'intervento dell'editor a livello sintattico o testuale del testo multimediale di partenza si rivela prezioso e fondamentale.

6. Il doppio-editing nella sottotitolazione

Nel live editing, non è d'obbligo che una persona debba riscoprire uno e un solo ruolo. Può accadere, invece, che la stessa persona abbia le abilità adeguate e necessarie a svolgere diverse funzioni relative a diversi profili.

Può accadere, infatti, che il respeaker sia anche editor (o viceversa). In un caso simile, è possibile la messa in gioco di una modalità di sottotitolazione ulteriore che rende possibile l'operazione di live editing grazie alla cooperazione dei due profili professionali che gestiscono la revisione del materiale multimediale in uscita, alternando le loro funzioni.

Nel caso specifico di **sottotitolazione con trascrizione automatica**, un respeaker che sia anche editor, potrà alternarsi e diventare editor a sua volta, prendendo così il posto del suo collega. In questo caso, il respeaker diventa live editor, e sarà in grado di utilizzare le skills necessarie per procedere alla correzione di eventuali errori durante la trascrizione automatica dei sottotitoli.

[Torna all'indice](#)

Nella **sottotitolazione con respeaking**, può verificarsi l'eventualità in cui entrambi i profili professionali siano interscambiabili nelle loro funzioni; il vantaggio che deriva dall'alternanza delle due figure diventa ancora più tangibile. Infatti, in un contesto simile, i due profili professionali (editor e respeaker) potranno scegliere di alternarsi a intervalli di tempo più o meno regolari, e di compiere, rispettivamente, le operazioni di dettatura del parlato multimediale (testo di partenza) e la revisione del testo di arrivo sostituendosi a vicenda.

7. Conclusioni

Diverse sono le modalità di sottotitolazione che coinvolgono le abilità e le funzionalità dell'editor: sottotitolazione con respeaking, trascrizione automatica, e doppio-editing. In un quadro complessivo, il live editing richiede una varietà di skills cognitive e non automatiche non indifferenti che rendono questo lavoro molto complesso. Per questa ragione, sono stati messi a punto numerosi esercizi di training, fondamentali per sviluppare e acquisire le tecniche operative adeguate, a garanzia del funzionamento di programmi specifici e dei giusti strumenti orientati al nuovo profilo professionale del futuro.

PARTE III: RICERCA E SVILUPPO

[10. Aspetti teorici del rispeaking \(Francesca Marchionne\)](#)

[11. Metodi e tempi di una verbalizzazione – prove tecniche \(Rossella D’Arcangelo e Francesco Cellini\)](#)

[12. Sottotitoli di conferenza intralinguistici e interlinguistici – strategie traduttive a confronto \(Silvia Bellucci\)](#)

[13. Il respeaking interlinguistico \(Francesca Marchionne\)](#)

[14. Progetti dell’Associazione Internazionale di Respeaking onA.I.R. \(Tiziana Trapani\)](#)

[15. Novità tecnologiche](#)

15a. [Il progetto europeo EU-Bridge – Un ponte per colmare il divario linguistico \(Margit Rödder\)](#)

15b. [Il progetto europeo SAVAS – Riconoscimento del parlato per respeaking e sottotitolazione in tempo reale \(Carlo Aliprandi\)](#)

15c. [Il progetto europeo SUMAT – Sottotitolazione assistita dalla traduzione automatica \(Panayota Georgakopoulou\)](#)

[Torna all'indice](#)

10) ASPETTI TEORICI DEL RISPEAKERAGGIO TELEVISIVO

Francesca Marchionne

1. Introduzione

Il respeaking è una tecnica relativamente nuova e pertanto il suo studio teorico non vanta di una lunga tradizione letteraria. Solo nell'ultimo decennio, infatti, professionisti e ricercatori del settore hanno dato il loro prezioso contributo teorico al respeaking. Così quella che era semplicemente una tecnica di scrittura veloce che sfrutta un software di riconoscimento del parlato sta diventando lentamente una vera e propria disciplina di studio. D'altronde è ciò che a suo tempo è successo con altre discipline umanistiche come la traduzione, quando persone come Jakobson e Holmes si battevano per uno studio teorico della materia. Il parallelismo non è casuale. Prima di tutto il respeaking è una forma di traduzione dal momento in cui per traduzione si intende un **processo** nel quale un testo di partenza (TP) è rielaborato da un professionista al fine di formulare un testo di arrivo (TA) - il **prodotto** - la cui **funzione** è quella di permettere a un pubblico definito di accedere a un testo altrimenti inaccessibile in parte o totalmente. In secondo luogo è mia intenzione proporre in questo articolo l'esplorazione del respeaking basandomi su una mappa che riprende quella elaborata da Holmes nel 1972 e modificata da Toury nel 1991 nell'ambito dei *Translation Studies*. Tale mappa si proponeva, infatti, di esemplificare i settori inerenti la ricerca traduttologica pratica e teorica e può essere facilmente adattata per approfondire lo studio del respeaking in tutti i suoi aspetti, secondo la riformulazione che si propone di seguito:

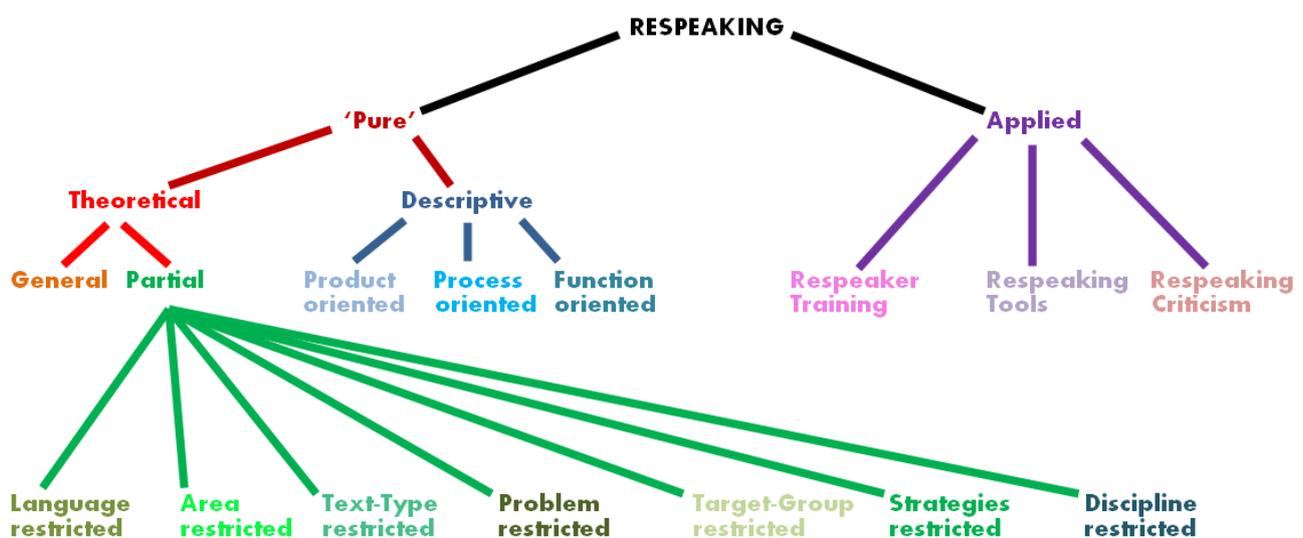


Figura 1 – "La mappa degli studi sul respeaking": mappa per uno studio teorico e pratico del respeaking.

[Torna all'indice](#)

Come si può evincere dalla mappa, teoria e pratica non possono essere scisse perché inevitabilmente interconnesse. Appare quindi ovvio che dalla sperimentazione e dalla ricerca vengono tratte delle linee teoriche che, a loro volta, possono dare spunto a nuove ricerche pratiche e così via, assicurando uno sviluppo sempre crescente alla disciplina. Per questi motivi il presente articolo non può di certo essere esauriente, ma può inquadrare in generale gli aspetti teorici del *respeaking* dando numerosi spunti per degli approfondimenti futuri.

2. "La mappa degli studi sul *respeaking*"

Secondo lo schema illustrato nella figura 1, vi sono due principali aree di ricerca, *pure* e *applied*.

- I. La **ricerca pura** ha due obiettivi legati l'uno all'altro:
- **descrittivo**: descrivere il fenomeno così come si manifesta nell'esperienza pratica (*Descriptive Respeaking Studies* – DRS);
 - **teorico**: stabilire dei principi generali per mezzo dei quali tale fenomeno può essere spiegato e praticato (*Respeaking Theory* – RTh).

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

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

La RTh si suddivide in:

- **teoria generale**: mira a spiegare e considerare l'atto del *respeaking* nei suoi elementi basilari, rilevanti per il *respeaking* nel suo complesso. È un obiettivo ambizioso in quanto vuole creare delle basi valide per qualsiasi possibile applicazione del *respeaking*;
- **teorie parziali**: vogliono spiegare e considerare l'atto del *respeaking* secondo parametri ben precisi: **lingua** (*respeaking* intralinguistico e interlinguistico); **area** (in una data area culturale, sociale, linguistica); **tipo di testo** da *respeaking* (programma televisivo, conferenza, lezione universitaria, programma radiofonico ecc.); **problema** (in base ai problemi specifici del *respeaking*, come l'interazione uomo-macchina, la correzione istantanea degli errori ecc.); **target** (sordi,

[Torna all'indice](#)

ipoudenti, stranieri e normoudenti); **strategie** utilizzate per effettuare il respeaking (omissione, espansione ecc.); **ambito di applicazione** (convegnistica, sottotitolazione TV, resocontazione ecc.).

II. La **ricerca applicata** si divide in tre ramificazioni a loro volta suddivise in settori specifici:

- **formazione** del respeaker: preparazione ed esperienza;
- **strumenti** del respeaking: uso delle nuove tecnologie (software di riconoscimento del parlato, software di sottotitolazione ecc.);
- **critica** del respeaking: valutazione dei risultati da parte di esperti e pubblico di destinazione.

Dal momento che approfondire tutti i punti appena elencati richiederebbe un volume a parte, si cercherà nel paragrafo successivo di ricavare una visione teorica generale del respeaking passando necessariamente per l'analisi degli studi descrittivi del respeaking e lasciando agli altri articoli del presente volume il compito di approfondire altri punti.

3. I DRS

1.1 Il prodotto

Il prodotto del respeaking è diverso a seconda delle seguenti variabili diversamente combinabili tra loro:

- ✓ campi applicativi → televisivo, teatrale, cinematografico, messale, radiofonico, universitario, convegnistico, concertistico ecc.
- ✓ fini → accessibilità in tempo reale, resocontazione, verbalizzazione, traduzione ecc.
- ✓ utenti finali → sordi oralisti, sordi segnanti, ipoudenti, udenti italiani, udenti stranieri
- ✓ lingua/e → nella stessa lingua del TP o in un'altra lingua
- ✓ condizioni di lavoro → in loco, a distanza; in diretta, in semidiretta o in differita
- ✓ layout → su una o più righe; a blocchi o parola per parola; con la presenza di video/immagini o senza ecc.
- ✓ convenzioni → colori, punteggiatura, segni grafici ecc.

Un verbale è sicuramente diverso da una sottotitolazione o da una lezione universitaria così come una sottotitolazione intralinguistica per non udenti è diversa da una sottotitolazione interlinguistica per udenti ecc. Tuttavia,

[Torna all'indice](#)

possiamo individuare un prodotto principale che racchiude tutti i singoli prodotti del respeaking: si tratta di un testo costituito da una o più righe che si compone pezzo per pezzo in una determinata lingua e per un determinato tipo di pubblico, con un ritardo di trasmissione variabile rispetto all'enunciazione del TP e a partire da un prodotto audiovisivo o solamente auditivo che può essere in diretta, in semidiretta oppure registrato. In particolare, il prodotto è, secondo Eugeni, una traduzione intersemiotica perché dal canale orale si passa a quello scritto; supersemiotica perché aggiunge la componente scritto-verbale al TP senza sostituirne le altre componenti (nel caso di sordi oralisti e normoudenti), ma diasemiotica quando sostituisce completamente la componente audio (nel caso di Sordi segnanti, i quali non sentono nessun suono e non leggono le labbra in quanto la loro lingua materna è la LIS), per cui i canali di comunicazione possono essere sempre due (video-non verbale e scritto-verbale), ma diversi rispetto al TP; intralinguistica o interlinguistica; contenente alcuni elementi audio verbali e non verbali del TP; asincrona perché trasmessa sullo schermo pochi secondi dopo l'enunciazione; realizzata secondo alcune convenzioni (Eugeni 2008:14).

1.2 Il processo

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

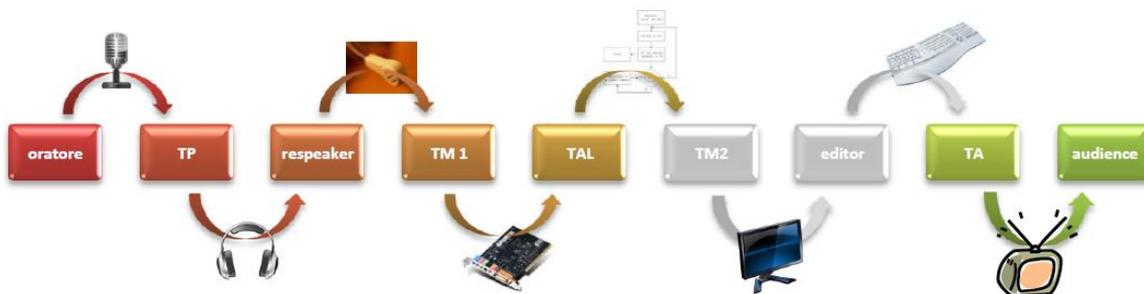


Figura 2 – Il processo del respeaking.

Il rispeaker ascolta ed eventualmente guarda il TP, lo riformula, lo traduce o lo ripete oralmente, creando un testo "di mezzo" (TM1) per la macchina che, tramite appositi software, elabora e trasforma l'input vocale in testo elettronico. Questo testo può costituire il TA nel caso in cui non sia prevista una fase di editing (da qui il colore grigio delle caselle), oppure può costituire il testo "di mezzo" 2 (TM2) da verificare, eventualmente correggere e trasmettere solo successivamente il testo agli utenti finali (TA).

Ovviamente, tutto ciò ha un prezzo: la percezione, la ricezione e la comprensione del TP, l'elaborazione mentale da parte del rispeaker per il passaggio dal TP al TM1, l'interazione uomo-macchina, la trasmissione da un software all'altro, la verifica e, infine, l'eventuale correzione della trascrizione e

[Torna all'indice](#)

l'invio del testo finale richiedono del tempo che oscilla dai 2 (senza correzione) agli 8 secondi, creando un ritardo tra la produzione del TP e la comparsa del testo corrispondente. In particolare, il processo è, secondo Eugeni, una traduzione interlinguistica o intralinguistica, simultanea, isosemiotica perché utilizza lo stesso numero di canali semiotici del TP per il TM1 e intrasemiotica perché utilizza lo stesso canale di produzione, cioè quello orale (Eugeni 2008:15).

1.3 La funzione del prodotto e la funzione del processo

Quanto alla funzione del prodotto, questa dipende ovviamente dal tipo di prodotto realizzato. In generale, il respeaking deve permettere, a un pubblico di destinazione ben definito e tramite specifici accorgimenti, la comprensione del TP, sia esso in tempo reale (accessibilità) o a posteriori (verbalizzazione, resocontazione e traduzione).

Nel caso di respeaking interlinguistico, pressoché inesistente, il testo deve consentire la comprensione del TP a un pubblico di lingua e cultura diverse rispetto a quelle cui è originariamente destinato.

Nel caso di respeaking intralinguistico, ormai sempre più praticato in vari Paesi, il testo deve rendere accessibile un TP prevalentemente a un pubblico di sordi. Ovviamente, diversi *target* richiedono diversi accorgimenti ed è perciò necessario partire sempre da un'analisi approfondita del bacino di utenza per poter poi riuscire nella realizzazione di un prodotto funzionale.

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

4. Conclusioni

Da questa breve analisi si può ricavare una visione teorica generale del respeaking, da concepire come una tecnica di scrittura veloce in cui un operatore ascolta l'audio originale di un qualsiasi evento e simultaneamente lo detta, includendo la punteggiatura ed eventualmente alcune informazioni para- ed extralinguistiche (nel caso il prodotto sia volto a un pubblico di audiolesi), a un software di riconoscimento del parlato che, interfacciandosi con un altro software di scrittura o di sottotitolazione, converte gli enunciati riconosciuti in testo visibile dagli utenti finali in diversi formati e su diversi supporti con un ritardo più o meno maggiore rispetto all'enunciazione del TP e con la presenza o meno di un editor che corregga il testo prima della sua messa in onda. In particolare il respeaking è una tecnica:

[Torna all'indice](#)

- ✓ il cui prodotto è un testo costituito da una o più righe che si compone pezzo per pezzo, con un ritardo di trasmissione variabile rispetto all'enunciazione del TP;
- ✓ sia per la produzione di testo on-line (in tempo reale), sia per la produzione di testo off-line (in preregistrato);
- ✓ il cui prodotto è fruibile sia in tempo reale, sia a posteriori;
- ✓ il cui processo implica necessariamente il coinvolgimento di un respeaker (ed eventualmente di un editor), l'utilizzo di un software di riconoscimento del parlato e un passaggio da un TP a un TM a un TA;
- ✓ il cui obiettivo è quello di permettere a un pubblico definito di accedere a un testo altrimenti inaccessibile in parte o totalmente;
- ✓ intralinguistica o interlinguistica;
- ✓ realizzabile in qualsiasi area geografica, culturale, sociale e linguistica (a patto che esista il software di riconoscimento del parlato in quella data lingua o un respeaker in grado di tradurre in un'altra lingua per cui esiste il software);
- ✓ applicabile a qualsiasi tipo di testo orale, sia esso un prodotto audiovisivo o solamente auditivo in diretta, in semidiretta oppure registrato;
- ✓ utilizzabile in qualsiasi evento, in diretta o in differita;
- ✓ che può presentare degli ostacoli tecnici e psico-cognitivi obiettivi quali l'interazione uomo-macchina, la trasmissione dell'audio, la simultaneità dell'ascolto del TP e della produzione del TM da parte del respeaker ecc.;
- ✓ che prevede l'uso di strategie volte ad aggirare i diversi ostacoli riscontrabili;
- ✓ che può essere appresa con una formazione adeguata in breve tempo e con l'esperienza;
- ✓ che necessita di continui miglioramenti ottenibili grazie alle critiche costruttive, alla pratica, alla sperimentazione e alla ricerca.

Creare una teoria generale del respeaking è un obiettivo ambizioso in quanto le variabili da tenere in considerazione, come abbiamo visto, sono veramente tante e le potenzialità del respeaking sono infinite. Tuttavia, tutti i punti appena elencati, per quanto necessariamente vaghi e generici, costituiscono la base della tecnica del respeaking. A questi bisogna aggiungere, per ogni voce de "la mappa degli studi sul respeaking", un approfondimento adeguato e non riassumibile in un articolo. Si possono, per esempio, compiere degli studi comparati a partire dall'analisi dei testi di diversi prodotti audiovisivi rispeakerati, confrontandoli tra loro.

Tali prodotti devono essere selezionati a loro volta in base ad alcuni parametri, secondo la ricerca che si vuole compiere. In campo televisivo, per esempio, se si vuole sapere quali sono le strategie utilizzate per sottotitolare un determinato genere televisivo, si prenderanno in considerazione diversi programmi appartenenti a un unico genere televisivo. Se si vuole capire, invece, la differenza tra le strategie utilizzate in un prodotto intralinguistico e quelle

[Torna all'indice](#)

utilizzate in un prodotto interlinguistico, si sceglierà uno stesso programma che, a parità di condizioni di lavoro, è stato rispeakerato nella stessa lingua del TP e in una diversa lingua di arrivo. Un altro esempio di studio comparato può avere come obiettivo quello di capire quale tipo di sottotitolazione è più consona per un certo tipo di programma. In questo caso, si metteranno a confronto sottotitoli *verbatim* (parola per parola) e sottotitoli adattati, sottotitoli per udenti e sottotitoli per non-udenti, sottotitoli con funzione didattico-formativa e sottotitoli informativi, di cui si valuteranno gli errori più frequenti, la struttura, la sintassi, il registro e il lessico utilizzati ecc.

Così facendo, si consente l'avanzamento della ricerca e si contribuisce al miglioramento della pratica del respeaking, in maniera tale che quest'ultimo sia il più funzionale possibile. Per esempio, lo studio approfondito del processo e della psicologia del respeaking ha dato degli spunti teorici per sviluppare nuove interfacce di respeaking e di editing, che sono sempre più fruibili. Da questo punto di vista anche la tecnica conosce sempre più avanzamenti, con gli operatori che imparano a gestire gli sforzi psico-cognitivi (soprattutto grazie ai contributi che provengono dagli studi sull'interpretazione simultanea).

La catena si conclude in questo caso con l'organizzazione e la realizzazione di corsi di formazione professionale sempre più mirati. Infine, non bisogna dimenticare che degli studi esplicativi sull'impegnativo e faticoso lavoro del respeaker potrebbero anche 'addolcire' le critiche sul respeaking, che portano il pubblico a lamentarsi della scarsa qualità dei prodotti sottotitolati in tempo reale, permettendo così la nascita di una collaborazione tra respeaker e pubblico per una sempre maggiore applicazione di tale tecnica.

11) METODI E TEMPI DI UNA VERBALIZZAZIONE – PROVE TECNICHE

Rossella D’Arcangelo e Francesco Cellini

1. Introduzione e metodologia

La sbobinatura o trascrizione di un brano “parlato” è la trasposizione dell’oralità in scrittura. In questo articolo analizzeremo metodi e tempi di una breve trascrizione. Più propriamente parleremo qui di verbalizzazione, cioè la pratica di trascrivere “*verbum pro verbo*” l’oralità in scrittura. Abbiamo ritenuto necessario affrontare delle prove tecniche usando tre diversi brani della durata di cinque minuti ciascuno che abbiamo denominato “Brano lento”, “Brano medio” e “Brano veloce”. I termini “lento”, “medio” e “veloce” si riferiscono alla velocità del parlato dell’oratore che si va a trascrivere: 90 parole al minuto per il Brano lento, 113 parole al minuto per il Brano Medio, 134 parole al minuto per il Brano veloce. Si è deciso nella fase delle prove di utilizzare tre brani selezionati in ambito assembleare, ma contenenti il maggior numero possibile di variabili differenti. Questo perché il parlato non è ben definibile e quindi non rappresentabile aprioristicamente. All’interno dello stesso testo sono contenute infatti troppe sottovariabili come diverse velocità di eloquio, diverse tonalità di voce, animosità, timbri, grammatica, solo per citarne alcune. Tuttavia l’ipotesi iniziale è che trascrivere un parlatore pacato, corretto e lineare nella sintassi, non troppo rapido e affatto eccentrico sia più facile e rapido da trascrivere rispetto a un testo avente caratteristiche diverse.

