

L'evoluzione tecnica: da Tirone al riconoscimento del parlato.

di Gian Paolo Trivulzio

I progressi realizzati negli ultimi due anni dai software di riconoscimento del parlato, unito alle maggiori velocità di elaborazione dei computer e di memorie di alta capacità, fanno intravedere la possibilità di utilizzare questo sistema di scrittura quale valida alternativa alla scrittura stenografica o stenotipica.

Prove pratiche hanno dimostrato che i programmi riescono oggi a trasformare le parole in segni, entro un limite molto limitato di tempo da quanto esse vengono pronunciate e che tale attività può svolgersi con durata nel tempo (le mie prove sono finora giunte ad un'ora e mezzo) senza particolari problematiche tecniche.

Diventa quindi ora possibile seguire la parola di un oratore, dettare il discorso al computer ed ottenere il testo scritto. La percentuale di fedeltà, con operatore opportunamente addestrato, raggiunge livelli superiori al 96 % in prima stesura.

Questi risultati, facilmente dimostrabili (come testimoniano le presentazioni fatte in Accademia a suo tempo dal prof. Galotta e successivamente dal Dottor Fabi) vengono facilitati dall'utilizzo di un microfono racchiuso all'interno di una maschera, denominata appunto steno-mask. Questa maschera è nata nel dopoguerra per risolvere il problema pratico di avere più controllori di volo nella stessa stanza che davano istruzioni agli aerei. Per evitare appunto interferenze fra di essi, venne realizzata dai militari una maschera di plastica. Essa ha trovato in America un'applicazione particolare: la resocontazione giudiziaria su nastro, il resocontista dettava con questa maschera il suo resoconto. La maschera isola dai rumori esterni e quindi la registrazione non è inficiata da informazioni spurie, come pure il resocontista che parla all'interno di questo strumento non disturba lo svolgimento dei lavori. Ebbi occasione di vedere questo strumento nei primi anni '60 su una rivista di organizzazione d'ufficio (L'Ufficio Moderno) che riportava una fotografia di una segretaria che con questo sistema prendeva appunti durante una riunione. Giudicai il sistema un po' macchinoso (anche perché allora i registratori erano già di per sé stessi macchinosi), ma comunque ingegnoso.

Il riconoscimento del parlato ha rivalutato questo strumento ed è nata la mini-mask, in pratica una versione meno 'militare', con forma e dimensioni più accattivanti, particolarmente adatta per il gentil sesso (anche se i maschietti la utilizzano correntemente).

Nel 2000 ho importato due di questi strumenti (uno tradizionale per me, ed uno mini per la mia collega) ed ho effettuato prove di messa a punto della tecnica di dettatura con questa steno-mask. Con essa mi sono anche presentato alle gare svizzere di stenografia, ed ho seguito il dettato a 320 sillabe (circa 140 parole al minuto) per tre minuti. Al termine ho corretto il testo, l'ho stampato e l'ho consegnato. La mia prestazione è stata ritenuta 'fuori concorso' in quanto non ho usato la stenografia tradizionale.

A parte le considerazioni politiche, dal punto di vista tecnico io sostengo che la scrittura con la voce è **una stenografia**. Poiché queste mie considerazioni sono in primo luogo rivolte ai colleghi accademici che ben conoscono la storia e la tecnica della stenografia, mi riporto alla definizione classica di Enrico Noe, "La stenografia è l'arte di segnare il discorso altrui, od i propri pensieri, nel più breve tempo e spazio possibile". In questa definizione è evidente (come dimostrato da secoli di pratica) che il verbo 'segnare' indica appunto 'trasformare in segni'. Tale trasformazione avviene con segni 'criptici' noti agli addetti, come la stenografia e la stenotipia tradizionale (il sistema Michela fino agli anni 80 aveva battute simili a quelle dei sistemi stenografici geometrici). E' da considerare che anche le macchine per stenotipia che usano le tradizionali lettere dell'alfabeto, sono comunque costrette a ricorrere a combinazioni particolari di due o tre lettere per razionalizzare le

battute e quindi la lettura della striscia richiede comunque sempre una conoscenza delle regole del sistema, come è richiesto a chi voglia leggere delle cartelle stenografiche. Solo negli ultimi 10 anni la stenopia, valendosi dell'accoppiamento con programmi informatici, ha consentito di avere un testo completamente leggibile in tempo reale.

La scrittura con la voce, invece, parafrasando e completando la definizione classica, è l'arte di segnare il discorso altrui (ed i propri pensieri) nel più breve tempo possibile ed **in caratteri leggibili da tutti**. Il perfezionamento tecnico quindi non dimentica una lunga storia, ma anzi da essa trae i motivi per un rinnovato vigore. L'opera dell'oscuro tachigrafo romano è assimilabile a quella del moderno resocontista con stenografia, stenopia o riconoscimento del parlato, anche se penso che oggi nessuno più consideri di utilizzare le Note Tironiane per fare un resoconto 'classico'.