I tre diversi brani si differenziano tra di loro anche per quanto riguarda la qualità dell’audio⁸, che risulta buona nel “Brano lento” e meno buona negli altri due brani, caratterizzati da una registrazione disturbata da “fruscio”. Le prove di verbalizzazione sono state eseguite da due operatori professionisti che denomineremo “Operatore 1” e “Operatore 2”. L’Operatore 1 ha utilizzato il metodo della Dattilografia e il metodo del Respeaking, l’Operatore 2 ha utilizzato il metodo della Stenotipia e della Dattilografia. Inoltre sia l’Operatore 1 sia l’Operatore 2 hanno lavorato con i rispettivi metodi anche su documenti trascritti con il sistema della Trascrizione automatica. Più nel dettaglio, ogni brano è stato trascritto con i seguenti metodi:

- Trascrizione Automatica
- Trascrizione Automatica corretta mediante Stenotipia
- Trascrizione Automatica corretta mediante Dattilografia

⁸ I fattori considerati per determinare la qualità dell’audio sono: l’estensione del file (AAC, AC3, AIFF, APE, AMR, BWF, Cda, FLAC, IFF, MID, MKA, MP1, MP2, MP3, ecc.), la velocità di campionamento (8000 Hz – telefono, 11015 Hz, 22050 Hz, 32000 Hz – DAT Long Play, 44100 Hz – CD, 48000 Hz – DAT, 96000 Hz – HQ Mastering; tipo di ambiente (chiuso o aperto); voce degli oratori; rumori ambientali; sistema di registrazione (analogico o digitale), impianto microfonico (panoramico, direzionale, mono, stereo, ecc.).

[Torna all'indice](#)

- Trascrizione Automatica corretta mediante Respeaking
- Trascrizione mediante Stenotipia in real time e successiva correzione in Dattilografia
- Trascrizione mediante Stenotipia in differita (con pedaliera per fermare e riascoltare il testo)
- Trascrizione mediante Respeaking in real time e successiva correzione in Dattilografia
- Trascrizione mediante Respeaking in differita (effettuando delle pause di dettatura)
- Trascrizione mediante Dattilografia

Da come è stato trattato ogni brano è comprensibile che i metodi per effettuare una trascrizione sono diversi e non sempre lo stesso metodo è efficace per tutti i tipi di brano. Per efficacia intendiamo il metodo migliore e più rapido per trascrivere un brano lento, medio o veloce che sia, ottenendo un prodotto finito che in questo caso è da intendersi la mera trascrizione parola per parola di ciò che si è sentito direttamente dall'oratore, completo di punteggiatura.

La buona riuscita della trascrizione di qualsiasi tipo di brano dipende anche dalle tecnologie che si vanno ad adoperare. Tuttavia, il loro utilizzo incide non tanto sul risultato finale, quanto sul tempo impiegato per ottenerlo. Nel nostro caso le tecnologie adoperate sono state le seguenti:

1. Hardware e software utilizzati dall'Operatore 1⁹:

- Dragon NaturallySpeaking 11.0 (software di dettatura)
- DSS Player 2002 Olympus (registratore digitale con software per l'ascolto dei file registrati). Non vi è uso di pedaliera poiché per tale software sono previsti 3 tasti-funzione sulla tastiera che sono F2, F3 ed F4. I primi due servono rispettivamente per indietreggiare e avanzare l'ascolto della registrazione con la funzione "backspace", che fa tornare indietro il brano di un tempo regolabile tra 0 a 3 secondi. L'ultimo svolge la funzione di play e pause o di stop.
- Winword 2007 (software di elaborazione del testo)
- PC HP Compaq
- Microfono a cuffia Hama (Mic e line) con volume regolabile

2. Hardware e software utilizzati dall'Operatore 2¹⁰:

- Macchina per Stenotipia Stenograph
- Express Scribe (Software di trascrizione)
- Pedaliera

⁹ La registrazione automatica che ha prodotto il testo trascritto è stata effettuata con il software Dragon Naturally Speaking 11.0 tramite computer HP Compaq eseguita dall'Operatore 1.

¹⁰ La registrazione automatica che ha prodotto il testo trascritto è stata effettuata con il software Dragon Naturally Speaking 11.0 tramite computer HP Compaq eseguita dall'Operatore 1.

[Torna all'indice](#)

- Winword 2007 (software di elaborazione testi)
- PC Notebook Packard Bell portatile
- Microfono a cuffia Hama (Mic e line) con volume regolabile

Nell'esperire le prove i due operatori per ogni brano hanno ritenuto di alternare i metodi utilizzati, per cui l'ordine delle metodologie adoperate è sempre stato invertito ogniqualvolta ogni nuovo brano è stato approcciato per la trascrizione. Ciò per garantire la genuinità della riproduzione del testo e per avvicinarsi il più possibile ad una situazione "reale".

Per la correzione del testo/bozza ottenuto sia con Trascrizione Automatica, sia con Stenotipia, sia con Respeaking si è ritenuto di procedere per l'integrazione utilizzando la Dattilografia e la Stenotipia, il Respeaking. Per quanto riguarda questo ultimo punto, considerando principalmente che la registrazione automatica produce un testo senza punteggiatura e che la correzione della stessa tramite dettatura non presenta vantaggi temporali rispetto alla semplice battitura di un tasto, ipotizziamo che il tempo di correzione sia decisamente più lungo rispetto ad altri sistemi, che apparentemente ci sembrano più versatili.

Per la trascrizione con Stenotipia si è proceduto con due modalità:

- In real time con produzione di bozza e quindi successiva e immediata revisione con correzione di ortografica e punteggiatura.
- In differita, ovvero con produzione del testo con immediate correzioni ortografiche e punteggiatura mediante l'uso della pedaliera.

Per la trascrizione con il Respeaking si è proceduto con due modalità:

- In real time con produzione di bozza e quindi successiva e immediata revisione e correzione di ortografica e punteggiatura;
- In differita, ovvero produzione del testo con immediate correzioni ortografiche e punteggiatura mediante l'uso delle funzioni del software Olympus.

Per la trascrizione con Dattilografia si è proceduto con un'unica modalità, ossia in differita, con produzione del testo seguita dall'immediata correzione dell'ortografica e della punteggiatura mediante l'uso delle funzioni del software Olympus. C'è da fare una breve considerazione riguardo alla produzione del testo prodotto con il programma di Trascrizione Automatica. Quando si intende utilizzare tale programma occorre avere delle accortezze: un oculato operatore prima di avviarlo si preoccupa di regolare in modo ottimale il volume di registrazione, di equalizzarlo, di tararlo e soprattutto di "ripulire" l'audio. Nel nostro caso il brano automaticamente trascritto ha avuto la "fortuna" di avere un audio ben regolato sotto tutti gli aspetti.

[Torna all'indice](#)

2. Brano Lento — Durata 5 minuti

Il Brano Lento è stato estrapolato da un consiglio comunale ben registrato dove a parlare è un assessore che ha un linguaggio chiaro, pulito e grammaticalmente buono e scandisce bene le parole. La registrazione è ottima. Il brano “produce” in media 196 sillabe/minuto ovvero 90 parole/minuto. I risultati delle prove effettuate sono espressi nella tabella 1:

Metodo	Parole totali	Parole minuto	Caratteri totali	Caratteri minuto	Sillabe totali	Sillabe minuto	Tempo M. S.	
Trascrizione automatica	447	89	3081	616				30 Errori
Revisionata Stenotipia	444	89	3108	622	974	195	06.17	
Revisionata Dattilografia	442	88	3103	621	974	195	05.58	
Revisionata Respeaking	444	89	3120	624	980	196	06.49	
Trascrizione Stenotipia real time con bozza	461	92	3162	632	995	199	10.00	
Trascrizione Respeaking real time con bozza	444	89	3114	623	980	196	10.00	
Trascrizione Dattilografia	444	89	3127	625	986	197	09.45	
Trascrizione Stenotipia con utilizzo pausa	446	89	3116	623	980	196	05.55	
Trascrizione Respeaking con utilizzo pausa	444	89	3120	624	980	196	06.15	

Tabella 1: risultati brano lento

Come si evidenzia in tabella il Brano Lento è stato ben ripreso dal programma di Trascrizione Automatica, producendo appena 30 errori, quindi una media di 6 errori o parole errate al minuto. Il testo ottenuto è stato corretto sia con la Stenotipia che con la Dattilografia e quest'ultima ha garantito miglior risultati per la correzione dal punto di vista della velocità. Infatti avendo un brano quasi perfetto, la Stenotipia, che usa il metodo sillabico per scrivere, è risultata meno versatile. La Dattilografia, con poche parole da correggere, “obbligando” le mani fisse sulla tastiera garantisce più speditezza nella correzione dell'elaborato. Per i risultati col metodo Respeaking, contrariamente a quanto si poteva ipotizzare inizialmente, la tecnica, benché macchinosa e poco agevole da usare per la correzione di un testo/bozza, non è risultata essere particolarmente lenta, visto che i tempi di correzione sono stati inferiori al minuto rispetto alle altre due tecniche, decisamente più versatili.

[Torna all'indice](#)

Il Brano lento è stato trattato con metodo Stenotipia in real time, quindi producendo una bozza preliminare che con successivo riascolto e correzione ha fornito il prodotto finale in 10 minuti. Mentre usando la Stenotipia in modalità differita, quindi con l'uso di pedaliera con correzione immediata, il prodotto finito è stato ottenuto in un tempo di minuti 5:55.

Il Brano lento è stato trattato con metodo Respeaking in real time, quindi producendo una bozza preliminare che con successivo riascolto e correzione ha fornito il prodotto finale in 10 minuti. Mentre usando il Respeaking in modalità differita, quindi con l'uso di software con backspace da tastiera e con correzione immediata, il prodotto finito è stato ottenuto in un tempo di minuti 6:15. Per cui con un siffatto brano la Stenotipia e il Respeaking si equivalgono perfettamente nella real time, mentre in modalità differita la Stenotipia garantisce il miglior risultato. Inutile dire che il metodo Dattilografia in differita risulta il meno vantaggioso.

Nella tabella 2 vengono messi a confronto a sinistra il testo trascritto con la registrazione automatica e a destra il testo precedentemente denominato "prodotto finito":

<p>Si è decifrati necessario affrontare notevoli specie obbligatorie quali fitti personale energia elettrica telefoni metano assicurazioni contratto rimborso di quote capitale su mutui che hanno assorbito l'attività amministrativa dell'ente per circa il 90% del complesso delle risorse si è dovuto altresì procedere al riconoscimento e all'approvazione di debiti fuori bilancio nonché alla stipula di transazioni si è dovuto altresì primariamente tener conto dei risultati della gestione 2000 e 10 che com'è noto si raggiunse con un rilevante disavanzo ammontante ad euro 4.630.966,53 e so come ricorderete era in buona parte scaturito dalle errate previsioni relative alle entrate Ici e più dettagliatamente con riguardo all'applicazione Ici sulle aree edificabili i cui relativi accertamenti pur notificati per una misura eccellente superiore a quelle programmate pur prudenzialmente si erano rivelati per la quasi totalità in essa d'altra parte il suo avanzo era depresso dall'attento monitoraggio effettuato in materia di riaccertamento dei siti attivi e passivi dell'ente per la prima volta era avvenuto che ogni singolo dirigente per ogni singolo settore in questo curato l'ente seguito di attenti sani e verifiche aveva</p>	<p>Si è reso infatti necessario affrontare notevoli spese obbligatorie quali fitti, personale, energia elettrica, telefoni, metano, assicurazioni, contratto, rimborso di quote capitale su mutui, che hanno assorbito l'attività amministrativa dell'ente per circa il 90% del complesso delle risorse. Si è dovuto altresì procedere al riconoscimento e all'approvazione di debiti fuori bilancio, nonché alla stipula di transazioni. Si è dovuto altresì primariamente tener conto dei risultati della gestione 2010 che, com'è noto si raggiunse con un rilevante disavanzo ammontante ad euro 4.630.966,53. Esso, come ricorderete, era in buona parte scaturito dalle errate previsioni relative alle entrate ICI e più dettagliatamente con riguardo all'applicazione ICI sulle aree edificabili i cui relativi accertamenti, pur notificati per una misura decisamente superiore a quelle programmate pur prudenzialmente, si erano rivelati, per la quasi totalità, inesatte. Per altra parte il di avanzo era dipeso dall'attento monitoraggio effettuato in materia di riaccertamento dei residui attivi e passivi dell'ente. Per la prima volta era avvenuto che ogni singolo dirigente, per ogni singolo settore in cui è strutturato l'ente, a seguito di attenti esami e verifiche,</p>
--	--

prodotto apposta determina di riaccertamento dei dissidi versamento profondità in ossequio anche alle raccomandazioni più volte espresse dalla corte dei conti sezione di controllo che ha portato ad eliminare somme avessimo che producevano solo un appesantimento infruttuoso dei conti dell'ente l'affermazione forte avvenute in sede di approvazione del bilancio previsionale 2000 e 11 ribadita come dicevo nella parte iniziale dell'intervento in facce di riequilibrio è confermata in occasione dell'asestamento l'affermazione di conti veritieri oggi ci trova piena concretizzazione e legittimazione il conto consuntivo del 1011 si chiude infatti con un avanzo di euro 69.790,33 per quanto riguarda la cessione di competenza un risultato estremamente positivo se solo si guarda agli illeciti dell'anno passato quando il disavanzo di competenza era montato in euro 2.000.985 mila 103,32 il risultato è stato possibile innanzitutto per il ragguardevole lavoro svolto dall'ufficio tributi che in ossequio all'impegno preso è stato riorganizzato rafforzato nelle risorse umane interna rizzato in gran parte dei servizi e trasferito anche di sede l'apposita istituzione dell'assessorato alle politiche delle entrate ha ulteriormente potenziato anche l'attività di indirizzo e di controllo i primi risultati di questa attività sono veramente incoraggianti tra l'altro e quindi tra le molteplici attività svolte e i risultati conseguiti sono stati individuati formati e notificati 1914 abissi di accertamento ha che da soli Ici con riferimento alle aree edificabili e ad altre unità immobiliari l'ammontare degli accertamenti frutto di queste intense qualificate attività è stata di euro € 3.312.176 peraltro una conferma sia pure ex supposta che quanto era stato previsto nel bilancio 2000 e 10 nella materia non era frutto di fantasia ma di previsione ponderata

aveva prodotto apposita determina di riaccertamento dei residui. Un esame in profondità in ossequio anche alle raccomandazioni più volte espresse dalla Corte dei Conti sezione di controllo, che ha portato ad eliminare somme a residuo che producevano solo un appesantimento infruttuoso dei conti dell'ente. L'affermazione forte avvenuta in sede di approvazione del bilancio previsionale 2011, ribadita – come dicevo nella parte iniziale del mio intervento – in fase di riequilibrio è confermata in occasione dell'asestamento all'affermazione di conti veritieri, oggi ci trova piena concretizzazione e legittimazione. Il conto consuntivo 2011 si chiude infatti con un avanzo di euro 69.790,33 per quanto riguarda la gestione di competenza. Un risultato estremamente positivo se solo si guarda agli esiti dell'anno passati quando il disavanzo di competenza era ammontato in euro 2.985.103,32. Il risultato è stato possibile innanzitutto per il ragguardevole lavoro svolto dall'Ufficio Tributi che in ossequio all'impegno preso è stato riorganizzato, rafforzato nelle risorse umane, internalizzato in gran parte dei servizi e trasferito anche di sede. L'apposita istituzione di un assessorato alle politiche delle entrate ha ulteriormente potenziato anche l'attività di indirizzo e di controllo. I primi risultati di questa attività sono veramente incoraggianti. Fra l'altro, e quindi tra le molteplici attività svolte e i risultati conseguiti, sono stati individuati, formati e notificati 1914 avvisi di accertamento a evasori ICI con riferimento alle aree edificabili e ad altre unità immobiliari. L'ammontare degli accertamenti, frutto di queste intense qualificate attività, è stata di euro 3.312.076,00. Peraltro una conferma sia pure ex post che quanto era stato previsto nel bilancio 2010 nella materia non era frutto di fantasia ma di previsione ponderata.

Tabella 2: testo trascritto con registrazione automatica e trascrizione brano lento

3. Brano Medio — Durata 5 minuti

Il Brano Medio è stato estrapolato da un consiglio provinciale dall'intervento di un assessore, la registrazione risulta non ottimale e quasi sufficiente come qualità, la voce dell'interlocutore è chiara e le parole sono ben

[Torna all'indice](#)

scandite. Il brano “produce” in media 238 sillabe/minuto ovvero 113 parole/minuto. I risultati delle prove effettuate sono espressi nella tabella 3:

Metodo	Parole totali	Parole minuto	Caratteri totali	Caratteri minuto	Sillabe totali	Sillabe minuto	Tempo M. S.	80 Errori
Trascrizione Automatica	548	110	3603	721				
Revisionata Stenotipia	564	113	3700	740	1185	237	08.02	
Revisionata Dattilografia	568	114	3736	747	1195	239	08.58	
Revisionata Respeaking	573	115	3728	746	1195	239	09.42	
Trascrizione Stenotipia real time con bozza	571	114	3721	744	1186	237	11.00	
Trascrizione Respeaking real time con bozza	570	114	3736	745	1194	239	11.30	
Trascrizione Dattilografia	572	114	3736	745	1186	237	12.38	
Trascrizione Stenotipia con utilizzo pausa	569	114	3715	743	1187	237	05.42	
Trascrizione Respeaking con utilizzo pausa	573	115	3728	746	1195	239	07.33	

Tabella 3: risultati brano medio

Come si evidenzia in tabella il Brano Medio non è stato ben ripreso dal programma di Trascrizione Automatica, infatti in questo caso gli errori prodotti sono risultati 80, quindi una media di 16 errori o parole errate al minuto, probabilmente a causa sia dell’audio non ottimale, seppur sufficiente, e sia per la velocità del parlato dell’oratore.

Il testo ottenuto è stato corretto sia con la Dattilografia sia con la Stenotipia e quest’ultima ha garantito miglior risultati per la correzione dal punto di vista della velocità. Questo perché la Stenotipia ha garantito più rapidità nella correzione degli errori in virtù del sistema sillabico che in questo caso ha risparmiato battute rispetto alla Dattilografia. Quindi più aumentano gli errori da correggere più è efficace il metodo sillabico.

Per i risultati col metodo Respeaking, risulta ancora più evidente rispetto a prima la difficoltà nell'uso. In questo caso, inoltre, l'operazione ha richiesto molto più tempo rispetto alle altre due tecniche. Chiaramente, il numero di operazioni da compiere aumenta esponenzialmente i tempi di correzione e

[Torna all'indice](#)

quello che sembrava un leggero ritardo nel test sul Brano Lento era in realtà più da imputare all'esiguità degli errori che alla snellezza della tecnica.

Il testo trattato con Stenotipia e Respeaking in real time con successiva correzione della bozza, ha fornito risultati quasi equivalenti ottenendo uno scarto di appena 30 secondi in più a vantaggio della Stenotipia. Come è evidente in tabella il Brano Medio trattato in differita è stato meglio elaborato con metodo Stenotipia in quanto il prodotto finito è stato ottenuto quasi in tempo reale, infatti con uno scarto temporale di soli 42 secondi.

Per quanto riguarda il Respeaking in modalità real time con produzione di bozza va precisato che più è veloce la sequenza sillabale più vi è possibilità di errore nella pronuncia del respeaker. Mentre in modalità differita la perdita di tempo è causata dal dover stoppare più volte il brano per memorizzare ciò che si ascolta e quindi ottenere una miglior dettatura. Inutile dire che il metodo Dattilografia in differita risulta il meno vantaggioso, che in riproduzione testo, in questo caso, ha un rapporto di 1:2,5; quindi quasi due volte e mezzo la durata della registrazione.

Nella tabella 4 vengono messi a confronto a sinistra il testo trascritto con la registrazione automatica e a destra il testo precedentemente denominato "prodotto finito"¹¹:

<p>Si (...) è in pretrattamento decisamente delicato perché siamo nell'ambito dei prodotti finanziari in derivati (..) quindi sostanzialmente entriamo in una materia di un tecnicismo esasperato e pertanto io personalmente conoscendo un po' i prodotti ho ritenuto opportuno che fosse presente in quest'incredibile il tecnico che aveva già sconvolta precedentemente collaborato con l'amministrazione al fine di possa essere più chiaro di metter delucidare tutti consiglieri sulle operazioni che andiamo diciamo rettili proponiamo di portare io in sintesi cercherò con parole meno tecniche di spiegare diciamo l'illeciti prodotti derivati da dove sono a tasso variabile sono fortemente collegati contro l'andamento dei mercati come dicevamo nella precedente relazione all'altezza momento di una grande instabilità internazionale della grande instabilità dei mercati finanziari di posso assicurare che mi costa personalmente che sul mercato internazionale non ci</p>	<p>Si! Questo è un argomento decisamente delicato perché siamo nell'ambito dei prodotti finanziari derivati e quindi sostanzialmente entriamo in una materia di un tecnicismo esasperato e pertanto io personalmente, pur conoscendo un po' i prodotti, ho ritenuto opportuno che fosse presente in questa assise il tecnico che aveva già a sua volta precedentemente collaborato con l'Amministrazione, al fine che possa essere più chiaro di me per delucidare tutti i Consiglieri sulle operazioni che andiamo diciamo... che vi proponiamo di votare. Io in sintesi cercherò con parole meno tecniche di spiegare diciamo la situazione. I prodotti derivati, laddove sono a tasso variabile, sono fortemente collegati con l'andamento dei mercati. Come dicevamo nella precedente relazione oggi c'è un momento di grande instabilità nazionale, c'è una grande instabilità dei mercati finanziari, vi posso assicurare che mi consta personalmente che sul</p>
--	---

¹¹ Il testo della trascrizione automatica risulta essere più corto perché molte parole non sono state riconosciute dal software.

sono soldi segue andata ad (..) un'operazione anche privatistica oggi addirittura le banche oltre gli interessi richiedono una una tantum se dovete farne una una tantum per avere l'eventuale mutuo se è una restrizione notevole quindi sicuramente in questo momento i tassi di nell'ambito di questa forbice tra il tetto minimo (..) il tetto massimo non possono che andare decisamente persona quindi con una penalizzazione perché (..) il prodotto derivato sicuramente a posteriori e facile giudicare dire sostanzialmente dal fatto che l'abbia fatto male perché (..) come quando anche i privati li hanno investito in borsa laddove andavano bene 1° quando Pol Pot andati male la colpa era di tutti il problema è che purtroppo può così come Giovanni Agnelli nell'ambito della sua anche relazione vi si si può sapere cosa succederà domani (...) perché le condizioni di variabilità sono innumerevoli vanno dal sociale alla guerra all'apertura del petrolio alla questione che domani mattina sostanzialmente basta che un immigrato arabo se la prenda con qualcuno e quindi fa in modo che il prezzo del petrolio vada oltremisura cespugli dire sono tantissime le condizioni la corte accerta che questo è il momento peggiore di andare (..) operare sui derivati e quindi ad andare ad mantenere quel prodotto che noi avevamo e quindi in Italia abbiamo chiesto immediatamente come giunta con il presente di inizi che ci fosse un intervento ponte che riducesse l'esposizione del momento dell'ente verso questa variabilità con l'attività dei tassi e ci consentisse anche di recuperare tempo al fine di valutare quale possibilità reale e su questo vi assicuro che personalmente cercherò di trovare i tipi come disimpegnarmi al massimo per trovare una soluzione alternative (ai) derivati perché considerazione personale (..) quindi non è assolutamente un giudizio per altro e sento che faccio una protezione anche economica e quindi alcune posizioni di ogni professionista sono parti conosciuti e io sono sempre stato per la certezza dei costi sia nell'ambito delle famiglie (..) addirittura anche nell'ambito dell'ente quindi un'eventuale esposizione dell'ente su la incertezza di quelle che potrebbero essere i costi di rateazione francamente non mi vede

mercato internazionale non ci sono soldi, se voi andate a fare anche un'operazione privatistica, oggi addirittura le banche oltre l'interesse vi chiedono una una tantum, cioè dovete pagare una una tantum per avere l'eventuale mutuo. C'è una restrizione notevole. Quindi sicuramente in questo momento i tassi nell'ambito di questa forbice tra il tetto minimo e il tetto massimo, non possono che andare decisamente verso l'alto. Quindi con una penalizzazione, perché sai, il prodotto derivato, cioè sicuramente a posteriori è facile giudicare e dire sostanzialmente "abbiamo fatto bene", "abbiamo fatto male", questo è come quando anche i privati hanno investiti in Borsa, laddove andavano bene, erano bravi, quando poi le cose sono andate male la colpa era di tutti. Il problema è che purtroppo, così come Giovanni Agnelli nell'ambito della sua più ampia relazione disse: chi può sapere cosa succederà domani in Borsa? Perché le condizioni di variabilità sono innumerevoli, vanno dal sociale, alla guerra, alla questione del petrolio, alla questione che domani mattina sostanzialmente basta che un emirato arabo se la prenda con qualcuno e quindi fa in modo che il prezzo del petrolio vada oltre misura, cioè, voglio dire, sono tantissime le condizioni. Una cosa è certa: che questo è il momento peggiore di andare ad operare sui derivati e quindi ad andare a mantenere quel prodotto che noi avevamo e quindi abbiamo chiesto immediatamente, come Giunta, con Presidente De Nisi, che ci fosse un intervento ponte che riducesse l'esposizione del momento dell'ente verso questa variabilità e volatilità dei tassi e ci consentisse anche di recuperare tempo al fine di valutare quale possibilità reale, e su questo vi assicuro che personalmente cercherò di trovare e di, come dire, di impegnarmi al massimo per trovare una soluzione alternativa ai derivati, perché, considerazione personale e quindi non è assolutamente un giudizio, per altro essendo che faccio una professione anche economica e quindi alcune posizioni di ogni professionista sono quasi conosciuti, io sono sempre stato per la certezza dei costi, sia nell'ambito delle famiglie e addirittura anche nell'ambito delle imprese. Quindi un'eventuale esposizione dell'ente su una incertezza di quelle che

felice pertanto egli dicevamo con il presidente di studiare ed è questa la proposta che viviamo una intervento ponte e ci dia la possibilità di far 31 12 di poter verificare come anche eventualmente uscire completamente e derivati per l'applicazione più tecnica per sé e diciamo è necessario come ritengo che se lo desiderate verrà e il dottore E (...) sicuramente spiegarle sotto un profilo tecnico meglio di me l'operazione che vi stiamo proponendo	potrebbero essere i costi di rateazione, francamente, non mi vede felice. Pertanto e, dicevamo con il Presidente, di studiare – ed è questa la proposta che vi diamo – un intervento ponte che ci dia la possibilità di qua al 31/12 di poter verificare come anche eventualmente uscire completamente dai derivati. Per la spiegazione più tecnica se diciamo è necessario, come ritengo che se lo desiderate avverrà, c'è il dottore che può sicuramente spiegare sotto un profilo tecnico meglio di me l'operazione che vi stiamo proponendo.
---	--

Tabella 4: testo trascritto con registrazione automatica e trascrizione brano medio

4. Brano Veloce — Durata 5 minuti

Il Brano Veloce è stato estrapolato da un consiglio comunale discretamente registrato dove il consigliere si esprime con voce chiara ma alterata in quanto lo stesso è emotivamente coinvolto dall'argomento, ciò rende il brano incostante riguardo la tonalità e la velocità che l'interlocutore usa, mentre le parole sono ben scandite. Il brano "produce" in media 260 sillabe/minuto ovvero 134 parole/minuto. I risultati delle prove effettuate sono espressi nella tabella 5:

Metodo	Parole totali	Parole minuto	Caratteri Totali	Caratteri minuto	Sillabe totali	Sillabe minuto	Tempo M. S.	
Trascrizione Automatica	435	87	2613	523	827			Oltre 400 Errori
Revisionata Stenotipia	668	134	4155	831	1303	261	10.05	
Revisionata Dattilografia	665	133	4140	828	1300	260	17.22	
Revisionata Respeaking	660	132	4122	824	1303	261	12.30	
Trascrizione Stenotipia real time con bozza	674	135	4148	830	1301	260	12.30	
Trascrizione Respeaking real time con bozza	660	132	4122	824	1303	261	13.10	
Trascrizione Dattilografia	670	134	4144	829	1304	261	14.45	
Trascrizione Stenotipia con utilizzo pausa	672	134	4135	827	1296	259	07.00	
Trascrizione Respeaking con utilizzo pausa	668	134	4137	835	1303	261	09.15	

Tabella 5: risultati brano veloce

[Torna all'indice](#)

Come si evidenzia in tabella il Brano Veloce è stato pessimamente ripreso dal programma di Trascrizione Automatica: qui abbiamo oltre 400 errori, dove per “errore” in questo caso abbiamo dovuto intendere sia parole non prodotte sia parole prodotte ma erroneamente trascritte. Questo risultato induce qualsiasi operatore a riscrivere totalmente il brano, in quanto qualsiasi metodo porterebbe a un lavoro lungo e faticoso per non dire snervante. Tuttavia abbiamo voluto trattare il brano come i precedenti e di questo abbiamo fornito i risultati.

Il testo ottenuto è stato corretto sia con la Dattilografia sia con la Stenotipia sia con il Respeaking. La Stenotipia ha garantito miglior risultati per la correzione dal punto di vista della velocità per i motivi già detti per il Brano Medio. Il testo trattato con Stenotipia e Respeaking in real time con successiva correzione della bozza, anche in questo caso, ha fornito risultati quasi equivalenti ottenendo uno scarto di 40 secondi in più a vantaggio della Stenotipia.

Come è evidente in tabella il Brano Veloce trattato in differita è stato meglio elaborato con metodo Stenotipia e per questo ci riportiamo alle considerazioni già espresse per il Brano Medio, evidenziando il fatto che più aumenta la velocità del parlato più è efficace il metodo sillabico della Stenotipia. Ovviamente risulta meno conveniente la Dattilografia in differita, che in riproduzione testo in questo caso ha un rapporto di 1:3, ovvero quasi tre volte la durata della registrazione.

Nella tabella 6 vengono messi a confronto a sinistra il testo trascritto con la registrazione automatica e a destra il testo precedentemente denominato “prodotto finito”:

<p>E già da tre cose all'interno dell'area revisori all'ione debiti fuori bilancio dell'uno e dell' ente che ha trovato traccia e non mi va di casa la somma perché secondo me anche quello per la serenità sarà un'ulteriore accumulo di debito o di dell'importo grazie anche dare il via l'oste dei ad ognuno di somme che posso capire l'perché lo era arrivato ed è una parte è stato utilizzato nel 1000 6 milioni ne è l'noi formalmente ora presenteremo anche dall'organo di ione e dell'organo fa niente vorrei che mi deve dire che l'organo grazie lei rivolgendole con il gruppo del partito democratico formalmente rappresenta una di chi non è opportuno farla mangiare anche accadere anche dei diritti dell'organicità vogliamo anche informalmente risposta sulla verifica della carta soprattutto i limiti col vincolo della mezzanotte è per capire la portata la e</p>	<p>Leggevo tra le altre cose, all'interno della relazione dei revisori che fanno menzione di due debiti fuori bilancio, del Settore 1 e del Settore 7, e non ho trovato traccia, e vorrei che mi venisse quantificata la somma, perché secondo me anche quello per non aggravare le condizioni di cassa sarà un ulteriore accumulo di debito. Mi può dire quant'è l'importo? Grazie. Ha certificato anche, la relazione e gli atti lo dicono lo stesso, che ad oggi le opere Pisu, al 2011, nessuna somma è stata impegnata. Posso capirlo perché l'anticipo che era arrivato del 20% buona parte è stato utilizzato per l'anticipo di Tesoreria. Due milioni e mezzo. Bene, sulla base di questo, noi formalmente ora presenteremo una richiesta all'organo di revisione, che tra i tanti adempimenti che lo stesso organo fa – Presidente vorrei che mi sentisse, Presidente dell'organo, grazie, a</p>
--	---

soprattutto di ad oggi delle e delle virtù degli accertamenti è anche vero che ad oggi siamo qui è un elemento che non saremo da qui a poche tutte le Usi non era delimitata a il fatto rimane prologo elettronico dell'ente e la mancanza di liquidità riguarda le ore della dei trasferimenti erariali e gradevoli perdiamo in un quadro normativo in continua e un'non è quindi oggi non vale quel detto ieri Domenico da chi non bada quel detto oggi in ogni caso per certificare le abbiamo del diamo analitica di vetro quando ad oggi la vera lealtà dei conti vincolo soprattutto in Italia netto forma è ribadisco di generare concludo il tempo limitato nella cucina perché vorrei che lei mi ha la voglia problema secondo cui il terreno dell'affidamento dell'incarico per la riscossione volontaria coattiva all'Italia è vero che mi abbiamo avuto difficoltà quando alcuni dei gregari nominavano di nominar dirigente segretario generale erano con loro c'era ancora oggi è generale il Comune come li pone o comunque i comuni non solo nell'eventualità di ricoprire tale eventualmente ne contempla e deve spiegare come la organizzativa all'interno del settore non può svolgere quella dei emotivo e generale Rai corte le attività di dire da controllore controllato ma così forse non è che mi ha fatto leggere di determinati atti ed è per abitudine convince coscienti che tutti quanti o buona parte di noi non gli vennero le è una grande offesa e quotidianamente legge le carte sono forse prediletti di arrivato il momento Mario lo dica per mettere tutti gli dei dell'ordine sociale che lasciar non aggiungiamo saranno quelli del no dicevo il gamma di bordo diceva Giovanni e fatti concreti fatti concreti sono negli istituti prego perché con attenzione è certo il reparto asco

lei mi sto rivolgendo, perché come gruppo del Partito Democratico formalmente ora presentiamo una richiesta –, perché secondo me è opportuno fare chiarezza, noi siamo pronti a formalizzare una richiesta, considerando il fatto che tra i tanti adempimenti che l'organo di revisione fa, vogliamo la certezza e quindi formalmente una risposta, sulla verifica della cassa, soprattutto in riferimento ai conti vincolo ed eventuale allineamento dello stesso. Questo a noi serve per capire il modus operandi portato avanti e capire soprattutto quali sono le condizioni reali ad oggi dell'ente. Perché è vero che tutto si basa su degli accertamenti, è anche vero che ad oggi la riscossione che dicevamo prima è un elemento che non dà certezze, è anche vero che ci troveremo da qui a poco a discutere su aliquota o meno dell'Imu e quant'altro, ma il fatto e il male cronico dell'ente è la mancanza di liquidità che riguarda tutti gli enti, vuoi per la riduzione dei trasferimenti erariali che tra le altre cose e ci troviamo in un quadro normativo in continua evoluzione e quindi oggi non vale quello che si è detto ieri e domani può darsi non valga quello che si è detto oggi, quindi in ogni caso per certificare questo noi abbiamo bisogno di questa risposta. Non chiediamo la verifica trimestrale per non stare troppo addosso, ma ad oggi la vera realtà sui conti vincolo soprattutto e quindi l'eventuale allineamento formalmente, lo ribadisco, l'organo di revisione ce ne deve fare copia. Concludo, visto il tempo... è illimitato? Benissimo, e allora continuo perché vorrei che mi venisse data risposta sul problema se secondo voi esiste o meno dell'affidamento dell'incarico per la riscossione volontaria e coattiva all'agenzia Equitalia e Europa 2000. Abbiamo avuto difficoltà anche quando alcuni decreti sindacali nominavano dirigenti o nominava dirigente il segretario generale, eravamo convinti allora e lo siamo ancora oggi che un segretario generale in un comune come Vibo Valentia o comunque nei comuni, soltanto nell'eventualità di mancanza di dirigenti può ricoprire tale ruolo, eventualmente non è contemplato. Perché se no mi si deve spiegare come mai una posizione organizzativa all'interno dello stesso settore non può svolgere quello che deve svolgere il segretario,

	<p>per il semplice motivo che il segretario generale partecipa e coordina l'attività dei dirigenti, in quel momento diventa controllore e controllato. Una certa, chiamiamola così, e forse non è il termine esatto, leggerezza, nel gestire determinati atti, non deve passare per abitudine convinti e coscienti che tutti quanti o buona parte di noi gli atti non li vede e non li legge. È una grande offesa, è una grande mancanza di rispetto a chi quotidianamente legge le carte, anche e solo forse per divertirsi. Io penso che è arrivato il momento, caro Mario, lo dico a te perché generalmente il referente tu lo utilizzi, di finirla col dire "questo ci avete lasciato", perché se noi aggiungiamo, chissà che troveranno quelli che verranno dopo di voi, per il semplice motivo che ora ci troviamo al giro di boa, come diceva Giovanni prima, e fatti concreti, fatti concreti ci sono pure, poi me li dici tu, ti prego, perché io con attenzione e in silenzio certolino come ho sempre fatto ascolto.</p>
--	---

Tabella 6: testo trascritto con registrazione automatica e trascrizione brano veloce

5. Risultati

Per quanto riguarda la Trascrizione Automatica, in condizioni ottimali offre all'operatore con buona conoscenza audio-informatica due vantaggi: fare il grosso del lavoro al suo posto dandogli la possibilità di riposarsi o lavorare altrove contemporaneamente per tutto il tempo che la macchina lo sostituisce. Comincia a risultare svantaggiosa quando il tempo necessario per correggere gli inevitabili errori supera quello guadagnato nell'automazione. Nei casi simili a quelli del Brano Veloce risulta essere del tutto inutile.

In merito alla Dattilografia, essa è fondamentale e parte integrante delle altre tecniche analizzate in quanto le completa e le coadiuva. L'intervento sulla tastiera sarà sempre necessario e inevitabile (es.: salvataggio di un file, inserimento punteggiatura, l'uso di Internet, ecc.).

Veniamo ora alle due tecniche di scrittura veloce più facilmente paragonabili tra di loro, Stenotipia e Respeaking. La Stenotipia è un metodo di ripresa del parlato molto efficace ma solo se chi la pratica sa padroneggiarne al meglio la tecnica. Imparare la Stenotipia richiede tempo, esercizio e costante pratica. Il periodo di formazione per diventare operatore stenotipista dura all'incirca dai due ai tre anni: da sei a nove mesi di formazione per apprendere le competenze di base; un anno-un anno e mezzo per diventare esperti in tempo reale nella ripresa del linguaggio parlato. Uno stenotipista necessita di

[Torna all'indice](#)

aggiornamenti continui sui software da utilizzare e periodica revisione sulle proprie attrezzature, le quali devono sempre funzionare al meglio, e deve costantemente tenersi in allenamento per garantire un'ottima performance. Uno stenotipista, per quanto bravo a saper riportare il parlato velocemente, correttamente e senza errori, dovrà sempre revisionare, attraverso un'attenta rilettura, il testo prodotto. Lo stenotipista lavora nel proprio "silenzio" e in autonomia per molte ore riuscendo anche ad astrarsi dal concettualizzare ciò che sente grazie al meccanismo dell'automatismo.

Il Respeaking è anch'esso molto efficace, se può contare su una consolidata professionalità tecnica. Infatti imparare la tecnica del Respeaking implica non solo saper ascoltare, capire e parlare allo stesso tempo, ma anche avere una buona pronuncia, un eloquio fluente e soprattutto la padronanza della propria voce, che deve essere ben "impostata" e possibilmente priva delle inflessioni dialettali più impattanti. Altra competenza è la capacità dell'operatore di gestire e ottimizzare la propria respirazione. Saper respirare bene è infatti fondamentale in quanto la "mancanza di ossigeno" potrebbe affaticare e compromettere tutto il lavoro.

Queste competenze si acquisiscono non solo con la pratica e l'esercizio costante, ma anche con la buona conoscenza delle funzionalità del software che si utilizza. Per arrivare a una buona padronanza della tecnica del Respeaking non c'è un periodo stabilito in quanto l'apprendimento delle competenze di base è un processo del tutto soggettivo. Chi approccia questa tecnica ed è intenzionato a praticarla, se è ben (pre)disposto, saprà utilizzarla bene anche dopo le prime settimane. Si potrà così godere del vantaggio di servirsi di pochi mezzi per arrivare al prodotto finito, ovvero il proprio computer e la propria voce. Tuttavia, da ricordare è che il respeaker, così come lo stenotipista, riesce a produrre in tempo reale un testo leggibile e completo, ma dovrà necessariamente revisionare il testo prodotto attraverso un'attenta rilettura che sia in grado di rilevare perfino gli errori più puntuali, se si vuole giungere a un risultato soddisfacente al 100%.

6. Conclusioni

In conclusione abbiamo ritenuto di effettuare queste prove per dimostrare che non esiste un metodo infallibile e preciso al 100% per la riproduzione fedele di un testo e che sta quindi all'operatore, risorsa umana insostituibile, saper valutare e di conseguenza scegliere (o alternare) il sistema migliore tra quelli sopra considerati per la trascrizione di un brano. Il professionista approccerà la giusta metodologia in base alla situazione che ha di fronte, valutando quindi l'ambiente di lavoro, l'attrezzatura di cui dispone, il tipo di parlato che si dovrà trascrivere e, perché no, le proprie condizioni psicofisiche, in quanto è evidente che una migliore performance la si ottiene nelle migliori condizioni di salute e intellettuali di chi opera.

[Torna all'indice](#)

Non risulta conveniente a un professionista del settore della “verbalizzazione” utilizzare una singola tecnica, in quanto limiterebbe i propri campi di azione a suo svantaggio. Nelle prove abbiamo approcciato ogni brano singolarmente e con singole metodologie, ma in situazioni “reali” è evidente che chi sa padroneggiare più di un metodo di scrittura veloce più riuscirà a intervenire contemporaneamente e con tutti i metodi su di un testo a seconda delle situazioni che si possono manifestare come quelle sopra esaminate nei test.

Nel nostro ambito il professionista “modello” sarebbe quello che sa dattilografare, stenotipare, rispikerare e adoperare tecnologie in generale. Ovviamente un individuo competente in tutti i metodi difficilmente riuscirà a eccellere in tutti allo stesso livello e probabilmente ne preferirà sempre uno rispetto agli altri anche per il tipo di formazione ricevuta. Tuttavia saper utilizzare più metodi mette a disposizione dell’operatore la possibilità di non escludere la trascrizione di qualsiasi tipo di evento. La professione del verbalizzatore, così come quella di altri ambiti affini (resocontazione assembleare e consiliare, sottotitolazione di programmi televisivi e non, perizia fonica forense ecc.), non si può improvvisare. Si acquisisce col tempo e con l’esperienza avendo la formazione giusta basata sulla conoscenza dei metodi di trascrizione che lo stesso deve continuamente aggiornare e praticare.

Questa breve relazione, più che paragonare tra loro le varie tecniche di scrittura, aveva come scopo quello di mostrare gli aspetti negativi e positivi di ogni singola tecnica, tramite una sperimentazione che fosse il più verosimile possibile, per quanto limitata a degli estratti e non estesa su una intera giornata. Ci auguriamo che essa possa stimolare il dibattito su queste due tecniche ancora poco conosciute non solo da parte della maggioranza della cittadinanza, ma neanche da parte dei professionisti del settore.

12) SOTTOTITOLI DI CONFERENZA INTRALINGUISTICI E INTERLINGUISTICI – STRATEGIE TRADUTTIVE A CONFRONTO

Silvia Bellucci

1. Introduzione

Questo mio contributo vuole fornire uno spunto di riflessione sulla tecnica del *respeaking*, andando a evidenziare le differenze tra *respeaking* intralinguistico e interlinguistico e mettendo quindi a confronto lo stesso con l'interpretazione simultanea. Nello specifico intendo presentare un esperimento svolto in ambito universitario, che è stato oggetto di tesi di laurea. Lo scopo dell'esperimento era dimostrare la fattibilità o meno del *respeaking* utilizzato congiuntamente all'interpretazione simultanea e analizzarne i risultati ottenuti. Il primo elemento che è stato necessario individuare è stata la metodologia con la quale i risultati finali sarebbero stati analizzati. Pertanto si è scelto di procedere con l'analisi strategica del discorso rispeakerato e, quindi, attraverso la tassonomia proposta da Gambier (2005) e adattata da Eugeni alla sottotitolazione live (cfr. Eugeni 2008).