Ritengo altresì che non ci sia nessuno che dubiti che un testo scritto con la scrittura ordinaria (e non 'in italiano' come viene spesso detto ed anche scritto in modo illogicamente assurdo in qualche occasione) sia più immediatamente utilizzabile di quello in caratteri stenografici (che richiede l'intervento di un interprete), di conseguenza la stenopia elettronica ed il riconoscimento del parlato hanno al momento e da questo punto di vista una superiorità pratica rispetto alla scrittura stenografica. (Tentativi di superare questo gap sono, come è noto, in atto in Italia e Giappone e sono stati tentati in Inghilterra e Francia).

Tutte queste tecnologie hanno comunque un'immagine esterna che non corrisponde alla realtà dei fatti: si ritiene che riprendere il discorso, scrivendo su un foglio, digitando su una tastiera o parlando al computer sia un'attività semplice, meccanica. La colpa di questa situazione è da imputare agli stessi operatori del settore che strillano che la loro tecnologia è migliore dell'altra (spesso senza dimostrarlo) e riducono il costo della loro prestazione al di sotto dei livelli di semplice manualità praticati da volonterosi e bisognosi extra-comunitari (in senso positivo).

Non è infrequente esperienza di trovarsi all'interno di un consesso e mentre si sta stenografando essere avvicinati da qualcuno che chiede un'informazione, come se chi stenografa (sempre in senso lato) non fa niente od è soltanto un automa che è messo lì come una pianta colorita per dare lustro all'ambiente. Se è pur vero (ed è stato asserito da valenti stenografi come Frittitta – Vercesi – Galletti per citare quelli che ho potuto avvicinare) che l'elevato automatismo raggiunto per seguire le più alte velocità d'eloquio lasci spazio ad altri pensieri nel caso di scrittura a velocità ridotta^[1] (tant'è che qualcuno teme una ridotta velocità che può portare ad una diminuzione del livello di attenzione) è comunque altrettanto vero che un resocontista non può distrarsi o lasciarsi distrarre nel corso di una ripresa e deve sforzarsi di seguire il discorso anche quando l'argomento è ostico, le idee sono in contrasto con quelle professate (e verrebbe la voglia di replicare), l'oratore parla malissimo (e lo si vorrebbe strozzare) , alcune necessità fisiologiche devono essere contenute e via discorrendo.

Proprio dimenticando tutti questi fatti, alcune persone anche del settore definiscono la scrittura con la voce, un'azione **da pappagallo**, dando italianamente a questo termine un contenuto dispregiativo.

Se parzialmente si può scusare chi, dall'esterno, non si è soffermato a valutare tutti gli aspetti prima precitati, penso che una tale affermazione da parte di persone del settore sia quanto meno superficiale, non corrispondente al vero ed atta a rafforzare un atteggiamento di svalutazione dell'attività del resocontista (con qualunque tecnologia venga attuata).

Si ingenera infatti l'idea che sia la tecnologia a nobilitare il resocontista, e non il contrario. Da tempo vado sostenendo che non si crea automaticamente il resocontista regalandogli una scatola del miglior programma di riconoscimento del parlato (o una avanzatissima macchina stenotipica), come io (e penso anche qualcuno dei lettori) non divento Schumaker e nemmeno Barrichello se mi regalano la Ferrari!

Il principale argomento addotto a sostegno della definizione 'da pappagallo' è quello che il resocontista non deve far altro (sic?!) che riprodurre le parole dell'oratore e che questa tecnologia

non consenta altre alternative, cosa che invece (sempre a loro dire) è possibile con la stenografia o con la stenotipia.

Evidentemente coloro che sostengono queste argomentazioni non conoscono o si sono dimenticati (si dice che i dolori per l'acquisizione della professionalità stenografica o stenotipica siano come quelli del parto, si dimenticano immediatamente dopo l'evento) che per arrivare ad avere il testo scritto (in prima stesura) si attua un complesso meccanismo mentale. Tale meccanismo può essere sinteticamente così indicato:

ascolto con attenzione	
Comprensione di quanto detto	
Eliminazione di elementi superflui (rumori – ripetizioni – esitazioni ecc.)	
Sintesi in base alle regole stenografiche – stenotipiche	
Invio del comando all'organo interessato (mano – mani)	
Controllo che la trasmissione del comando non contenga elementi che lo rendano equivoco	
Esecuzione del comando (scrittura con la mano – battuta stenotipica) controllo dell'esecuzione affinché vengano chiaramente indicati elementi distintivi (es. rafforzamento per la stenografia) e non si introducano elementi modificatori (deformazione del segno stenografico oltre i livelli accettabili – errata diteggiatura che porti alla scrittura di una battuta errata)	

Ho lasciato volutamente in bianco la colonna di destra, dove ognuno possa riportare quali di queste fasi non avvengono nel caso di scrittura con la voce. Salvo errori ed omissioni, esse sono tutte applicabili alla specifica tecnologia.