2. Metodologia

Gambier è noto come l'ideatore dell'analisi strategica fondata su tre macro-strategie: *réduction*, *simplification* e *expansion*. La strategia della *réduction* è la più utilizzata, in quanto consente di rispettare i limiti spazio-temporali che caratterizzano questa professione operando attraverso una riduzione quantitativa del testo originale, che permette così di seguire con maggiore facilità la velocità di eloquio e di eliminare tutti quei tratti propri dell'oralità, che risultano essere inutili ai fini della comprensione del testo stesso. La strategia della *simplification* viene invece applicata a livello sintattico al fine di favorire la leggibilità dei sottotitoli. L'ultima strategia, l'*expansion*, viene utilizzata per spiegare un concetto che altrimenti risulterebbe di difficile comprensione per l'utente finale. Partendo da queste tre macro-strategie, Eugeni è intervenuto apportando modifiche e facendo in modo che tale tassonomia potesse adattarsi all'analisi di un qualsiasi testo affrontato in *respeaking*. Vediamo pertanto che la prima distinzione proposta da Eugeni riguarda la segmentazione del testo in macro-unità concettuali (ogni proposizione di senso compiuto) e micro-unità concettuali (tutte le componenti dotate di significato presenti in una frase). Una volta individuati tali elementi è quindi necessario entrare più nello specifico e individuare omissioni di macro-unità, ripetizioni di macro-unità e alterazioni, siano esse espansioni o riduzioni delle macro-unità del testo di arrivo rispetto al testo di partenza (vedi Figura 1). Inoltre è anche stata aggiunta una sezione riguardante gli errori, possibili e frequenti nella sottotitolazione in tempo reale se paragonata alla sottotitolazione preregistrata.

MACRO-UNITA' CONCETTUALI					
Macro-unità non rese		Macro-unità rese			
Omissioni	Errori	Ripetizioni	Alterazioni		
		Espansioni	Riduzioni		Errori
			Compressione di micro-unità concettuali	Omissione di micro-unità concettuali	

Figura 1 – Tabella riassuntiva della tassonomia proposta da Eugeni.

3. L'esperimento

Terminata questa prima fase di individuazione della metodologia, si è potuto procedere alla scelta del materiale e dei soggetti sottoposti all'esperimento, il quale si è basato su una *mock conference* svoltasi in ambito universitario, con gli strumenti generalmente utilizzati per svolgere esercitazioni in simultanea; la postazione del *respeaker*, inoltre, è stata dotata di un computer nel quale era stato installato un software di riconoscimento del parlato, Dragon NaturallySpeaking, uno dei software *speaker-dependent* maggiormente utilizzati nell'ambito del riconoscimento del parlato in genere. Per quanto concerne il testo, si è scelto di sottoporne uno – appositamente scelto dopo attenta analisi – a dei professionisti nel campo dell'interpretazione simultanea e del *respeaking*, entrambi francesisti, al fine di riprodurre nel modo più corretto possibile la situazione lavorativa reale e di ottenere un prodotto finale di qualità. La lettura del testo in lingua originale è stata affidata a una professoressa madrelingua, anch'essa interprete. Il testo scelto, particolarmente lungo al fine di permetterne la suddivisione in due parti, è la trascrizione di un discorso in lingua francese pronunciato da Marie-Thérèse L'Huillier¹² durante un intervento nell'ambito di una conferenza svoltasi in Francia in cui si parla dell'analfabetismo nelle persone sorde. La suddivisione del testo in due parti ha permesso di gestire la prima parte in simultanea e in *respeaking* intralinguistico e la seconda parte solamente in *respeaking* interlinguistico; questo perché l'obiettivo finale dell'esperimento era mettere a confronto i risultati ottenuti con l'intervento dell'interprete e del *respeaker* contemporaneamente e con il solo intervento del *respeaker*, al fine di comprendere quali sono le possibili interazioni tra *respeaking* e simultanea.

¹² **Marie-Thérèse L'Huillier** è una ex produttrice televisiva francese sorda: è nota per aver presentato il programma televisivo “*Mes mains ont la parole*” in lingua dei segni francese nel 1979, come anche il programma “*L'œil et la Main*” per un pubblico di sordi e ipoudenti sull'emittente televisiva *France 5* tra il 1994 e il 2003. Dal 2010 è ingegnere presso il CNRS (*Centre national de la recherche scientifique*).

[Torna all'indice](#)

Dal punto di vista prettamente tecnico, i tre professionisti si sono sistemati su tre postazioni diverse. È stata creata una sorta di catena *ad hoc* in cui, la professionista che leggeva il testo di partenza era direttamente collegata (attraverso cuffie e software) alla simultaneista che ha tradotto il testo in italiano, mentre il *respeaker* ha lavorato in una postazione indipendente ricevendo il testo tradotto in italiano e dettandolo al software di riconoscimento del parlato per generare la trascrizione.

4. Risultati della prima parte dell'esperimento – la sottotitolazione intralinguistica

Al fine di rendere l'esperimento quanto più corretto possibile, si è ritenuto opportuno fare un breve raffronto sulle *performance* dei professionisti, in quanto vi sono stati diversi fattori che hanno influito sul risultato finale della ricerca. Innanzi tutto il *respeaker* si è trovato a lavorare in una postazione di sottotitolazione a lui non familiare e ha dovuto interfacciarsi con un software che, non essendo stato sufficientemente addestrato, non rispondeva in modo appropriato agli input vocali da lui dettati. Pertanto i problemi riscontrati nell'interazione con la macchina si sono tradotti in un elevato numero di errori imputabili al software. Quanto all'interprete invece, il fatto di conoscere le caratteristiche di eloquio della lettrice madrelingua (le due professioniste lavorano insieme) ha contribuito in modo molto positivo alla resa. L'ultimo elemento riguarda il testo: il fatto di averne scelto uno particolarmente lungo – la durata della prima parte è stata di circa 25 minuti – non ha aiutato né l'interprete né il *respeaker*, dal momento che entrambi hanno lavorato ininterrottamente e sono stati sottoposti a un affaticamento mnemonico molto consistente.

Premesse a parte, da questa prima fase si è evinto che la presenza della simultaneista ha aiutato in modo considerevole il *respeaker*; esso infatti, si è trovato di fronte a un testo già di per sé corretto, coerente e coeso e, pertanto, ha avuto la possibilità di focalizzarsi soprattutto sulla dettatura al software senza dover fare importanti interventi nel testo a livello semantico. Nel momento in cui la simultaneista ha apportato in corso d'opera delle correzioni a quanto aveva detto, il *respeaker* è stato in grado di cogliere la correzione e non riportarla nei sottotitoli.

La velocità di eloquio della lettrice non era particolarmente elevata, ma il fatto di avere un passaggio aggiuntivo tra oratore e *respeaker* – l'interprete – ha fatto sì che il *respeaker* procedesse a una velocità un po' più sostenuta andando quindi a generare un testo alterato rispetto all'originale nella maggior parte dei casi. Passando all'analisi vera e propria dei risultati ottenuti è necessario premettere che non tutti gli elementi previsti dalla tassonomia di Eugeni sono stati presi in considerazione, pertanto la tabella sotto riportata (Figura 2) evidenzia soltanto ciò che è stato rilevato e analizzato ai fini della presente analisi strategica. I principali elementi presi in considerazione sono: omissioni,

[Torna all'indice](#)

errori, ripetizioni e alterazioni. In termini numerici, riferendoci alla tassonomia sopra presentata, le alterazioni hanno riguardato il 67% del testo. La densità lessicale del testo, unita ai problemi del software, ha dato luogo a una elevata percentuale di omissioni (19%) a cui il *respeaker* è ricorso per non sovraccaricare troppo la propria memoria a breve termine o per recuperare il passo con il testo della simultanea. Il restante 14% è rappresentato dalle ripetizioni (12%) e dalle unità non rese in simultanea (2%). Il 12% di ripetizioni è sintomo del fatto che il *respeaker* ha avuto la possibilità di dettare il testo così com'è senza sovraccaricare maggiormente la propria memoria.

Un ulteriore dato importante è rappresentato dagli errori. L'intera analisi si basa sulla trascrizione fedele, parola per parola, di quanto pronunciato dai professionisti. È fondamentale sottolineare questo elemento, in quanto non è stato possibile analizzare i sottotitoli generati dal software che, come accennato in precedenza, non era stato sufficientemente addestrato dal *respeaker* e ciò ha fatto sì che i sottotitoli generati contenessero una percentuale di errore molto elevata. Un esempio fra tanti è dato dai nomi propri delle persone citate nel testo originale: al momento della dettatura, questi ultimi sono stati tutti compresi correttamente dal software perché il *respeaker*, prima di iniziare l'esperimento, ha introdotto i nomi propri digitandoli nel modo corretto e successivamente pronunciandoli per far sì che a un determinato nome corrispondesse una determinata pronuncia.

Di seguito si riporta la tabella riassuntiva dei risultati ottenuti generata partendo dalla tassonomia proposta da Eugeni e adattata alle esigenze specifiche di tale analisi.

MACRO-UNITA' CONCETTUALI			
Macro-unità non rese		Macro-unità rese	
Omissioni	Errori	Ripetizioni	Alterazioni
19%	2% ¹³	12%	67%

Figura 2 – Tabella riassuntiva dei risultati ottenuti nella sottotitolazione intralinguistica.

5. Risultati della seconda parte dell'esperimento – la sottotitolazione interlinguistica

La caratteristica del testo di essere un discorso orale destinato alla comprensione immediata da parte del pubblico presente, in questo caso, così come in quello precedente, non favorisce la coesione e la coerenza del testo

¹³ Tale dato si riferisce alle unità non rese in simultanea che pertanto non risultano nel testo di arrivo prodotto dal *respeaker*.

[Torna all'indice](#)

stesso. Anche nella seconda parte quindi, si riscontra una grande quantità di informazioni che, in certe occasioni, risultano sconnesse. Il software ha giocato sempre un ruolo molto importante nel generare proposizioni difficili da comprendere o contenenti una percentuale molto elevata di errori ortografici che non facilitano la comprensione. Proprio per questo motivo si è deciso, come per la parte precedente, di non basare l'analisi strategica sul testo dei sottotitoli, ma sul testo prodotto dal *respeaker* in fase di dettatura.

Dall'analisi (Figura 3) risulta chiaro che, a differenza della parte precedente, vi è una maggior presenza di unità ripetute, ossia tradotte letteralmente, parola per parola dal *respeaker* (44%). Tale dato è sintomatico del fatto che il *respeaker*, al fine di evitare un aumento di *décalage* rispetto al testo di partenza e sovraccaricare la sua memoria, ha preferito procedere con una traduzione letterale, non riformulando le proposizioni.

Per quanto riguarda le omissioni, esse rappresentano il 10% del testo e riguardano unità del testo che non sono di importanza fondamentale, quindi si presuppone che il *respeaker* abbia deciso di tralasciarle per non sovraccaricare troppo la propria capacità mnemonica a breve termine o per recuperare il passo con il testo di partenza. Quanto alle alterazioni, in questa seconda parte esse sono risultate in percentuale minore rispetto alla parte precedente, anche se pur sempre elevate (46%) e includono le riduzioni, le espansioni e gli errori. A tale proposito è necessario specificare che, se nella prima parte non si è ritenuto di rilevante importanza procedere all'analisi degli errori, in quanto non ve ne erano in quantità consistente, in questa seconda parte si è deciso di analizzare anche quest'ultimi, perché il dato risulta più importante. Come si può evincere dalla tabella sottostante (Figura 4), la percentuale di errore contenuta nelle alterazioni è pari all'8%: tale dato è da ritenersi fondamentale in quanto la presenza di errori (in parte di senso) va a inficiare la comprensione stessa e la fruibilità del testo.

Anche in questo caso, le tabelle riportate sono state generate partendo dalla tassonomia proposta da Eugeni e adattate alle esigenze specifiche di tale analisi, pertanto alcuni elementi risulteranno mancanti.

MACRO-UNITA' CONCETTUALI		
Macro-unità non rese	Macro-unità rese	
Omissioni	Traduzioni letterali	Alterazioni*
10%	44%	46%

Figura 3 – Tabella riassuntiva dei risultati ottenuti nella sottotitolazione interlinguistica.

[Torna all'indice](#)

*ALTERAZIONI	
Alterazioni Corrette	Alterazioni Errate
92%	8%

Figura 4 – Del 46% delle alterazioni totali, il 92% è risultato corretto, mentre l'8% errato.

6. Conclusioni

Da quanto ottenuto dall'analisi di entrambe le parti di testo, si è concluso che, nella prima parte, la presenza del testo tradotto attraverso la simultanea ha permesso al *respeaker* di produrre un testo di ottima qualità, senza incappare in grandi problemi. Così non è stato, invece, per la seconda parte, in cui il *respeaker* ha fatto maggiore fatica ed è stato sottoposto a uno stress maggiore, dovuto al fatto di dover tradurre, prestare attenzione a quanto compreso e trascritto dal software e correggere eventuali errori. Il tutto evitando di allungare troppo il *décalage* e ovviamente evitando di perdere informazioni essenziali alla comprensione del testo stesso.

Si è concluso, in definitiva, che il *respeaking* intralinguistico rimane a tutt'oggi nettamente superiore rispetto al *respeaking* interlinguistico. Le ragioni sono da ricercarsi nel maggiore sforzo richiesto al *respeaker* nella prova interlinguistica e nel poco tempo a disposizione per riformulare la frase e renderla sintatticamente e semanticamente corretta per il pubblico finale al quale i sottotitoli sono destinati. Tuttavia, la prova interlinguistica è stata considerata accettabile e persino incoraggiante a livello sperimentale.

13) IL RESPEAKING INTERLINGUISTICO

Francesca Marchionne

“L'oggetto di questo articolo è la sottotitolazione interlinguistica in tempo reale. Con ciò intendo la sottotitolazione di un programma televisivo in diretta, da una lingua all'altra. Deve essere il sogno di ogni produttore di film: realizzi un film formidabile, un documentario affascinante o un qualsiasi altro tipo di programma televisivo e lo puoi far arrivare a tutti ovunque nel mondo perché, come per magia, ha dei sottotitoli nella lingua desiderata in tempo reale. Questo è ciò che la sottotitolazione in tempo reale può fare: trasmettere un programma con i sottotitoli desiderati. Quando dico "come per magia", ovviamente sono cinico. Non è un lavoro semplice.”

(libera traduzione da den Boer, 2001: 167)

1. Introduzione

Già nel 2001 den Boer promuoveva l'utilità della sottotitolazione interlinguistica in tempo reale, cioè la produzione di sottotitoli in una lingua diversa rispetto a quella del prodotto televisivo in onda. Nel mondo non mancano i tentativi di sottotitolazione interlinguistica in tempo reale. Negli Stati Uniti si utilizza un software di traduzione automatica dall'inglese ad altre lingue a partire dai sottotitoli intralinguistici prodotti, mentre altri Paesi, come l'Olanda e il Belgio, praticano la sottotitolazione interlinguistica in tempo reale da parecchi anni, ma non tramite respeaking. Per esempio, la *NOB Cross Media Facilities*, un'affermata azienda di sottotitolazione olandese, offre un servizio di *live interlingual subtitling* fin dagli anni novanta. Tale servizio si è rivelato molto utile, registrando un incremento di audience grazie alla sottotitolazione di eventi internazionali importanti. La *NOB Cross Media Facilities* utilizza due sottotitolatori professionisti per *équipe*, uno come interprete, l'altro come dattilografo, che non trascrivono l'intero testo tradotto, ma lo sintetizzano quanto più possibile. Trattandosi di un lavoro molto stancante, due *équipe* si alternano in turni di 15-20 minuti per sottotitolare lo stesso programma. Un ritardo di 10 secondi nella trasmissione del segnale permette di sincronizzare i sottotitoli e correggere gli errori. Il metodo adottato dalla *NOB* da una parte aggira l'ostacolo dell'interazione uomo-macchina, dall'altra è più lento e costoso rispetto al respeaking. L'emittente televisiva olandese *SBS*, che negli scorsi anni ha trasmesso degli eventi speciali in diretta con sottotitoli interlinguistici, nonostante qualche problema tecnico, ha registrato un picco di ascolti superiore a quello delle reti che hanno trasmesso lo stesso evento senza sottotitoli. Inoltre, seppur isolati o sporadici, non mancano nemmeno gli esempi di respeaking interlinguistico: la *BBC Welsh* offre un servizio di sottotitolazione in tempo reale tramite respeaking, dal gallese all'inglese ma non viceversa, ed è l'unica emittente televisiva al mondo a utilizzare questa tecnica; in Danimarca, sono stati fatti alcuni esperimenti con i discorsi di Obama e altri eventi internazionali importanti; nelle Fiandre, un primo test interlinguistico è stato effettuato con

[Torna all'indice](#)

successo in occasione del discorso inaugurale del presidente Obama. Quest'ultimo esperimento è stato seguito dal respeaking di altri programmi, come *So You Think You Can Dance*, *The X Factor* (inglese-olandese) e *Henin* (francese-olandese). Un caso anomalo lo ritroviamo in Svizzera, dove i respeaker si trovano spesso a dover tradurre il tedesco svizzero in tedesco standard, da dettare al software di riconoscimento del parlato. La Svizzera sarebbe un ottimo palcoscenico per la pratica del respeaking interlinguistico perché la TV pubblica svizzera è vista da francesi, italiani, tedeschi e romanci. La Svizzera grazie alla *SRG SSR*, cioè la Società Svizzera di Radiotelevisione, offre già diversi canali televisivi in ognuna di queste lingue con sottotitolazione in tempo reale intralinguistica.

Tutto ciò fa già capire che la sottotitolazione interlinguistica in tempo reale non è un'utopia. È vero che den Boer e gli esempi riportati si limitano al campo televisivo, ma perché non andare oltre? Nell'era del digitale e della globalizzazione, rendere accessibili programmi ed eventi che si svolgono in diretta in lingua straniera per un pubblico desideroso di poterne usufruire pur non conoscendo la lingua, o conoscendola poco, costituisce un servizio in più al cittadino. Perciò il presente articolo vuole introdurre il respeaking interlinguistico prevalentemente in campo televisivo e in quello convegnistico internazionale, tenendo in considerazione quali sono le sue potenzialità, la sua utilità e le sue difficoltà e riportando delle sperimentazioni a sostegno dell'utilità di questa pratica.

2. Perché e per chi

2.1 Il respeaking interlinguistico in TV

Prima di tutto è importante trovare un motivo valido per fornire un servizio di respeaking interlinguistico. In realtà i motivi sono tanti e vari a seconda dell'ambito di utilizzo. Iniziamo proprio dall'ambito televisivo a cui si è già fatto cenno. Già da diversi anni, la TV satellitare e Internet danno la possibilità di accedere a dei canali stranieri senza muoversi da casa. Questo permette un'eccezionale diffusione degli eventi nazionali e internazionali e dell'informazione su ciò che succede in tutto il mondo in tempo reale. Si pensi ai discorsi di insediamento dei Presidenti USA, agli eventi di portata internazionale come le Olimpiadi o la Notte degli Oscar o la notizia dell'ultimo minuto in diretta da una parte remota del mondo. Spesso questi eventi vengono commentati da giornalisti, presentatori televisivi e inviati speciali nei canali dei diversi Paesi quando sarebbe sufficiente un solo canale di diffusione in onda in tutto il mondo e direttamente accessibile da tutti nella propria lingua grazie ai sottotitoli interlinguistici. Inoltre, alcune aziende, come Sky già producono sottotitoli intralinguistici e interlinguistici di numerosi prodotti audiovisivi preregistrati, anche se non di quelli in diretta e in semidiretta. Avere sempre a disposizione canali televisivi in una data lingua con diverse tipologie di programmi (e quindi diversi registri, diversi livelli di difficoltà riguardo la comprensione, i lessici, gli atti

[Torna all'indice](#)

di comunicazione e l'organizzazione del discorso) sottotitolati in un'altra lingua dà l'opportunità a un immigrato, un bambino bilingue o un discendente di apprendere meglio, allenare e approfondire il proprio livello di conoscenza di una lingua straniera. Per quanto riguarda le emittenti televisive, rivolgendosi non solo al pubblico nazionale, ma anche a quello appartenente a un'altra realtà linguistica e culturale, queste hanno la possibilità di aumentare la propria audience e la propria visibilità facendosi conoscere a livello internazionale. Realizzare sottotitoli interlinguistici significa, infatti, aumentare indirettamente l'offerta televisiva per i telespettatori. La presenza dei sottotitoli significa che il telespettatore medio ha la possibilità di guardare e capire programmi diversi da quelli contenuti nel palinsesto nazionale, scegliendo tra un'offerta televisiva molto più ampia.

Infine, un'attenzione particolare meritano i telespettatori di quei Paesi dove si parlano due o più lingue come la Svizzera o il Principato di Monaco, ma anche alcune regioni dell'Italia del nord. Queste zone di solito hanno canali in diverse lingue. Nel Principato di Monaco, dove la lingua ufficiale è il francese (affiancata dal monegasco), ma dove si parla anche inglese e italiano, dal giugno 2010, grazie all'introduzione del digitale terrestre, si trasmettono 18 canali: *TF1, France2, France3, France4, France5, M6, TMC, Monaco Info, Arte, Canal+, BFM TV, iTele, Euronews, Rai Uno, Rai Due, Rai Tre, Canale5 e CNBC*. I monegaschi hanno quindi l'opportunità di guardare canali sia in italiano, sia in francese con rispettiva sottotitolazione intralinguistica. Se a ciò aggiungiamo il fatto che a Monaco il digitale (via cavo, satellite, ADSL o TNT) permette già di scegliere la lingua dei sottotitoli (in francese o in lingua originale) per prodotti preregistrati come i film, possiamo ipotizzare l'inserimento di sottotitoli chiusi, cioè selezionabili dall'utente e non sempre impressi sullo schermo, in italiano per programmi francesi in diretta, in semidiretta e in differita (o viceversa).

Quindi, nei contesti summenzionati, la sottotitolazione interlinguistica in tempo reale ha quattro benefici principali: permette al pubblico di tenersi aggiornato sugli eventi di altri Paesi e su quelli internazionali in diretta dal mondo, è un buon strumento per l'apprendimento di una lingua, aumenta l'offerta televisiva e consente alle reti televisive di aumentare l'audience.

Finora abbiamo individuato due soggetti che trarrebbero vantaggio da questa tecnica, ma esiste un terzo soggetto: le aziende di sottotitolazione, che oltre ad avere più lavoro e quindi più introiti, avrebbero l'opportunità di sfruttare maggiormente la tecnologia a loro disposizione per effettuare sia il *respeaking* intralinguistico, sia il *respeaking* interlinguistico. Inoltre, potrebbero sfruttare il *respeaking* interlinguistico anche per produrre sottotitoli in preregistrato a costi inferiori perché si ridurrebbero i tempi di produzione. Infine, nelle realtà dove la sottotitolazione interlinguistica in tempo reale è già marginalmente presente, scegliendo la tecnica del *respeaking* rispetto ad altre alternative, come nel caso dell'olandese *NOB* che utilizza un interprete simultaneo affiancato da un

dattilografo o da uno stenotipista, le aziende e le televisioni avrebbero dei costi contenuti di produzione e gestione del servizio.

2.2 Il respeaking interlinguistico in campo convegnistico

Per quanto riguarda invece l'ambito convegnistico, soprattutto quello internazionale, l'utilità del respeaking interlinguistico è evidente. Si pensi alle riunioni, alle conferenze, ai convegni, ai seminari, a eventi che si tengono a livello internazionale (incontri medici, Consigli di Amministrazione ecc.). In caso di un budget sufficiente a coprire le spese di interpretariato, i diversi interventi di persone provenienti dai quattro angoli del mondo possono essere tradotti nella lingua materna dei partecipanti o perlomeno nelle lingue più conosciute, ma sempre escludendo la parte di popolazione non udente o quella parte di popolazione che non conosce bene la lingua di arrivo del testo.

Nel caso di budget ridotti che impediscono il ricorso all'interpretazione simultanea, i partecipanti si trovano addirittura costretti a dover fare i conti con una sola lingua veicolare, spesso e volentieri l'inglese non raramente mal parlato e mal compreso impedendo così la comunicazione. In entrambi i casi citati l'ausilio dei sottotitoli è fondamentale non solo per comprendere meglio ciò che viene detto grazie a una lettura del testo in tempo reale, ma anche e soprattutto (perché l'interpretariato non lo consente) di avere un testo scritto pronto da consultare, nella lingua prescelta, subito dopo la fine dell'incontro. Si pensi ora alle sedute del Parlamento europeo o a un progetto europeo come quello dell'Euroscola che riunisce le scuole superiori dei diversi Paesi europei per passare una giornata nei panni degli europarlamentari.

Qui le principali lingue veicolari sono solitamente l'inglese, il francese e il tedesco e i partecipanti agli incontri non hanno sempre una buona conoscenza di una di queste tre lingue da e verso cui è assicurato l'interpretariato.

Ancora una volta, l'eventuale lettura di ciò che viene detto può sicuramente aiutare la comprensione. In più, considerando l'esistenza di un *live streaming* dell'evento, che va quindi in diretta in tutto il mondo grazie a Internet, la sottotitolazione in tempo reale interlinguistica tramite web permetterebbe di raggiungere un pubblico di udenti e non udenti molto più vasto rispetto a quello presente nell'emiciclo del Parlamento a Strasburgo. Il tutto allo stesso costo che viene già speso per l'interpretazione. Infatti, con il software adatto e con una formazione adeguata, una sola persona potrebbe fornire tre servizi in uno: il servizio normale di interpretariato; il servizio di sottotitolazione in loco; il servizio di sottotitolazione in tempo reale via web per il resto del mondo (sordi, ipoudenti o persone con una scarsa comprensione della lingua dell'interpretazione orale). A conferma di quanto detto, vi sono le sperimentazioni svolte l'8 e il 14 febbraio da onA.I.R., proprio nell'ambito del progetto Euroscola, a favore di un nostro socio italiano sordo, Giacomo, che

[Torna all'indice](#)

svolgeva un tirocinio al Parlamento europeo senza poter accedere adeguatamente agli incontri su cui, tra l'altro, si basava il suo tirocinio. Per rendere queste due sessioni accessibili, alcuni membri onA.I.R. hanno sperimentato il respeaking interlinguistico a distanza traducendo in tempo reale dall'inglese e dal francese per creare dei sottotitoli in italiano immediatamente fruibili da Giacomo a Strasburgo tramite PC e smartphone. In questo modo Giacomo ha potuto seguire gli incontri, le discussioni, i dibattiti e non solo ha potuto godere di un diritto fondamentale come quello della partecipazione attiva nella società di cui fa parte, ma ha potuto anche svolgere meglio il suo lavoro al Parlamento.

2.3 Il target di riferimento

Nel caso del respeaking interlinguistico quindi il *target* privilegiato non sarà più costituito prevalentemente da non udenti e ipoudenti, ma perlopiù da un pubblico di lingua e cultura diverse rispetto a quelle dell'evento. Ovviamente una sottotitolazione interlinguistica in tempo reale, seppur non appositamente pensata per un pubblico di sordi, permette anche a questi ultimi di accedere al contenuto. In questo caso la persona udente e la persona sorda sono quasi sullo stesso piano in quanto entrambe non percepiscono la componente audio verbale del testo di partenza.

Tenendo conto di tali premesse (e cioè la presenza di un pubblico con esigenze specifiche da soddisfare), è evidente che le applicazioni del respeaking interlinguistico sono potenzialmente le stesse di quelle del respeaking intralinguistico, ma con una difficoltà per il respeaker in più rispetto a quelle già affrontate negli altri articoli: il trasferimento linguistico in tempo reale da una lingua di partenza (LP) a una lingua di arrivo (LA). Questa attività è molto complessa perché diversi meccanismi mentali devono attivarsi in pochi secondi nella testa del respeaker, aumentando il carico di stress con cui quest'ultimo si ritrova a dover lavorare. L'unica differenza che separa il respeaking interlinguistico da quello intralinguistico, cioè il passaggio da una lingua all'altra, avvicina definitivamente questa tecnica a quella dell'interpretazione simultanea. Mi piacerebbe affrontare l'argomento spiegando nel dettaglio quali sono le somiglianze e le differenze tra le due professioni, ma gli elementi chiave, a mio parere, sono fondamentalmente due. Uno lega imprescindibilmente le due realtà, l'altro le divide violentemente:

- un respeaker interlinguistico deve essere prima di tutto un interprete professionista formato adeguatamente al respeaking;
- un interprete parla direttamente al suo pubblico, mentre un respeaker "parla" al pubblico tramite un software per cui il suo eloquio deve essere pulito e chiaro.

3. Sperimentazioni e difficoltà

3.1 Il progetto

A sostegno della fattibilità del respeaking interlinguistico in campo televisivo, anche se sono presenti esempi più autorevoli in diversi Paesi europei, vorrei riportare alcune osservazioni derivate dalla sperimentazione svolta nel settembre 2011 in occasione della stesura della mia tesi di laurea sul respeaking interlinguistico. La sperimentazione prevedeva la sottotitolazione in tempo reale tramite respeaking dal francese in italiano di 5 tipologie di programmi televisivi diversi (sport, meteo, TG, intervista politica, *talk show*) coinvolgendo tre interpreti professionisti italiani, di cui solo uno respeaker professionista, un suggeritore e quattro editor italiani con una buona conoscenza del francese. Il suggeritore è una figura presente perlopiù in Francia (*souffleur*) che ha il compito di confrontare il TM1 e il TM2 (testo di mezzo 1 e 2, per maggiori informazioni vedi articoli 1 e 8) e, se una frase ha bisogno di modifiche, sollecitare l'*editor*. I prodotti realizzati da ogni singolo respeaker (sottotitoli interlinguistici) sono poi stati comparati tra loro e il migliore di ogni genere televisivo è stato confrontato con il suo corrispettivo francese (sottotitoli intralinguistici) nell'ambito di un'analisi comparativa delle strategie messe in atto nel processo traduttivo dai sottotitolatori francesi e da quelli italiani. Queste operazioni hanno consentito di indagare a fondo le operazioni applicate dai respeaker nella sottotitolazione di programmi diversi tra loro per genere, durata e grado di difficoltà linguistica e hanno permesso di stabilire oggettivamente la validità delle prove effettuate nel corso della fase di sperimentazione del progetto con buoni risultati.

3.2 I risultati

Su 17 prove in diretta analizzate, 8 sottotitolazione interlinguistiche sono risultate utilizzabili come prodotto finale, una è risultata accettabile ma non utilizzabile, 2 utilizzabili solo in parte e 6 sono state scartate completamente. I problemi riscontrati sono principalmente dovuti alla lingua, non solo a causa del passaggio interlinguistico, ma anche a causa delle caratteristiche della LP, come il registro, la densità lessicale e la complessità grammaticale. Infatti se un testo di partenza (TP) ha un linguaggio chiaro, una grammatica semplice e un registro formale, è più facile da sottotitolare anche se avesse un'elevata densità lessicale; un TP in cui il linguaggio è poco chiaro, familiare e intricato a livello grammaticale è più complicato da capire e tradurre in così poco tempo.

Altri problemi sono dovuti alle diverse velocità che il respeaker deve gestire: la velocità di eloquio del TP e la velocità di eloquio del TM1. Sulla prima non si può intervenire. Diversi TP hanno infatti diversi cambi di turno ciascuno con diversi ritmi di eloquio a cui il respeaker non può far altro che adattarsi. In genere TP più lenti e/o con meno locutori portano il respeaker a seguire più da vicino il testo e alla generazione di meno errori; TP più veloci e/o con più locutori portano a una maggiore riduzione del testo con una prevalenza di omissioni e un

[Torna all'indice](#)

numero elevato di errori. Il secondo tipo di velocità, quella del TM1, è, invece, decisa dal respeaker, anche se dipende in gran parte dalla precedente. Per TP lenti, il respeaker si può dedicare maggiormente alla cura del TM1 e della sua enunciazione per un buon riconoscimento da parte del software; per TP veloci, il respeaker sarà costretto a ridurre il testo e a modulare il ritmo secondo le esigenze del software ricordandosi che non si tratta di un'interpretazione simultanea, ma di un respeaking interlinguistico che passa quindi per una macchina. A queste si aggiunge la velocità di correzione da parte dell'editor che a volte cancella sottotitoli corretti nel tentativo di recuperare il ritardo accumulato.

Nella comparazione con l'intralinguistico, si è notato che il respeaker interlinguistico tende a usare più la riformulazione tramite un uso maggiore di compressioni ed espansioni, mentre il respeaker intralinguistico cerca di creare sottotitoli *verbatim* (parola per parola) nei limiti consentiti dal TP. Solo in un'occasione le unità concettuali rese dal respeaker italiano sono risultate in numero maggiore rispetto a quelle del respeaker francese. Per il resto, non si sono notate grandi differenze di omissione a livello di unità concettuali.

Un altro elemento interessante da analizzare nel respeaking, in particolar modo interlinguistico, è il maggiore o minore peso semiotico della componente video rispetto a quella audio nel TP. In questo caso le omissioni di unità concettuali non semioticamente rilevanti dipendono, oltre che da tutto il prodotto audiovisivo, soprattutto dalla presenza di tratti dell'oralità nel TP, che ne diminuiscono la rilevanza. Si è infine rilevato, in generale, un alto numero di errori evitabili nel respeaking interlinguistico grazie a una formazione mirata del respeaker, molto esercizio e una maggiore attenzione al TP e alla produzione del TM1.

Per quanto riguarda le figure coinvolte, il suggeritore si è rivelato utile in più di un'occasione. Tuttavia, il fatto che le prove in cui non è stato coinvolto siano state comunque migliori di quelle in cui è stato coinvolto significa che il respeaking interlinguistico è fattibile senza una figura in più da pagare, a patto che si fornisca una formazione adeguata al respeaker e all'editor. Se proprio si vuole pagare una persona in più, allora è meglio investire su un secondo editor in aiuto al primo visto il carico e la tipologia di errori che questi ha dovuto gestire.

Nonostante la sperimentazione riguardi esclusivamente l'ambito televisivo, le osservazioni sopra citate possono a mio parere essere valide anche in altri ambiti in cui è possibile utilizzare il respeaking interlinguistico, come per esempio quello convegnistico. Questo perché i TP presentavano velocità di eloquio, registri, cambi turno, tipi di eloquio molto differenti tra loro; nell'analisi non sono stati presi in considerazione né il ritardo, né il tempo di permanenza dei sottotitoli trasmessi sullo schermo, né la concomitanza con le immagini; l'esperimento si è caratterizzato anche per la diversità delle prove in modo da

[Torna all'indice](#)

testare più metodologie di lavoro come l'aggiunta di un suggeritore, l'inserimento della punteggiatura da parte del respeaker o da parte dell'editor, l'inserimento del cambio locutore e il respeaking interlinguistico in semidiretta. Ovviamente, questa è solo una prima sperimentazione che richiederebbe una ricerca approfondita da diversi punti di vista, come sul pubblico di destinazione (Esigenze? Aspettative? Ricezione? Percezione? Comprensione del TA? Suggestimenti per miglioramenti?) e sul tempo di realizzazione e trasmissione dei sottotitoli (Troppo ritardo? Si può ridurre? Come?).

4. Conclusioni

Il respeaking interlinguistico è in definitiva concretamente fattibile e si possono ottenere buoni risultati se si fornisce una giusta preparazione agli operatori e si prendono degli accordi ben precisi con i clienti che ne fanno richiesta, riguardo per esempio la disponibilità del testo prima dell'evento e/o di tutte le informazioni necessarie per una buona sottotitolazione. Perché, allora, non si fornisce ancora un servizio di sottotitolazione di tutti i programmi trasmessi via satellite e di tutti gli eventi di interesse internazionale che si svolgono in tempo reale? Le risposte possono essere diverse. Innanzitutto, perché manca un interesse da parte dei *broadcaster* in campo televisivo e degli organizzatori degli eventi negli altri casi di investire in un servizio in più al pubblico. In secondo luogo, non tutti i Paesi hanno sviluppato dei software di riconoscimento del parlato per la propria lingua. Infine il respeaking intralinguistico ha impiegato diversi anni per farsi strada e quindi convincere i clienti ad affidarsi anche al respeaking interlinguistico, che mostra maggiori ostacoli e difficoltà rispetto al primo, sarà un'impresa ancora più ardua: il futuro, l'incognito, la novità fanno infatti paura. Come è già successo col respeaking intralinguistico, bisognerebbe ragionare prima in termini di quantità, cioè sulla diffusione dell'attività, buttandosi nelle sperimentazioni pubbliche, dimostrandone l'utilità, sbagliando e migliorando, formando i nuovi respeaker interlinguistici e solo dopo si potrà ragionare in termini di qualità per un servizio sempre migliore e una pratica che si spera diventi "normalmente" diffusa.

14) SVILUPPI FUTURI – PROGETTI DELL'ASSOCIAZIONE INTERNAZIONALE DI RESPEAKING onA.I.R.

Tiziana Trapani

1. Introduzione

“Spesso le idee si accendono l’una con l’altra, come scintille elettriche”. A dirlo, nel XIX secolo, fu Friedrich Engels, economista, filosofo e politico tedesco. E se ci riflettiamo meglio ci rendiamo conto di come, in realtà, le idee non siano altro che energia creatrice, scaturita direttamente dalle nostre menti, che si nutre dell’energia creatrice delle menti altrui, generando un processo senza fine che ha condotto l’umanità al progresso. I tempi odierni, purtroppo, tendono a soffocare questa energia tramite una cappa di disillusioni e sfiducia nel futuro: la crisi economica, la precarietà del lavoro, la disinformazione dei media, la mancanza di fiducia nei valori creano una cupola di negatività che inibisce l’energia creativa delle menti più giovani. Credo sia stato questo, in fondo, il motivo che ha portato, il 4 marzo 2012, alla nascita di onA.I.R. – Associazione Internazionale Respeaking: non permettere che le idee partorite così, per caso, attraverso la quotidianità e le esperienze di ognuno di noi, andassero perse nel grigiore dei tempi moderni, ma, al contrario, lasciare che si accendessero le une con le altre.

2. OltreSuoni – Dalla voce alla scrittura per una radio accessibile a tutti

Il primo progetto dell’associazione, **OltreSuoni**, rappresenta in un certo senso il nostro primo tentativo di incanalare un’idea empirica in un’applicazione pratica utile per due categorie di persone svantaggiate da un punto di vista comunicativo: i sordi, per motivi sensoriali, e gli immigrati, per motivi linguistici. OltreSuoni non vuole rappresentare in alcun modo un punto di arrivo quanto piuttosto un punto di partenza, un esperimento per individuare tutte le problematiche pratiche e tecniche e potere, di conseguenza, migliorare l’idea iniziale per poi espanderla su più vasta scala, coinvolgendo le comunità di tutta Italia e, perché no, anche estere.

“Oltresuoni” è stato ideato in collaborazione con l’Associazione TDM 2000, che consiste nella sottotitolazione in tempo reale tramite respeaking di un’emittente radiofonica web (Unica Radio, la radio dell’università di Cagliari) per sordi e per persone con una conoscenza limitata della lingua italiana, quali stranieri di recente immigrazione.

La durata della sperimentazione sarà di sei mesi e coinvolgerà un campione di 20 individui, di età ed estrazione sociale eterogenea, ripartiti equamente tra normoudenti e persone con deficit uditivi o limitata conoscenza della lingua italiana. I soggetti del campione dovranno seguire dei programmi specifici che saranno trasmessi sia in modalità tradizionale, sia sottotitolata e

[Torna all'indice](#)

dovranno confrontarsi su di essi, rispondendo a domande di comprensione e condividendo opinioni e punti di vista in appositi spazi web.

Tale condivisione rappresenterà il materiale la cui analisi, svolta da un team di linguisti, sociologi e psicologi, misurerà la capacità del *respeaking* di ridurre il gap esistente tra udenti e non udenti nella fruizione di un contenuto radiofonico.

3. RadioText – Leggendo anche noi ascoltiamo la radio

Il progetto “RadioText – leggendo anche noi ascoltiamo la radio” si propone di sviluppare il discorso estendendolo nella portata. Tra i soci onA.I.R. ci sono anche traduttori e interpreti professionisti che lavorano su varie combinazioni linguistiche, permettendo all’Associazione stessa di potersi cimentare anche con sfide innovative, come la sottotitolazione interlinguistica in tempo reale tramite *respeaking*. Un esperimento compiuto dalla presidente Francesca Marchionne nell’ambito della stesura della sua tesi di laurea magistrale, infatti, ha dimostrato come sia possibile per un *respeaker* professionista con una data conoscenza di una lingua straniera e delle tecniche di interpretazione simultanea, sottotitolare in tempo reale un programma televisivo da una lingua di partenza (nel caso specifico dell’esperimento, il francese) a una lingua di arrivo (l’italiano).

Naturalmente occorre avere un addestramento e una preparazione adeguati perché l’attività del *respeaker* è complessa già a un livello intralinguistico, cioè quando il passaggio dal canale orale a quello scritto avviene all’interno della stessa lingua. Il processo di traduzione simultanea che deve avvenire nella mente del *respeaker* interlinguistico aggiunge un ulteriore tasso di difficoltà, ma le potenzialità sono straordinarie. Nell’era della globalizzazione e dell’interattività multimediale sarebbe impensabile negare alle persone con disabilità sensoriali la possibilità di fruire della programmazione televisiva in diretta o di eventi di particolare rilievo culturale: l’integrazione è frutto di una serie di processi sociali e culturali che rendono l’individuo membro di una società, ma se l’individuo non ha accesso alla totalità di questi processi, potrà mai dirsi veramente integrato nella società di cui fa parte? E ovviamente in questo caso il termine società non comprende solo i membri della comunità linguistica di appartenenza, ma viene inteso in senso generalizzato come la società umana: perché dunque non permettere di seguire in tempo reale programmi televisivi di interesse internazionale, come i discorsi dei più importanti presidenti stranieri, abbattendo contemporaneamente le barriere linguistiche e quelle della disabilità?

[Torna all'indice](#)

4. PALiOSS – Piattaforma per l'Accessibilità alla Lingua Orale per Sordi e *Stranieri*

Ed è votato proprio a questo duplice scopo un altro progetto su cui onA.I.R. sta lavorando. È stato chiamato “**PALiOSS**”, acronimo di Piattaforma per l'Accessibilità alla Lingua Orale per Sordi e Stranieri. Il progetto, che attualmente si trova ancora allo stadio embrionale, prevede la creazione di una piattaforma accessibile tramite Internet a cui sordi e stranieri possono rivolgersi per richiedere vari tipi di servizi che vanno dalla traduzione simultanea alla sottotitolazione in tempo reale. Un po' come avere sempre a disposizione un *respeaker*/traduttore personale in ogni angolo dell'Europa e del mondo. OnA.I.R. è costruita intorno alle idee dei propri soci, intorno all'osservazione diretta dei fatti della vita quotidiana e intorno alla ricerca di soluzioni che possano garantire la piena integrazione di ogni tipo di cittadino, senza discriminazione alcuna. In una mail che ci è stata inoltrata dall'Ing. Pirelli, una ragazza bolognese, audiolesa dall'età di 3 anni ha portato l'attenzione sull'utilità di sviluppare un *software* che permetta la scrittura simultanea di un colloquio telefonico sul computer, collegato al cellulare o ad altro apparecchio telefonico, senza previo addestramento della voce dell'interlocutore. Da un lato, infatti, i *software* cosiddetti *speaker-dependent* garantiscono una maggiore accuratezza rispetto ai *software speaker-independent*, ma hanno lo svantaggio che è necessario addestrarli tramite la creazione di un profilo vocale: e si capisce come, nel caso delle telefonate, sia difficile, se non impossibile, creare un profilo vocale per ogni contatto. Questa è senza dubbio una questione interessante che sarebbe un peccato non approfondire.

5. **A, E, I, OnAIR – Il prof. si sottotitola da solo**

Attraverso il nuovissimo progetto “**A, E, I, OnAIR**.” si promuovono, all'interno delle scuole superiori nelle cui classi ci sono studenti sordi, delle attività mirate alla promozione del *respeaking* come forma di integrazione di persone con disabilità sensoriali. In particolare, la lezione verrà *respeakerata* direttamente dal docente e resa immediatamente disponibile per lo studente sordo che potrà leggerla su uno schermo in tempo reale. Successivamente il file di testo ricavato verrà affidato a due studenti della stessa classe che avranno il compito di correggerlo, adattarlo e inviarlo al docente. L'insegnante ne controllerà la correttezza formale e contenutistica e lo distribuirà alla classe, creando così delle dispense su cui poter studiare.