Le varie regole stenografiche/stenotipiche tendono soprattutto a dare maggiore precisione e differenziazione ai segni stenografici, anche nel riconoscimento vocale occorre prestare attenzione per la sicura distinzione di parole che potrebbero ingenerare equivoci.^[2] Nella scrittura stenografica o stenotipica si ricorre a sigle per facilitare l'esecuzione di locuzioni particolari o ripetitive, analogo procedimento è sviluppato col riconoscimento del parlato tramite macro-comandi a voci. In America, dove la tecnologia di riconoscimento vocale è stata particolarmente messa a punto, traendo proprio spunto dall'esperienza stenotipica (che a sua volta riposa su quella stenografica) i programmi informatici sono denominati SpeechCAT, in analogia a quelli CAT in uso per la stenotipia.

Se, come io sostengo, il procedimento mentale-pratico è identico a quello utilizzato per la stenografia-stenotipia, in contrapposto alla dicitura 'da pappagallo', si potrebbe dire che le altre tecnologie sono 'da urango' in quanto quest'ultimo animale imita fortemente gli atteggiamenti umani. Ma non voglio alimentare inutili polemiche 'semantiche' ed entrare più specificamente in due aspetti a mio avviso fondamentali.

Il primo riguarda l'asserita impossibilità, da parte di chi resoconta con il riconoscimento del parlato di svolgere un'attività di sintesi o di eliminazione o correzione di elementi spuri, contrariamente a quanto si dice possibile con la stenografia o stenotipia.

La possibilità di svolgere un'attività di 'correzione' in diretta riposa a mio avviso su due elementi:

- 1 – la possibilità in assoluto di essere realizzata
- 2 – la possibilità di farla nel caso concreto

Per quanto riguarda il primo aspetto, se tale attività (che è principalmente di tipo mentale) può essere svolta da un resocontista stenografo o stenotipista non esistono preclusioni per il resocontista che utilizzi la tecnologia di riconoscimento della voce. Si tratta, ripeto, di un processo mentale. Anche qui si può ricordare che non è la tecnologia che fa il resocontista, ma il contrario.

A sostegno di questa mia tesi mi permetto far osservare una similitudine tra questo procedimento e quello di ascolto e traduzione tipica degli interpreti di simultanea. E' indubbio che un interprete di simultanea aggiunga, alle fasi citate per la resocontazione, quelle della esigenza di riformulazione della frase in relazione alla lingua di destinazione. Si pensi ad un semplice esempio: quello della traduzione tedesco-italiano, dove nella prima lingua il verbo viene solitamente posto al termine della frase mentre in italiano è necessario anteporlo, pena anche l'incomprensione o quanto meno una notevole cacofonia della frase tradotta. Monique Politi ^[3], in una relazione intitolata 'Interpretazione simultanea tra due lingue romanze. Problemi di riformulazione in lingua straniera', sottolinea che l'attività di interpretazione simultanea si trova "in una specifica dimensione temporale: l'immediatezza del processo" e che "tale attività richiede tre sforzi concorrenti (Gile 1988) ascolto-analisi-comprensione, memorizzazione e riformulazione nel rispetto dei codici linguistici della lingua d'arrivo." Se tale attività è (con adeguata formazione) possibile a così alto livello, non ci sono dubbi che possa essere svolta (sempre con adeguata formazione) anche senza la necessità di tradurre in altra lingua, ma soltanto in messaggio comprensibile nella lingua base.

Il processo mentale necessario per questa attività è ben indagato nello studio 'L'interpretazione simultanea: premesse di neuro linguistica' a cura di Laura Gran dell'Università di Trieste. Questo studio è riprodotto nel volume già indicato nella nota 1 e da essa traggio una informazione utile ai fini didattici, confermata nei corsi svolti unitamente a Maria Luisa Corti Crippa ^[4] quale primo approccio alla tecnologia di riconoscimento della voce. Tale informazione dice "in compiti verbali di attenzione divisa, in particolare nell'interpretazione simultanea, soggetti non allenati tendono a spostare l'attenzione, alternandola tra il messaggio in LS e quello in LA (*in pratica tra quello che si sente e quello che si forma in uscita n.d.a.*), nonché ad **umentare l'intensità vocale**"

Nessun dubbio quindi che '*con adeguata formazione*' questo processo possa essere svolto.