6. **Lectio subtutulis – Prendo appunti per te**

Una sorta di evoluzione del precedente progetto è rappresentata da “**Lectio subtutulis**” che si propone di facilitare e migliorare il percorso accademico di studenti sordi, stranieri e normoudenti. Prendere appunti diventa più facile attraverso la sottotitolazione delle lezioni universitarie, non solo per gli studenti sordi che così possono seguire autonomamente la lezione, senza l'intervento di un tutor che prenda appunti, ma anche per gli stranieri e i normoudenti. Capita,

[Torna all'indice](#)

infatti, che ci si distraiga e si perda il filo del discorso e questo può risultare ancora più problematico per gli studenti stranieri (per esempio i partecipanti a un programma Erasmus): l'ausilio dei sottotitoli permette di ritrovare il filo senza perdere troppo tempo. Inoltre, viene facilitata anche la creazione di dispense da parte del professore che potrà elaborarle sulla base del file contenente i sottotitoli, disponibile già al termine della lezione.

7. TVonA.I.R. – L'informazione è uguale per tutti

Dalla collaborazione tra Libera.TV e onA.I.R. è nata **TVonA.I.R.**, la prima social web TV espressamente concepita per rendere accessibile sia alla popolazione non udente sia a quella straniera i contenuti di interesse generale presenti sul web.

L'accordo di collaborazione prevede la gestione da parte di onA.I.R. di uno spazio tutto suo per trasmettere video di interesse generale sottotitolati dai propri membri. Così Libera.TV potrà raggiungere finalmente anche quella parte di popolazione che a causa di problemi uditivi o linguistici prima restava fuori dalla comunicazione di una delle social web tv più seguite e premiate d'Italia. Dal canto suo onA.I.R. avrà una grande opportunità per realizzare sia i propri scopi statutari sia per diffondere le proprie attività alle più svariate categorie di utenti.

I contenuti di TVonA.I.R. sono divisi per canali a seconda dell'area di interesse. Per il momento ci sono solo due canali, uno dedicato alla ripubblicazione dei servizi di maggiore interesse sociale proposti da Libera.TV; e uno dedicato alla conferenza di presentazione della nostra associazione. Tuttavia è in progetto l'apertura di altri canali tematici di interesse nazionale e sociale per donne, bambini, immigrati, lavoratori e altre categorie sociali che potranno comprendere tutti quei servizi audiovisivi prodotti dai soci onA.I.R., trovati su altri siti di condivisione o inediti, sia in lingua italiana, sia in altre lingue che i soci possono proporre di far sottotitolare al team che attualmente sta già lavorando alle prime sottotitolazioni.

8. Note Vibranti – Concerti accessibili per i sordi

È proprio per questo che il 2013 di onA.I.R. è iniziato con "**Note vibranti**": un evento simbolo della volontà di garantire la maggiore accessibilità possibile, anche a costo di affrontare quelli che sembrano dei paradossi. "Concerto per sordi", infatti, potrebbe sembrare una contraddizione in termini ma se si cerca di pensare con meno superficialità ai problemi collegati all'ascolto della musica da parte di una persona sorda si può raggiungere un obiettivo importante: quello di regalare le emozioni della musica a un pubblico che non può accedervi in maniera diretta. Nella splendida cornice del Molino Excelsior, a Valderice (TP), l'Associazione Internazionale di Respeaking ha organizzato il 2 gennaio 2013 un vero e proprio concerto a cui hanno partecipato udenti e sordi con un entusiasmo quasi commovente.

[Torna all'indice](#)

La collaborazione con i Matrioskando, un quartetto locale specializzato in musica di tradizione irlandese e scozzese, e con l'Accademia Europea Sordi ONLUS di Palermo ha permesso di sperimentare un ascolto diverso della musica, con gli occhi e con il cuore invece che con le orecchie. I canti tradizionali in gaelico irlandese e scozzese, infatti, sono stati interpretati artisticamente in LIS mentre su uno schermo antistante i musicisti, i *respeaker* dell'associazione hanno proiettato i sottotitoli in tempo reale delle parti parlate e dei testi originali delle canzoni, con la relativa traduzione. Ma per permettere alla musica di arrivare direttamente al cuore sono stati distribuiti all'inizio della serata dei palloncini colorati: l'aria contenuta all'interno, infatti, consente la diffusione delle onde sonore che, intrappolate dal lattice, possono essere percepite semplicemente appoggiando le mani o qualsiasi altra parte del corpo sul palloncino. Vedere l'aria estasiata di un bambino sordo che, appoggiando il palloncino all'orecchio, riusciva a sentire le vibrazioni musicali è stata la migliore ricompensa che ci si potesse aspettare. È proprio per questo che "Note vibranti" è diventato il nome di un progetto che onA.I.R. si sta impegnando a riproporre su più larga scala (introducendo, ad esempio, la traduzione in colori dei suoni) in tutta Italia, contribuendo non solo a vincere le barriere sensoriali ma anche a sostenere il talento artistico di giovani professori di musica.

Lo scrittore Lev Tolstoj diceva: "Tutte le *idee* che hanno enormi conseguenze sono sempre *idee semplici*." E forse, nel nostro piccolo, grazie alla semplicità delle nostre idee e alla convinzione e all'entusiasmo con cui le portiamo avanti, cercando di realizzarle, potremo dare origine a grandi innovazioni in grado di migliorare la vita di un'importante fetta di popolazione, a volte poco considerata.

15) NOVITA' TECNOLOGICHE

15a) IL PROGETTO EUROPEO EU BRIDGE – UN PONTE PER COLMARE IL DIVARIO LINGUISTICO

Margit Rödder (traduzione di Silvia Bellucci)

1. L'obiettivo

Il progetto ha lo scopo di sviluppare la trascrizione automatica e le tecnologie relative alla traduzione che consentono di progredire nel campo del sottotitolaggio multimediale e dei servizi di traduzione di documenti audiovisivi tra lingue europee ed extra europee. Il progetto fornirà tecnologie in streaming che potranno trasformare i discorsi di conferenze, le riunioni e le conversazioni telefoniche in testo prodotto in un'altra lingua.

Pertanto il progetto EU-BRIDGE si pone l'obiettivo di riunire competenze accademiche, ingegneristiche e aziendali al fine di creare offerte competitive per l'esistente domanda di traduzioni, comunicazione, elaborazione di contenuti e editoria. I quattro ambiti di applicazione sono: la traduzione di sottotitoli per trasmissioni televisive, le traduzioni di lezioni universitarie, le traduzioni per il Parlamento Europeo, le traduzioni di comunicazioni per dispositivi portatili. I potenziali utenti finali del progetto sono le aziende europee che operano sul mercato audiovisivo (in particolare nell'ambito del sottotitolaggio televisivo e della traduzione).

2. Obiettivi e innovazioni

I partner di EU-Bridge stanno sviluppando le funzionalità delle più recenti tecniche di traduzione automatica citate in precedenza, in vista di nuovi e più ambiziosi ambiti di applicazione a livello aziendale. Ciò significa approdare ad avanzate tecnologie per il linguaggio parlato che permettano di elaborare e trasferire i contenuti delle informazioni umane da una lingua in un'altra, in situazioni che fino a oggi non potevano essere gestite con tecniche automatiche.

Un'ulteriore importante sfida è ridurre i costi della raccolta dati e la dipendenza da grandi quantità di dati raccolti e annotati manualmente. Il gruppo di partner propone la creazione di una delle più grandi raccolte combinate di lingue disponibili sia per il riconoscimento del parlato che per la traduzione.

L'ultimo obiettivo riguarda la rapida transizione tecnologica e l'inserimento sul mercato, da attuare attraverso la creazione di servizi erogati in sostituzione del trasferimento di software, rendendo così l'installazione parte del progetto.

[Torna all'indice](#)

3. Risultati

Si potranno avere tecnologie di traduzione automatica ad alte prestazioni, sistemi avanzati sviluppati durante il progetto nel servizio di trascrizione e traduzione di informazioni commerciali, più redditizi rispetto a quelli attualmente esistenti.

4. Impatto

I meccanismi di ricerca e inserimento sul mercato di tale progetto avranno un impatto sulla posizione competitiva dell'Europa sul mercato multilingue, attraverso la fornitura di prodotti e servizi migliori ai cittadini e alle imprese.

Il consorzio, composto da alcune tra le massime istituzioni europee nel campo della ricerca, ha l'opportunità di diventare un leader scientifico e tecnologico grazie a una ipotesi largamente accettata: il progetto dimostra come le tecnologie avanzate nella traduzione del parlato possono favorire le imprese e la società.

Il progetto rafforzerà la cooperazione, il dialogo e i partenariati tra ricerca e industria e permetterà una migliore comprensione delle necessità dell'utente finale.

5. I partner del progetto

I partner del progetto sono:

- Istituto di tecnologia di Karlsruhe (KIT), Germania
- Fondazione Bruno Kessler, Italia
- Polish Japanese Institute of Information Technology, Polonia
- Università tecnica di Aquisgrana (RWTH), Germania
- Università di Edimburgo, Regno Unito
- Università di Scienza e Tecnologia di Hong Kong, Cina
- Red Bee Media Limited, Regno Unito
- Mobile Technologies GmbH, Germania
- Pervoice SPA, Italia
- Accipio Projects GmbH, Germania
- Andrexen, Francia

Coordinatore del progetto è l'Istituto di tecnologia di Karlsruhe (KIT) nella persona del Prof. Alex Waibel, che dirigerà i lavori a partire dal 1° febbraio 2012 al 31 gennaio 2015. Il lavoro, reso possibile grazie all'ottenimento di un finanziamento erogato dall'Unione Europea attraverso l'accordo di sovvenzione n° 287658, è disponibile negli aggiornamenti al sito web: www.eu-bridge.eu

15b) IL PROGETTO EUROPEO SAVAS – RICONOSCIMENTO DEL PARLATO PER RESPEAKING E SOTTOTITOLAZIONE IN TEMPO REALE

Carlo Aliprandi

1. Introduzione

A seguito delle recenti leggi nazionali e direttive internazionali, la domanda e le aspettative in ambito di sottotitolaggio di programmi televisivi sono progressivamente cresciute. Le tecniche di sottotitolaggio basate sulle sole risorse umane stanno diventando sempre meno praticabili e diffuse, anche in relazione al costo del processo, che spesso coinvolge diversi operatori di alta professionalità. Di conseguenza, gli operatori televisivi e le aziende del settore sono alla ricerca di tecniche alternative, che, mantenendo l'attuale livello qualitativo, permettano una produttività più elevata rispetto ai tradizionali processi di sottotitolaggio.

In questo quadro la precisione ottenuta dalla tecnologia ASR può, in alcuni domini e ambiti applicativi, essere particolarmente alta. Tuttavia, gli attuali sistemi di respeaking, basati su sistemi commerciali di ASR per il parlato continuo, soffrono di alcune limitazioni: tali sistemi, infatti, sono stati sviluppati per applicazioni di dettatura e le loro performance possono essere limitate per certe attività di sottotitolaggio, specialmente in domini specifici. I sistemi attuali sono di tipo Speaker Dependent, ossia devono essere addestrati sulla voce del parlatore per fornire prestazioni ottimali, operazione che può richiedere anche uno sforzo significativo e continuativo nel tempo. I sistemi attuali, inoltre, sono stati sviluppati per le principali lingue (ad esempio inglese, francese, spagnolo, tedesco) e pertanto non sono disponibili per diverse lingue europee, soprattutto per le lingue regionali e minoritarie. L'alto costo associato alla collezione e annotazione delle risorse linguistiche indispensabili per creare sistemi ASR è una delle principali ragioni che ha sinora ostacolato lo sviluppo di nuove lingue e di nuovi domini applicativi. Quindi, sebbene la tecnologia ASR si stia dimostrando molto utile per il sottotitolaggio, una delle barriere alla sua diffusione consiste nella limitazione degli attuali sistemi e negli alti costi legati allo sviluppo di nuove lingue o nell'adattamento a nuovi domini.

2. Il progetto SAVAS

Il progetto SAVAS (Sharing AudioVisual language resources for Automatic Subtitling), cofinanziato dalla Comunità Europea, si pone come obiettivo principale il superamento di questa barriera.

SAVAS sta raccogliendo risorse linguistiche audiovisive per diverse lingue europee, alcune delle quali sono tipicamente considerate lingue minoritarie, e sta realizzando un'innovativa tecnologia ASR dedicata in maniera specifica alle

[Torna all'indice](#)

esigenze di sottotitolazione automatizzata per il settore del Broadcasting e dei New Media.

1. Obiettivi

Più in particolare, gli obiettivi di SAVAS sono:

1. Definire una metodologia efficace per l'acquisizione e l'annotazione di risorse linguistiche audiovisive prodotte da Broadcaster e operatori dell'industria dei Media;
2. Sviluppare una piattaforma per la condivisione delle risorse tra l'industria dei Broadcaster e gli sviluppatori di sistemi ASR, che consenta lo sfruttamento commerciale dei risultati, nel pieno rispetto di eventuali vincoli legali;
3. Sviluppare una nuova tecnologia ASR utilizzabile per la produzione di sottotitoli multilingua in sistemi di respeaking e di sottotitolazione;
4. Integrare e valutare i sistemi in diversi scenari e in 3 domini applicativi (News, Sport e Interviste/Dibattiti), dimostrando l'impatto della condivisione delle risorse audiovisive nello sviluppo di soluzioni alternative ai tradizionali processi di sottotitolazione.

2. Aspetti tecnologici

La tecnologia di ASR di base è stata sviluppata per trattare efficacemente il *parlato spontaneo*, ed è in grado, elaborando in maniera robusta il segnale audio, di eliminare esitazioni e balbettii, di identificare ripetizioni e interloquzioni, di eliminare rumori di fondo e musica.

Si tratta di una tecnologia di tipo Speaker Independent, basata su modelli acustici opportunamente ottimizzati, che non richiede alcun addestramento alle voci dei parlatori, ed è in grado di trascrivere correttamente il parlato anche in presenza di accenti dialettali oppure pronunce straniere.

3. Principali risultati

Nuove applicazioni dedicate alla sottotitolazione sono ora disponibili per il settore del Broadcasting e dei New Media:

- **S.Live!**, il sistema di **sottotitolazione automatica online**, in grado di operare in tempo reale, producendo con operatività online trascrizioni e sottotitoli allineati con audio e video in broadcasting e in streaming;
- **S.Server!**, il sistema di **sottotitolazione offline**, in grado di trascrivere in modalità batch file audio e video, producendo sottotitoli con segmentazione automatica e riconoscimento del cambio di oratore;

[Torna all'indice](#)

- **S.Respeak!**, il nuovo sistema di **Respeaking collaborativo** per la sottotitolazione live da remoto.

La tecnologia è attualmente disponibile in 10 lingue: italiano, inglese, spagnolo, francese, tedesco, inglese-americano, portoghese, basco, brasiliano e svizzero-italiano. Altre lingue sono attualmente in fase avanzata di sviluppo.

4. **Sviluppi futuri e prodotti**

La metodologia SAVAS sta dimostrando di essere utile per lo sviluppo di tecnologie ASR anche per lingue minoritarie o regionali, come il basco o lo svizzero-tedesco, che non vengono normalmente sviluppate per carenza di risorse. Su questa linea, sono attualmente oggetto di attività di Ricerca e Sviluppo altre metodologie innovative per il tuning e l'adattamento automatico sia del Dizionario di Base sia del Modello del Linguaggio, tali da permettere un'ulteriore riduzione dei tempi di adattamento ad un nuovo dominio e dei costi per lo sviluppo di una nuova lingua.

Ma minori risorse non vuol dire minor qualità. I primi test effettuati hanno infatti dimostrato che la tecnologia ASR sviluppata fornisce una buona copertura di partenza nel dominio giornalistico di riferimento, fornita dal Dizionario di Base. Poiché, tuttavia, il linguaggio naturale è per natura un fenomeno altamente dinamico sia in linea teorica (ad es. neologismi, nomi di persone o nuovi toponimi) sia in linea pratica (ad es. parole poco frequenti o forme flesse di alcune parole), alcune parole possono mancare dal Dizionario di Base. Tali parole vengono definite Out of Vocabulary word (OOV). Normalmente, più alta è la percentuale di OOV in un dato dominio, maggiore probabilità si ha di avere una degradazione della precisione di trascrizione.

Per maggiori informazioni e per tenersi aggiornati sugli sviluppi del progetto SAVAS, consultare il sito <http://www.fp7-savas.eu/>

15c) IL PROGETTO EUROPEO SUMAT – SOTTOTITOLAZIONE ASSISTITA DALLA TRADUZIONE AUTOMATICA

Panayota Georgakopoulou (traduzione di Silvia Velardi)

1. Introduzione

La sottotitolazione gioca un ruolo chiave nel rendere accessibile il contenuto audiovisivo in più lingue. Grazie all'uso e alla trasmissione crescenti di contenuti digitali multimediali e multilingue attraverso il web, i DVD e il cinema e grazie alle attuali politiche europee e nazionali che promuovono la sottotitolazione di contenuti trasmessi dalle reti televisive pubbliche, la domanda per la produzione di sottotitoli è in continuo aumento. Questo aumento della domanda fa crescere la pressione sulle aziende di sottotitolazione per ridurre i tempi di consegna, mantenendo la qualità alta ed i costi di gestione bassi. Con l'aumento del volume di contenuti, basarsi unicamente sulla traduzione umana non rappresenta più un modello sostenibile. Vi è la necessità di nuove tecnologie che consentano all'industria di stare al passo con le richieste del mercato.

La sottotitolazione e la traduzione audiovisiva sono state riconosciute come settori che potrebbero trarre grandi vantaggi dall'introduzione della traduzione automatica statistica (**SMT** – *Statistical Machine Translation*), una tecnologia volta ad aumentare la produttività del traduttore (cfr. Volk, 2008; Hardmeier e Volk, 2009; De Sousa et al, 2011). La traduzione automatica statistica (Koehn, 2010) è un paradigma di traduzione automatica che ha avuto un impulso enorme, sia nella comunità di ricerca sia nel settore commerciale. In questo approccio, un sistema di traduzione automatica apprende regole di traduzione da grandi quantità di testi tradotti, allineando i segmenti nelle lingue di origine e di destinazione e calcolando le statistiche sui dati. Inoltre, tali sistemi includono un sistema di misurazione della fruibilità delle traduzioni proposte, tramite l'apprendimento automatico delle ricorrenze da dati aggiuntivi nella lingua di arrivo.

Negli ultimi dieci anni, l'introduzione e l'uso di "template" (cioè file contenenti sottotitoli con *timecode* nella lingua di partenza per la traduzione in altre lingue) ai fini della sottotitolazione ha fatto sì che un numero crescente di corpora paralleli di sottotitoli sia stato e continui a essere generato dalle aziende di sottotitolazione. I sistemi SMT possono essere adattati alle esigenze della sottotitolazione incorporando tali dati professionalmente tradotti da una varietà di generi e domini di sottotitolazione.

2. Il progetto

Il progetto **SUMAT** (*SUMAT: An Online Service for SUBtitling by MAchine Translation* – Servizio Online per la Sottotitolazione tramite Traduzione Automatica) si propone di verificare la possibilità di integrare la SMT nel processo di traduzione dei sottotitoli e quindi di fornire uno strumento innovativo in grado di supportare i processi di traduzione dei sottotitolatori freelance e delle aziende

[Torna all'indice](#)

di sottotitolazione. Si sta sviluppando un servizio online di traduzione di sottotitoli in nove lingue europee e 14 coppie di lingue diverse, che sarà disponibile nel 2014. Le combinazioni linguistiche bi-direzionali sono: inglese-olandese, inglese-francese, inglese-tedesco, inglese-portoghese, inglese-spagnolo, inglese-svedese e serbo-sloveno. Queste coppie di lingue coprono circa il 62,43% delle lingue parlate nei 27 Stati membri della UE.

Il motivo principale per cui non ci sono ancora strumenti efficaci per i servizi di sottotitolazione interlinguistica automatica è la mancanza di una qualità sufficientemente alta dei corpora paralleli di sottotitoli che possano addestrare i modelli SMT. Le risorse disponibili gratuitamente, come il corpus OPUS OpenSubtitle, pur essendo disponibili in alcune lingue (Tiedemann, 2009), non hanno una qualità confermata, quindi è ancora da determinare quanto possano essere utili. Inoltre, quei corpora di sottotitoli prodotti professionalmente e di alta qualità sono di proprietà delle aziende di sottotitolazione o dei loro clienti.

Il progetto SUMAT è una collaborazione tra quattro aziende di sottotitolazione e cinque partner tecnici. Le aziende di sottotitolazione sono Deluxe Digital Studios (<http://www.bydeluxe.com>), Voice & Script International (<http://www.vsi.tv/>), InVision Ordertiteling (<http://www.subtitling.net/>) e Titelbild Subtitling and Translation (<http://www.titelbild.de/>). Le prime due sono grandi multinazionali di sottotitolazione con sedi in tutto il mondo, le ultime due sono leader nella sottotitolazione all'interno dei rispettivi mercati locali. I partner tecnici sono ATC (www.atc.gr), Capita Translation and Interpreting (www.capitatrtranslationinterpreting.com), TextShuttle (www.textshuttle.ch), l'Università di Maribor (www.um.si) e Vicomtech (www.vicomtech.es), una vasta gamma di organizzazioni che include dipartimenti di ricerca universitari e aziende specializzate in soluzioni di traduzione automatica per l'industria.

3. Avanzamento del progetto

Attraverso questa collaborazione, abbiamo raccolto dei corpora di sottotitoli tradotti in sette coppie di lingue, prodotti da traduttori di sottotitoli esperti professionalmente qualificati che lavorano nella loro lingua materna e che sono stati assunti dalle suddette aziende di sottotitolazione dopo rigorosi test e procedure di assunzione.

Questi corpora sono stati poi elaborati per realizzare un corpus parallelo per ogni coppia di lingue. I corpora contengono dai 200.000 ai 2 milioni di sottotitoli per ogni coppia di lingue. Inoltre abbiamo raccolto altri corpora monolingue nelle lingue di cui sopra per un ulteriore addestramento dei sistemi SMT e completato i nostri corpora paralleli con altri corpora di alta qualità disponibili (per esempio, il corpus EUROPARLTV).

A partire dal febbraio 2013, sono stati costruiti i nostri sistemi SMT standard nelle coppie linguistiche di cui sopra ed è stata testata l'introduzione di

[Torna all'indice](#)

informazioni linguistiche in questi sistemi. I nostri sistemi avanzati saranno pronti dal marzo 2013, quando inizierà una vera e propria valutazione, che continuerà per diversi mesi, con la collaborazione di sottotitolatori professionisti delle quattro aziende di sottotitolazione del consorzio. Un primo case study sulla coppia di lingue inglese e svedese è già stato completato utilizzando i sistemi standard, al fine di misurare l'iniziale guadagno di produttività e, soprattutto, di prendere nota dei fattori che influenzano il processo di post-editing in modo che siano trattati adeguatamente durante la valutazione su larga scala. Lo scopo di questa valutazione è, da una parte, individuare nell'output della traduzione automatica errori sistematici e ricorrenti che possono essere eliminati; dall'altra, raccogliere dati umani sull'adeguatezza e la fruibilità dell'output della traduzione automatica e dati sull'incremento di produttività nella fase di post-editing dell'output della traduzione automatica rispetto alla traduzione convenzionale dei file di sottotitoli.

Il sistema on-line che integra questo servizio sarà presentato al pubblico in occasione della conferenza *Media4ALL* di Dubrovnik nel settembre 2013, in cui un panel di utenti finali avrà la possibilità di testare il sistema SUMAT in tempo reale e di fornire così un feedback sulla sua usabilità e utilità.

Il progetto SUMAT (www.sumat-project.eu) è finanziato dall'EU ICT Policy Support Programme (2011-2014) nel quadro dell'accordo di sovvenzione n. 270919.

POSTFAZIONE: IL FUTURO DELLA TRASCRIZIONE

Luigi Zambelli

Il 12 settembre 2012, nella sala convegni di Palazzo Marini a Roma, si è svolta una Conferenza stampa nel corso della quale il direttivo dell'Associazione onA.I.R. ha presentato l'attività della stessa e, in particolare, il progetto "OltreSuoni".

Non entro nei dettagli dello stesso che sono stati ampiamente illustrati dal direttivo nei comunicati stampa e nelle comunicazioni diffuse in rete (<http://www.youtube.com/watch?v=E2ZyVSfw4io>), né mi soffermo sull'analisi degli altri ambiziosi obiettivi che il direttivo della neo costituita Associazione si è posta (vedi articolo 15). Vorrei soltanto evidenziare alcuni aspetti che mi hanno favorevolmente colpito e che ritengo dover sottolineare perché fortemente significativi e strategici, anche per le sinergie che si potranno attivare con l'Accademia "Giuseppe Aliprandi-Flaviano Rodriguez" e, più in generale, perché potranno permettere l'avvio di proficui rapporti con altri organismi, anche a livello internazionale come Intersteno.

Già dai primi documenti elaborati emergono il livello di preparazione, l'entusiasmo, le forti motivazioni e la progettualità condivisa che hanno portato a maturare nei giovani membri dell'Associazione la consapevolezza che, allorquando vi è una chiara consapevolezza degli obiettivi da perseguire, è possibile passare dalla fase della mera riflessione teorica a quella operativa e dare concretezza a un programma di lavoro organico e partecipato, sia negli strumenti, sia nelle strategie. Inoltre l'impianto teorico è ben strutturato – prevede una calendarizzazione degli impegni definita e la verifica dei risultati con delle valutazioni in itinere. Questi elementi devono costituire uno dei punti di forza e di credibilità di ogni intervento e fanno superare il rischio – come purtroppo spesso accade negli organismi assembleari – che i soggetti promotori di un'iniziativa, anche di rilevante spessore, disperdano energie preziose in sterili e defatiganti discussioni, talvolta fuorvianti, o in analisi astratte. E tutto questo perché spesso si fraintende e si sovraccarica di significato il principio attorno al quale ruota tutta l'attività associativa, cioè quello di dare a tutti la possibilità di un'ampia, consapevole e democratica partecipazione alle scelte comuni.

Ebbene, fin da questi primi approcci con la realtà esterna, onA.I.R. ha saputo dimostrare concretamente che, allorquando tra gli obiettivi di un progetto vi è anche quello dell'apertura alle istanze del sociale unitamente alla determinazione di cogliere e fare propri alcuni bisogni che la società odierna non sempre chiaramente sa esprimere, il percorso intrapreso sarà contrassegnato da inevitabili ostacoli che, tuttavia, possono essere superati. La sfida diventa particolarmente impegnativa quando si tenta di dare risposte a bisogni ai quali le strutture pubbliche non sanno dare risposte qualificate, professionali e tempestive. Ma, proprio per l'impegno richiesto, tale progetto acquista maggiore

credibilità verso l'esterno, si fa inoltre catalizzatore di nuove energie e, nel caso specifico, di adesione di nuovi soci: di questa capacità di aggregazione e di stimolo ne sono tangibile testimonianza le numerose domande di ammissione a soci che in questi mesi sono pervenute all'organo direttivo. Ciò non solo da parte di giovani neolaureati, ma anche da trascrittori, stenotipisti e altri professionisti della comunicazione scritta come traduttori e interpreti che in onA.I.R. ritengono di poter trovare un riferimento preciso per affrontare temi di interesse comune. Tutto questo con l'obiettivo di perseguire un progressivo arricchimento collettivo, superando un muro di indifferenza se non, a volte, di ostilità che ha impedito di affrontare le problematiche comuni, tra le quali l'obiettivo di raggiungere un vero riconoscimento – e non solo in termini economici – della loro professionalità. onA.I.R., con la rete di soci diffusa su tutto il territorio nazionale e con i rapporti che sta instaurando con il mondo accademico e professionale e che sicuramente saprà consolidare già nel breve periodo, ha quindi tutti i requisiti per poter diventare un qualificato soggetto *super partes*, in grado di valorizzare le multiformi professionalità che fanno riferimento all'elaborazione del parlato, indipendentemente dagli strumenti e dalle tecniche utilizzate.

Purtroppo, come spesso accade, anche per onA.I.R. si sono fatte sentire le prime critiche. Com'è stato messo in evidenza in una delle relazioni presentate nel corso della conferenza da Francesco Cellini, già nella fase propedeutica all'avvio "sul campo" del progetto "OltreSuoni" e, più in generale, all'attività dell'Associazione, non sono mancati giudizi connotati di scetticismo verso questa "ventata di aria nuova" che, appunto faticosamente, cerca di permeare il mondo dell'elaborazione del linguaggio. Si tratta di giudizi aprioristicamente negativi, frutto della mancanza di conoscenza e, spesso, della malafede di personaggi che si contraddistinguono per una visione parcellizzata della professione che non è più accettabile; questi tentativi di accendere o riaccendere sterili polemiche sulla presunta superiorità di uno strumento di lavoro o di una tecnologia rispetto ad un'altra ignorano una realtà e un principio generale sul quale tutti ormai concordano: oggi, e ancor più in futuro, non saranno soltanto gli strumenti a essere sottoposti a un processo di costante e frenetica evoluzione e conseguentemente di altrettanto rapida obsolescenza, ma sarà proprio, oltre alle competenze e alla conoscenza, l'atteggiamento degli operatori, la loro *forma mentis* che dovrà entrare in simbiosi con l'evoluzione tecnologica. Ogni professionista deve trovare quella malleabilità e duttilità che gli consente di adattarsi a una professionalità sottoposta a continue sollecitazioni e stimoli e che si arricchisce di nuovi contenuti. Quella della presunta superiorità di uno strumento rispetto a un altro genera unicamente sterili polemiche dannose per tutti, che tuttavia non spaventano questi giovani che sono ben consapevoli delle sfide con cui dovranno affrontarsi e, anzi, da tali critiche deriveranno nuovi stimoli.

[Torna all'indice](#)

Anche in quest'occasione – come anche nel dibattito che si è sviluppato in queste ultime settimane in rete – ho colto l'attualità di quanto hanno sempre sostenuto e dimostrato nell'attività professionale autorevoli colleghi tra i quali non posso non citare Gian Paolo Trivulzio e cioè che non è unicamente lo strumento, la tecnologia, che dà spessore e qualità al lavoro dei resocontisti, ma la preparazione culturale. La stessa è l'elemento cardine, fondante, che caratterizza e dà un forte valore aggiunto a ogni prestazione; la cultura coniugata con l'abilità nell'uso intelligente e consapevole delle tecnologie, dei più sofisticati software, permetterà di ottenere dalle stesse i migliori risultati. È questo un principio trasversale che “contamina positivamente”, in senso lato, qualsiasi attività umana. In particolare, il professionista della documentazione scritta, che con il suo lavoro assume il ruolo di fedele “mediatore” tra un soggetto (la fonte della comunicazione) e un altro a cui l'informazione è destinata, deve essere consapevole che la ripresa e la trasposizione del parlato nella forma scritta richiede conoscenze trasversali, forte motivazione e tensione all'autoaggiornamento, alla formazione continua o, per usare un termine forse più attuale, al lifelong learning.

Era tangibile, nel direttivo di onA.I.R. e nei soci presenti all'assemblea svoltasi nella mattinata presso l'Università Roma Tre, lo stesso entusiasmo che contraddistingueva quanti in passato, in ambito stenografico, si batterono con forza per superare le diatribe tra i sistemi e le “correnti di pensiero” e le sottostanti aree di interesse economico; sono state proprio queste divisioni che hanno impedito, in anni non molto lontani, di sviluppare progetti di ampio respiro orientati al reale progresso e alla ricerca.

Al contrario, la determinazione di onA.I.R. nel perseguire gli obiettivi dell'Associazione è confermata dalla rete di contatti che già il direttivo ha attivato, sia in ambito accademico e della ricerca applicata sia tra i referenti del mondo economico, sociale e anche politico.

All'iniziativa hanno dato il loro autorevole appoggio esponenti politici, come l'on. Marco Beltrandi e l'europarlamentare Mario Mauro, che con un diretto impegno hanno contribuito a licenziare in sede parlamentare provvedimenti legislativi che potranno costituire un preciso punto di riferimento per le future azioni che onA.I.R. ha in animo di attivare sul territorio nazionale. In ambito universitario si è aperto un nuovo orizzonte per quanti operano nell'area dell'elaborazione elettronica del parlato, ovvero del respeaking o anche della stenotipia. Ne sono una dimostrazione concreta i progetti di eccellenza attivati presso alcune Università, tra cui Roma Tre e Parma, che hanno attivato dei corsi e dei Master post laurea inserendo l'insegnamento della sottotitolazione in differita e in tempo reale. Esperienze queste che trasferiscono sul piano della concretezza e in un ambito privilegiato il risultato degli studi e delle ricerche iniziate già nel lontano 2004 da Carlo Eugeni presso l'Università Federico II di Napoli. Il contatto e la sensibilizzazione del mondo accademico era un auspicio

[Torna all'indice](#)

con il quale gli insegnanti di stenografia e dattilografia – più sensibili ai nuovi scenari che si stavano prospettando già negli anni '70 – chiudevano i loro convegni e incontri di studio.

Un'ultima considerazione: sempre con riferimento alle indagini e agli studi in ambito accademico sempre più tesi di laurea indagano le problematiche connesse alla documentazione scritta focalizzando l'indagine sull'attività dei resocontisti in ambito giudiziario e assembleare. In tali lavori si auspicano specifici interventi di formazione del personale addetto alla trascrizione in genere con l'obiettivo di consentire allo stesso di fornire prestazioni di elevato standard qualitativo. Tutti questi approfondimenti in ambito accademico dimostrano che si è aperta una nuova fase che contribuirà efficacemente allo sviluppo delle tecniche di ripresa del parlato: riprendono nel contempo, con rinnovato vigore, scientificità e autorevolezza, le ricerche e gli studi in ambito grafico (e stenografico) che sembravano invece destinati a rimanere patrimonio del passato. Con soddisfazione prendiamo atto che oggi giovani studiosi, come loro, si stanno impegnando sulla strada della ricerca e degli studi grafici tracciata dai grandi Maestri, dagli autori dei sistemi stenografici agli studiosi come Giulietti, Aliprandi, Rodriguez. A tutti loro auguro il migliore successo.

BIBLIOGRAFIA

Andreou A., Kamm T. e Cohen J. (1994), "Experiments in vocal tract normalization". Proceedings CAIP/Rutgers Workshop: Frontiers in Speech Recognition II.

Arumí Ribas M. e Romero Fresco P. (2008), "A Practical Proposal for the Training of Respeakers". JoSTrans (The Journal of Specialized Translation), Issue 10: Articles.

<http://www.jostrans.org/archive.php?display=10>

Bearing 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 on Respeaking.

http://www.intralinea.it/specials/respeaking/ita_more.php?id=446_0_41_0_M

Baker R. (1981), *ORACLE Subtitling for the Deaf and Hard-of-hearing*. Southampton: Southampton University.

Baker R. et al. (1984), *Handbook for Television Subtitlers*. London: Independent Broadcasting Authority. Engineering division.

Barik H. C. (1973), "Simultaneous interpretation: temporal and quantitative data". In *Language and Speech n. 16*, pp. 237-270.

Barthes R. (1970), *S/Z*. Parigi: Seuil.

Bartoll E. (2004), "Parameters for the classification of subtitles". In Orero P. (a cura di), *Topics in Audiovisual Translation*. Amsterdam and Philadelphia: John Benjamins.

Bassnett-McGuire S. e Lefevere A. a cura di (1990), *Translation, History and Culture*. Londra e New York: Pinter Publishers.

Baum L., Petri E., Soules T., Weiss G. e N. (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 (1994), *The BBC subtitling guide*. Londra: BBC.

Bell A. (1984), "Language style as audience design". In *Language in society 13*.

Bell A. e Garrett P. a cura di (1998), *Approaches to Media Discourse*. Oxford: Blackwell.

Bellabarba V. (2010), *Il rispeakeraggio della telecronaca sportiva. Analisi delle strategie linguistiche adottate da Canal+*. Tesi di Laurea non pubblicata. Università degli Studi di Macerata.

Bernero R. e Bothwell H. (1996), *Relationships of hearing impairment to educational needs*. Springfield: Illinois DPHI & Office SPI.

Bhatia V. K. (1993), *Analysis Genre. Language Use in Professional Settings*. Londra: 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 and 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.

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, pp.167-172.

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.

Boulianne, Gilles Beaumont J. F., Boisvert M., Brousseau J., Cardinal P., Chapdelaine C., Comeau M., Ouellet P., Osterrath F. e P. Dumouchel (2009), "Shadow speaking for real-time closed-captioning of TV broadcasts in French". In Matamala A. e Orero P. (a cura di), *Listening to Subtitles. Subtitles for the Deaf and Hard-of-Hearing*. Berna: Peter Lang.

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. Manchester: University of Manchester.

Brette N. (1982), "Sous-titres, le crève-cœur des réalisateurs". In *Cahiers du cinéma* 338.

Bristow N. (1987), *Teletext sub-titles. A preliminary exploration*. Londra: BBC Broadcasting research.

Brunetto F. (2005), *I sottotitoli per non udenti: considerazioni generali e applicazioni pratiche in Italia, Regno Unito e Belgio*. Tesi di laurea. Forlì: Università di Bologna.

Cameron D. (2001), *Working with spoken discourse*. Londra: Sage.

Captionmax (2002), *Captioning styles*. www.captionmax.com/pages/ViewerInfo/VI_style.html.

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.

Chaume F. (2002), "Models of research in audiovisual translation". In *Babel* 48 (1).

Chiaro D. (2006), "Verbally Expressed Humor on Screen: translation, reception and perception". In Diaz-Cintas J., Orero P. e Ramael A. (a cura di), *JoSTrans* 6: 1.

Clark M. D. et al. a cura di (2001), *Context, cognition and deafness*. Washington : Gallaudet university press.

Colby (2011) *Company*. colby: Company.

<http://www.colby.it/pagine.php?p=content&txt=company&lang=it>

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. e Radkte E. (a cura di), *Gesprochen Italienisch in Geschichte und Gegenwart*. Tübingen: Universitat Tübingen.

Cosi P. e E. Magno Caldognetto (1996), "Lips and Jaw Movements for Vowels and Consonants: Spatio-Temporal Characteristics and Bimodal Recognition Applications". In Starke D. G. e Henneke M. E., *Speech reading by Humans and Machine: Models, Systems and Applications*. NATO ASI, vol. 150, Springer-Verlag.

Cox R.V. *et al.* (2000), "Speech and language processing for next-millennium communication services". In *Proceedings IEEE*, vol. 88.

CSA (2011a) "Méthodologie de constatation de la reprise des sous-titres à destination des personnes sourdes ou malentendantes". In CSA, *Brochures d'information*. csa: publications du CSA, télévision. <http://www.csa.fr/Media/Files/Television/Le-suivi-des-programmes/Methodologie-de-constatation-de-la-reprise-des-sous-titres-a-destination-des-personnes-sourdes-ou-malentendantes>

CSA (2011b) "Pour les personnes sourdes ou malentendantes : le sous-titrage". In CSA, *L'accessibilité des programmes*. csa: publications du CSA, télévision. <http://www.csa.fr/Television/Le-suivi-des-programmes/L-accessibilite-des-programmes/Pour-les-personnes-sourdes-ou-malentendantes-le-sous-titrage>

Davis K. H., Biddulph R. e S. Balashek (1952), "Automatic recognition of spoken digits". In *Journal of the Acoustical Society of America*, vol. 24, n° 6.

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 on Respeaking.

http://www.intralinea.it/specials/respeaking/ita_more.php?id=454_0_41_0_M

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 média audiovisuels*. Parigi : Presses Universitaires du Septentrion.

De Seriis L. (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.

Desmedt (2002), *Le sous-titrage pour sourds et malentendants*. Tesi di DESS. Bruxelles: ISTI-HEB.

de Sousa S., Aziz W. E L. Specia (2011), "Torna all'indice" t-Editing Effort for Automatic and

[Torna all'indice](#)

Semi-Automatic Translations of DVD Subtitles”. In *Proceedings of Recent Advances in Natural Language Processing Conference (RANLP-2011)*, Hissar, Bulgaria.

Díaz-Cintas J. (2001), *La traducción audiovisual. El subtitulado*. Salamanca: Almar.

Díaz-Cintas J. (2003), “Audiovisual translation in the third millennium”. In Andermann G. e Rogers M. (a cura di), *Translation today. Trends and perspectives*. Clevedon: multilingual matters, pp. 192-204.

Díaz-Cintas J., Orero P. e A. Remael a cura di (2007), *Media for all: subtitling for the deaf, audio description, and sign language*. Amsterdam: Rodopi.

Di Silvio M. (2006), “Le necessità dei sordi: aspetti tecnici”. In Eugeni C. e G. Mack (a cura di), *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking.

http://www.intralea.it/specials/respeaking/ita_more.php?id=482_0_41_0_M

Durante M. M. (2005), *La televisione e la sordità: un’analisi della sottotitolazione per non udenti delle emittenti televisive italiane*. Tesi di laurea. Bruxelles: HEB ISTI.

D’Ydewalle G. et al. (1991), *Watching subtitled television: automatic reading behaviour*. *Communication Research* 18:5, October 1991, p.650-666.

EBU (2004), *EBU Report on Access Services*. European Broadcasting Union Committee - EBU Technical-Information R-144-2004.

Eco U. (2003), *Dire quasi la stessa cosa*. Milano: Bompiani.

ESIST (European Association for Studies in Screen Translation), Code of Good Subtitle Practice: <http://www.esist.org/Code.pdf>.

Eugeni C. (2006a), “For a didactics of respeaking”. Comunicazione alla conferenza internazionale Languages and the Media 2006. Berlino: Hotel Intercontinentale.

Eugeni C. (2006b), “Introduzione al rispeaking televisivo”. In C. Eugeni e G. Mack (a cura di) *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking.

http://www.intralea.it/specials/respeaking/ita_more.php?id=444_0_41_0_M

Eugeni C. e G. Mack a cura di (2006), *Intralea, Special Issue on New Technologies in Real Time Intralingual Subtitling*. http://www.intralea.it/specials/respeaking/eng_open.php [01/01/2013]

Eugeni C. (2007), “Il rispeaking televisivo per sordi. Per una sottotitolazione mirata del TG”. inTRAlinea, Vol. 9. http://www.intralea.it/volumes/eng_open.php?id=P513

Eugeni C. (2008a), “A Sociolinguistic Approach to Real-time Subtitling: Respeaking vs. Shadowing and Simultaneous Interpreting”. In Cynthia J., Bidoli K. e Ochse E. (a cura di), *English in International Deaf Communication*, Linguistic Insights series, Vol. 72, Bern, Peter Lang, pp. 357-382.

[Torna all'indice](#)

Eugeni C. (2008b), *Le sous-titrage en direct pour sourds et malentendants: aspects théoriques, professionnels et didactiques du respeakerage télévisue.*, Macerata, CEUM.

Eugeni C. (2008c), "Respeaking the TV for the Deaf: For a Real Special Needs-Oriented Subtitling". In *Studies in English Language and Literature*, n. 21, pp. 37-47.

Eugeni C. (2009a), *La sottotitolazione in diretta TV. Analisi strategica del respeakeraggio verbatim di BBC News*, Tesi di Dottorato non pubblicata, Università Degli Studi Di Napoli Federico II.

Eugeni C. (2009b), "Respeaking the BBC News. A strategic analysis of Respeaking on the BBC". In *The Sign Language Translator and Interpreter (SLTI)*, Vol. 3, n. 1, Manchester, St. Jerome Publishing, pp. 29-68.

Evans M. J. (2003), "Speech Recognition in Assisted and Live Subtitling". In *BBC Research and Development White Papers*. www.bbc.co.uk.

FCC (1997), *Captioning*. Federal communications commission. www.fcc.gov.

Ferguson J. a cura di (1980), *Hidden Markov Models for Speech*. Princeton: IDA.

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, n° 11.

Fry D. B. (1959), "Theoretical aspects of the mechanical speech recognition". In *Journal of the British Institute of Radio Engineers*. Vol. 19, n° 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". *Proceedings ESCA-NATO Workshop on Robust Speech Recognition for Unknown Communication Channels*. Pont-a-Mouson, France , pp. 11–20.

Galiano N., Refice, M. e M. Savino (2004), "Il riconoscimento bimodale del parlato: un esempio di implementazione per l'italiano". In *Proceedings of AISV 2004*. Padova.