Per quanto riguarda il secondo aspetto, ossia la possibilità di svolgere questa attività nel caso concreto, mi pare ovvio che ciò risulti da due fattori: l'oratore ed il resocontista. Cito l'oratore in quanto se si resoconta un oratore che svolge il proprio intervento in modo chiaro e corretto, poche o nulla saranno le modifiche da apportare (e questo vale con qualunque tipo di tecnologia) mentre il resocontista può strategicamente agire in due modi. Il primo è appunto quello di apportare subito le correzioni ritenute utili, il secondo è quello di riprodurre fedelmente il discorso riservandosi una successiva elaborazione. Il tutto è condizionato, a mio giudizio, oltretutto alla indubbia capacità del resocontista, all'utilizzo del testo finale. Oggi l'orientamento della clientela (intesa in senso lato) privilegia l'esigenza di disporre del testo in tempi brevi dal termine dell'evento, od addirittura in tempo-reale. Si consideri che ad esempio, limitandoci solo all'Italia, che le maggiori assemblee parlamentari, Senato, Camera ed Assemblea Siciliana, privilegiano la diffusione del resoconto scritto con buona qualità linguistica e grafica contemporaneamente all'andamento dei lavori (di norma ad intervalli di mezz'ora). Il Senato invia in Internet sia il resoconto stenografico integrale che quello sommario (di notevole utilità ad esempio per i giornalisti e chiunque voglia prendere visione rapida degli argomenti discussi e delle conclusioni prese). Nel caso di utilizzo del testo in tempi brevi, sempre privilegiando il concetto di fedeltà al pensiero esposto, è logico ridurre la 'limatura' del testo che verrà invece riservata per atti destinati a passare alla storia (anche se non è facile il distinguo).

Un'altra considerazione mi corre obbligo, per completezza informativa, ed è che l'attività di riformulazione da parte del resocontista è possibile se 'si padroneggia' l'oratore sia in termine di conoscenza dell'argomento e della terminologia, sia in termini di rapidità di risposta (velocità di ripresa almeno uguale a quella dell'eloquio). E' una considerazione ovvia, ma non sempre ricordata. Altrettanto ovvio, a mio giudizio, è che se si possiede l'abilità di effettuare una resocontazione

sommaria, tale abilità può valersi della tecnica di scrittura con la voce anziché quello delle altre scritture più o meno veloci.

Molti sono comunque gli aspetti utili da approfondire, in questo l'Accademia deve svolgere il suo ruolo di riferimento imparziale, rivalutando le esperienze del passato alla luce dei nuovi e continui perfezionamenti tecnici (per questo l'abbiamo ridefinita 'multimedialità della scrittura e dell'informazione), eliminando il 'pappagalismo', (questa volta il termine è invece perfettamente azzeccato) ripetitivo di concetti ritriti.

In questo senso auspico uno sforzo comune per la messa a punto di programmi formativi che prescindano dalla tecnologia (stenografica – stenotipica o di riconoscimento della voce) e si focalizzino sulle qualità di base necessarie al resocontista, senza le quali appunto la formazione tecnica non serve quasi a nulla. Questi programmi debbono tener presenti le nostre esperienze e, perché no, anche quelle di altri in settori consimili quali quello da me citato. Disponendo di tali programmi formativi (in cui occorre la confluenza di più esperienze ed abilità) sarà più facile presentare proposte concrete ai vari Enti interessati ad un discorso strategico.

A titolo di curiosità rilevo infine che nella bibliografia del predetto volume è riportato uno studio del prof. M. Cortellazzo dell'Università di Padova, risalente al 1985 che si intitola "Dal parlato al (tra)scritto: i resoconti stenografici dei discorsi parlamentari", pubblicato (ovviamente) all'estero e poco conosciuto e meditato nei nostri ambienti.

^[1] Questo comportamento è spiegato dalle indagini di psicologia sperimentale. Per maggiori informazioni si legga il punto 14.5.1 della relazione di Laura Gran nel volume 'Interpretazione simultanea e consecutiva' – Hoepli 1999

^[2] Es. provare – trovare; pelle – belle, contatti-contratti e simili anche se lo sviluppo delle pratiche di 'intelligenza artificiale', termine da me non completamente condiviso, legate al contesto, oggi riducono notevolmente la possibilità di forme errate.

^[3] Sempre nel volume 'Interpretazione simultanea e consecutiva' – problemi teorici e metodologie didattiche – Hoepli 1999

^[4] Asfor Milano – 1999-2000