Gambier Y. (1992), "La reformulation – pratique linguistique et interlinguistique". In *KOINE. 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 Benjami

[Torna all'indice](#)

- 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. Parigi: FIT.
- Gambier Y. (2003), "Screen Transadaptation. Perception and Reception". In *The translator* 9 (2).
- Gambier Y. (2004), La traduction audiovisuelle: un genre en expansion. *Meta*, 49(1), pp. 1-11.
- Gambier Y. (2005a), "Screen translation: subtitling". In *The Encyclopedia of Languages and Linguistics* (2a edizione), Oxford: Elsevier.
- Gambier Y. (2005b), "Orientation de la recherche en traduction audiovisuelle". In *Target* 17 (1).
- Gambier Y. (2006a), "Le sous-titrage: une traduction sélective? ". In Tommola J. e Gambier Y. (a cura di), *Translation and interpreting. Training and research*. Turku : university of Turku, pp. 21-37.
- Gambier Y. (2006b), Discorso di chiusura della conferenza internazionale Languages and the Media 2006. Berlino: Hotel Intercontinental.
- Gile D. (1995), *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam, Benjamins.
- Gile D. (2005), *La traduction. La comprendre, l'apprendre*. Parigi, Presses Universitaires de France.
- 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.
- 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: John Benjamins.
- Gottlieb H. (1994), "Subtitling: Diagonal Translation". In *Perspectives: Studies in Translatology*, vol. 2, n. 1, Routledge, pp. 101-121.
- Gottlieb H. (1997), *Subtitles, translation and idioms*. Tesi di dottorato. Copenaghen: university of Copenaghen.
- Gran L. (1992), *Aspetti dell'organizzazione cerebrale del linguaggio: dal monolinguisma all'interpretazione simultanea*. Udine: Campanotto.

[Torna all'indice](#)

- Gran L. (1999), "L'interpretazione simultanea: premesse di neurolinguistica". In Falbo C., Russo M. e 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". *ESCA Workshop on Auditory Basis of Speech Perception*, pp. 1-8, Keele, UK.
- Halliday M. A. K. e R. Hasan (1985), *Language, Context and Text: a Social Semiotic Perspective*. Oxford, Oxford University Press.
- Hardmeier C. e M. Volk (2009), "Using Linguistic Annotations in Statistical Machine Translation of Film Subtitles". In: *Proceedings of 17th Nordic Conference on Computational Linguistics (Nodalida)*. Odense, Danimarca, pp. 57-64.
- Haton J.-P. (2006), *Reconnaissance automatique de la parole: du signal à son interprétation*. Parigi : Dunod.
- Hermansky H. (1990), "Perceptual linear predictive (PLP) analysis of speech". In *Journal of the Acoustical Society of America*, vol. 87, n° 4.
- Hermansky H. e N. Morgan (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 seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking.
http://www.intralinea.it/specials/respeaking/eng_more.php?id=480_0_41_0_M
- Hillier H. (2004), *Analysing Real Texts*. Basingstoke: Palgrave Macmillan.
- Hindmarsh R. (1985), *Language problems in European television. A feasibility study*. Manchester: European Institute for the Media.
- Hymes, D. H. (1972) "On Communicative Competence" In J.B. Pride and J. Holmes (a cura di) *Sociolinguistics. Selected Readings*. Harmondsworth: Penguin.
- Itakura F. (1975), "Minimum prediction residual applied to speech recognition". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. ASSP-23.
- ITC (1999), *ITC Guidance on standards for subtitling*. Londra: ITC, www.itc.org.uk.
- ITC (2001), *ITC code on subtitling, dign 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*. Stoccolma: Transedit.
- Jääskeläinen R. (a cura di), *Tapping and mapping the processes of translation and interpretation*. Amsterdam e Philadelphia: John Benjamin

[Torna all'indice](#)

Jakobson R. (1959), "Aspetti linguistici della traduzione". In Nergaard S. a cura di (1995), *Teorie contemporanee della traduzione*, Milano, Bompiani, pp. 51-62.

Jelinek F. (1997), *Statistical methods for speeche recognition*. Cambridge, MA: MIT Press.

Jelinek F. (1985), "The development of an experimental discrete dictation recognizer". In *Proceedings IEEE*, vol. 73, novembre 1985.

Jelinek F., Bahl L.R. e L. R. Mercer (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, n° 4.

Jensema C. (1999), *Caption speed and viewer comprehension of television programs final report*. Washington: US department of education.

Jensema C., McCann R. e S. Ramsey (1996), "Closed captioned television, presentation speed and vocabulary". In *American Annals of the Deaf*. Volume 144, n° 4.

Juang B. H. (1985), "Maximum likelihood estimation for mixture multivariate stochastic observations of Markov chains". *AT&T Technology Journal*, vol. 64, Morristown: Association for Computational Linguistics.

Juang B. H. (1996), "Automatic speech recognition: Problems, progress & prospects". In *IEEE Workshop Neural Networks for Signal Processing*, Kyoto, Japan.

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, pp. 257-265.

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, pp. 1142-1165.

Junqua J.-C. e J.-P. Haton (1996), *Robustness in Automatic Speech Recognition*. Boston, MA: Kluwer.

Karamitroglou F. (1998), "A proposed set of subtitling standards in Europe". In *Translation journal* 2 (2).

Kawahara T., Lee C.H. e Juang B.H. (1997), "Combining key-phrase detection and subword based verification for flexible speech understanding". In *Proceedings IEEE ICASSP97*.

Kintsch W. e T. A. van Dijk (1978), "Toward a Model of Text Comprehension and Production". In *Psychological Review* 85.

Koehn Ph. (2010), *Statistical Machine Translation*. Cambridge University Press, Cambridge, United Kingdom.

[Torna all'indice](#)

- Kohn K. e S. Kalina (1996), "The strategic dimension of interpreting". In *Meta*, 41 (1).
- Kovačič I. (1995), "Subtitling strategies: a flexible hierarchy of priorities". In Heiss C. e R. M. Bollettieri Bosinelli a cura di (1996), *Traduzione multimediale per il cinema, la televisione e la scena*. Bologna: CLUEB.
- Kovačič I. (1995), "Reception of subtitles. The non-existent ideal viewer". In *Translatio* XIV (3-4).
- Kutz W. (1997), "Compression of the source message during simultaneous interpretation". In Ministero per i beni culturali e ambientali. *La traduzione. Saggi e documenti III. Quaderni di libri e riviste d'Italia* 33. Roma: Ministero per i beni culturali e ambientali.
- Kyle J. (1996), "Switched on: deaf people's view on television subtitling previous reports". In www.deafstudiestrust.ork.U.K.
- Lambert S. (1989), "La formation d'interprètes: la méthode cognitive". *Meta* 34 (4), pp. 736-744.
- Lambert S. (1992a), "Shadowing". *The Interpreters' Newsletter* 4, pp. 15-24.
- Lambert S. (1992b), "The Cloze Technique as a Pedagogical Tool for the Training of Translators and Interpreters". *Target* 4:2. Amsterdam: John Benjamins, pp. 223-236.
- 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.
- Lambourne A., Hewitt J., Lyon C. e S. Warren (2004), *Speech-Based Real-Time Subtitling Services*. *International Journal of Speech Technology*, vol.7 (4), pp. 269-279.
- Lambourne A. (2006), "Subtitle respeaking". In Eugeni C. e G. Mack (a cura di), *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking. http://www.intralinea.it/specials/respeaking/eng_more.php?id=484_0_41_0_M
- Lambourne A. (2007), "Real-time subtitling: extreme audiovisual translation". Intervento alla conferenza internazionale *LSP Translation Scenarios*. Vienna.
- Lavour J.M. e A. Şerban, a cura di (2008), *La traduction audiovisuelle. Approche interdisciplinaire du sous-titrage*. Bruxelles, Groupe De Boeck.
- 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.
- Lepot-Froment C. (1986), *Vivre sourd – Communication et surdit e aujourd'hui... et demain?* Bruxelles : EDIRSA.

[Torna all'indice](#)

- 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.
- Lo Presti L. (2007), *La sottotitolazione per sordi: linee guida ragionate*. Tesi di laurea. Forlì: università di Bologna.
- Luyckx Bieke, Delbeke, Van Waes, Leijten e Remael (2010), "Live Subtitling with Speech Recognition: Causes and Consequences of Text Reduction". Anversa: Artesis working Papers in Translation Studies 2010-1.
http://www.artesis.be/vertalertolk/upload/docs/onderzoek/Artesis_VT_working_paper_2010-1_Luyckx_et_alii.pdf.
- Luyken G.-M., Herbst T., Langham-Brown J., Reid H. e H. Spinhof (1991), *Overcoming language barriers in television. Dubbing and subtitling for the European audience*. Manchester: the European Institute for the Media.
- Mack G. (2002), "New perspectives and challenges for interpretation: the example of television". In Garzone, G. e Viezzi, M. (a cura di) *Interpreting in the 21st century – Challenges and Opportunities*. Amsterdam e Philadelphia: John Benjamins.
- Mack G. (2006), "Detto scritto: un fenomeno, tanti nomi". In Eugeni C. e G. Mack (a cura di), *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking.
http://www.intralinea.it/specials/respeaking/eng_more.php?id=464_0_41_0_M
- Maragna S. (2000), *La sordità: educazione, scuola, lavoro e integrazione sociale*. Milano: Hoepli.
- Marchionne F. (2011), *Il respeaking interlinguistico. Sperimentazioni per una evoluzione della sottotitolazione in diretta*. Tesi di laurea magistrale non pubblicata. Università degli Studi di Macerata.
- Margaret A., "Audiolesi: diktat nelle scuole?" In *Salute. La Repubblica*, 9 novembre 2006.
- 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), pp. 271-285.
- Marsh A. (2004), *Simultaneous interpreting and respeaking: a comparison*. Tesi di laurea magistrale non pubblicata. Londra: University of Westminster.
- Marsh A. (2006), "Respeaking for the BBC". In C. Eugeni e G. Mack (a cura di) *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking. http://www.intralinea.it/specials/respeaking/eng_more.php?id=484_0_41_0_M
- Martin T. B., Nelson A. e H. J. L. Zadell (1964), *Speech recognition by feature abstraction techniques*. Air Force Avionics Lab, Tech [Torna all'indice](#) 176.

- Martínez-Pérez J. (2012), "Applying a Punctuation-based Segmentation to a new Add-on Display Mode of Respoken Subtitles". In Remael A., Orero P. e Carroll M. (a cura di), *Audiovisual Translation and Media Accessibility at the Crossroads. Media for All 3*, Amsterdam and New York: Rodopi, pp. 133-147.
- Mason I. (1989), "Speaker meaning and reader meaning: preserving coherence in screen translating". In Kölner R. e Payne J. (a cura di), *Babel: the cultural and linguistic barriers between nations*. Aberdeen: Aberdeen university press.
- Massaro D. W. (1972), "Perceptual images, processing time and perceptual units in auditory perception". *Psychological Review*, 79 (2): 124-145.
- McGurk H. e J. McDonald (1976), "Hearing lips and seeing voices". *Nature* 264 (5588), London, Pubmed, pp. 746-748.
- Mereghetti E. (2006), "Le necessità dei sordi: TV e vita quotidiana". In Eugeni C. e Mack G. (a cura di), *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking.
http://www.intralinea.it/specials/respeaking/ita_more.php?id=481_0_41_0_M
- Miller L. G. e A. Gorin (1993), "Structured networks for adaptive language acquisition". *International Journal of the Pattern Recognition and Artificial Intelligence* (Special Issue on Neural Networks), vol. 7, no. 4, pp. 873–898.
- Minchinton J. (1993), *Sub-titling*. Hertfordshire: Minchinton J.
- Moore R. C. (1997), "Using natural-language knowledge sources in speech recognition". In Ponting K., *Computational Models of Speech Pattern Processing*. Ed. Berlin, Germany: Springer-Verlag, pp. 304–327.
- Muzii L. (2006), "Respeaking e localizzazione". In Eugeni C. e G. Mack (a cura di), *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking. http://www.intralinea.it/specials/respeaking/ita_more.php?id=449_0_41_0_M
- Myklebust H.R. (1964), *The psychology of deafness*. New York: Grune and Stetton.
- Nadas A., Nahamoo D. e M. A. Picheny (1988), "On a model-robust training method for speech recognition". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. 36, pp. 1432–1436.
- 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.
- Nencioni G. (1983), *Di scritto e di parlato, discorsi linguistici*. Bologna: Zanichelli.
- Neves J. (2004) "Subtitling: written interpretation?" In *Génesis – Revista de tradução do ISAI 4*.
- Neves J. (2005), *Audiovisual translation. Subtitling for the Deaf and Hard-of-Hearing*. Roehampton University, University of Surrey.

[Torna all'indice](#)

Ney H. e S. Ortmanns (2000), "Progress in dynamic programming search for LVCSR". In *Proceedings IEEE*, vol.88.

Nothlev P. (2008), "Respeaking for a Danish broadcaster". Intervento al *Second International Seminar on Real-Time Intralingual Subtitling*. <http://www.respeaking.net/programme/nothlev.pdf>

O'Connel D. e S. Kowal (1994), "Some current transcription systems for spoken discourse: a critical analysis". In *Pragmatics 4*.

OFCOM, Office of Communication, linee guida

http://www.ofcom.org.uk/tv/ifi/guidance/tv_access_serv/guidelines/#top.

Orero P. (2004), *Audiovisual translation: a new dynamic umbrella*. In Orero P. (a cura di), *Topics in audiovisual translation*. Amsterdam e Philadelphia: John Benjamins.

Orero P. (2006), "Real-time subtitling in Spain". 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.

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. Università di Napoli Federico II.

Pieraccini R. e E. Lavin (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*. Roma: Armando.

Porcari Li Destri G. e V. Volterra a cura di (1995), *Passato e presente: uno sguardo sull'educazione dei sordi in Italia*. Napoli: Gnocchi.

Praet C. et al. (1990), "A one line text is not half a two line text". In Groner R. et al. (a cura di), *From eye to mind: information acquisition in perception, search and reading*. North-Holland: Elsevier science publishers BV.

Prandi M. (2004), "Riformulazione e condivisione". In *Rassegna italiana di linguistica applicata*, n.1. Roma: 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". *Proceedings IEEE*, vol. 77, pp. 257–286.

[Torna all'indice](#)

Rabiner L. R. e B. H. Juang (1993), *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice-Hall.

RAI (2011a), "Capitolato Tecnico per la fornitura del servizio di sottotitolazione di programmi televisivi preregistrati e in diretta per non udenti, per Televideo". Allegato 3 al Disciplinare di Gara per il Bando sulla *Fornitura del servizio di sottotitolazione di programmi televisivi preregistrati e in diretta per non udenti, per Televideo*, Roma, Bando di Gara.

RAI (2011b), "Livelli di servizio previsti per la sottotitolazione dei programmi televisivi per Televideo". Appendice 4 al Capitolato Tecnico per il Bando sulla *Fornitura del servizio di sottotitolazione di programmi televisivi preregistrati e in diretta per non udenti, per Televideo*, Roma, Bando di Gara.

RAI (2011c), "Regole di composizione dei sottotitoli per i non udenti". Appendice 3 al Capitolato Tecnico per il Bando sulla *Fornitura del servizio di sottotitolazione di programmi televisivi preregistrati e in diretta per non udenti, per Televideo*, Roma, Bando di Gara.

RAI (2011d), "STLive. Specifiche tecniche per contribuzioni sottotitoli "live" effettuate da terze parti". Appendice 2 al Capitolato Tecnico per il Bando sulla *Fornitura del servizio di sottotitolazione di programmi televisivi preregistrati e in diretta per non udenti, per Televideo*, Roma, Bando di Gara.

RAI e Ministero dello Sviluppo Economico (2011) "Contratto di Servizio 2010-2012". In G. De Cesare (a cura di) *Leggi e informazione radiotelevisiva*. USIGRai (Unione Sindacale Giornalisti Rai). http://www.decesare.info/contratto_servizio_20092012.htm

Rampelli S. (1990), *Aspetti linguistici dei sottotitoli e fruibilità da parte delle persone sorde. Prospettive teoriche e ambiti applicativi*. Tesi di laurea. Università "La Sapienza" di Roma.

RBM France (2011) *Travailler pour Red Bee Media France*. redbeemedia: emploi. <http://www.redbeemedia.fr/paris-jobs.html>

Redattore speciale (2011) ENS: "LA RAI TAGLIA SUI SOTTOTITOLI PER I SORDI". <http://www.disabili.com/home/ultimora/22977-ens-gla-rai-taglia-sui-sottotitoli-per-i-sordiq>

Remael A. e B. van der Veer (2006), "Real-time Subtitling in Flanders: Needs and Teaching". In Eugeni C. e G. Mack (a cura di), *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking. http://www.intralinea.it/specials/respeaking/ita_more.php?id=492_0_41_0_M

Remael A. e B. van der Veer (2007), "Teaching live-subtitling with speech recognition". Comunicazione alla conferenza internazionale Multidimensional Translation: LSP translation scenarios. Vienna: Centre for Translation Studies.

Remael A. (2007), "Sampling Subtitling for the Deaf and the Hard-of-Hearing in Europe". In Díaz Cintas J., Remael A. e P. Orero (a cura di), *Media for All*. Amsterdam: Rodopi, pp. 23-52.

Robson G. (2004), *The closed captioning handbook*. Oxford: Elsevier.

[Torna all'indice](#)

Romero-Fresco P. (2008), "La subtitulación rehablada: palabras que no se lleva el viento". In Pérez-Ugena A. e R. Vizcaíno-Laorga (a cura di), *ULISES: Hacia el desarrollo de tecnologías comunicativas para la igualdad de Oportunidades*. Madrid: Observatorio de las Realidades Sociales y de la Comunicación, pp. 49-73.

Romero-Fresco P. (2011), *Subtitling through Speech Recognition: Respeaking*. Manchester: St. Jerome.

Romero-Fresco P. (2012), "Quality in Live Subtitling: The Reception of Respoken Subtitles in the UK", in Remael A., Orero P. e Carroll M. (a cura di), *Audiovisual Translation and Media Accessibility at the Crossroads. Media for All 3*. Amsterdam: Rodopi, pp. 111-133.

Rubbi M. (2003), "Sottotitoli in Europa: i teleAspettatori". In *HP/Accaparlante n. 3*. Bologna: Centro Documentazione Handicap.

Russello C. (2009), *Respeaking e interpretazione simultanea: un'analisi comparata e un contributo sperimentale*. Tesi di Laurea Specialistica non pubblicata, Università di Roma LUSPioV.

Russello C. (2010), "Teaching respeaking to conference interpreters", Intersteno <http://www.intersteno.it/materiale/ComitScientifico/EducationCommittee/Russello2010Teaching%20Respeaking%20to%20Conference%20Interpreters.pdf>.

Sacks O. (1990), *Vedere voci. Un viaggio nel mondo dei sordi*. Milano: Adelphi.

Sakoe H. e S. Chiba (1978), "Dynamic programming algorithm optimization for spoken word recognition". In *IEEE Transactions Acoustics, Speech, Signal Processing*, vol. ASSP-26, pp. 43–49.

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*. Genova. <http://ia.di.uniba.it>

Shire M. L. (1997), *Syllable onsets detection from acoustics*. Tesi di master. UC Berkeley.

Silver J. et al. (1995), "Text display preferences on self-service terminals by visually disabled people". In *Optometry Today* 35 (2), pp. 24-27.

Silver J. (2000), *A new font for digital television subtitles*. www.tiresias.org.

ST'501 (2011) *ST'501*. [st501.fr: qui sommes-nous](http://www.st501.fr/index.php/qui-sommes-nous/st-501). <http://www.st501.fr/index.php/qui-sommes-nous/st-501>

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.

Tiedemann J. (2009), "News from OPUS - A Collection of Multi-lingual Parallel Corpora with Tools and Interfaces". In Nicolov N., Bontcheva K., Angelova G. e R. Mitkov (a cura di), *Recent Advances in Natural Language Processing (vol V)*. Amsterdam e Philadelphia: John Benjamins, pp. 237–248.

[Torna all'indice](#)

Titford C. (1982), "Sub-titling – constrained translation". In *Lebende Sprachen* 27 (3).

Toury, G. (1991), "What are descriptive studies into translation likely to yield apart from isolated descriptions?" In K. van Leuven-Zwart, e T. Naaijken (a cura di), *Translation Studies: The State of the Art*. Proceedings from the First James S. Holmes Symposium on Translation Studies. Amsterdam: Rodopi.

Trivulzio G. (2000), *Da Tirone al riconoscimento del parlato*. Milano: Asfor.

Trivulzio G. P. (2006), "Natura non facit saltus". In Eugeni C. e G. Mack (a cura di), *Proceedings of the first international seminar on real-time intralingual subtitling*. inTRAlinea, Special Issue on Respeaking. http://www.intralea.it/specials/respeaking/ita_more.php?id=451_0_41_0_M

Tucci L. (2000), "...a pagina 777 di televideo". In Taylor C. (a cura di), *Tradurre il cinema. Atti del convegno*. Trieste: Dipartimento di scienza del linguaggio dell'interpretazione e della traduzione, pp. 91-95.

University of Westminster, Master in Conference Interpreting.
<http://www.westminster.ac.uk/schools/humanities/postgraduate2/modern-and-applied-languages2/ma-conferencing-interpreting>.

Vázquez M. (2009), *El rehabilitado: un paso más hacia la accesibilidad universal*. Unpublished MA thesis. Vigo: Universidade de Vigo.

Vintsyuk T.K. (1968), "Speech discrimination by dynamic programming". In *Kibernetika*, vol.4.

Volk M. (2008), *The Automatic Translation of Film Subtitles. A Machine Translation Success Story?* In Nivre J., Dahllöf M. e B. Megyesi (a cura di), *Resourceful Language Technology: Festschrift in Honor of Anna Sægvall Hein*. Volume 7 of *Studia Linguistica Upsaliensia*, Uppsala, Sweden, pp. 202–214.

Volterra V. a cura di (1981), *I segni come parole: la comunicazione dei sordi*. Torino: Boringhieri.

Volterra V. (1986), *Il linguaggio dei sottotitoli e gli spettatori sordi*. Rapporto n. 5, Roma: RAI Televideo Roma.

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.

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. 38.

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, pp. 1870–1878.

[Torna all'indice](#)