

||  
*Consiglio Nazionale delle Ricerche*

||  
**ISTITUTO DI ELABORAZIONE  
DELLA INFORMAZIONE**

**PISA**

||  
Pubblicazione B0-02

---

**E. Locuratolo**

**TRASCRIZIONE AUTOMATICA:  
UN'ESPERIENZA**

---

Estratto da: MANAGEMENT E INFORMATICA, XXIV, n. 3 (1986), 191-195

# TRASCRIZIONE AUTOMATICA: UN'ESPERIENZA

Elvira Locuratolo\*

Lo sviluppo di un sistema di trascrizione automatica basato sulla macchina stenografica «Michela» in uso presso il Senato della Repubblica. Lo scopo del sistema è trasformare i testi stenotipati in testi leggibili facilitando così il lavoro di decrittazione.

## Il metodo Michela

Il metodo di stenotipia utilizzato in Italia, chiamato «Michela» dal nome del suo inventore, permette di codificare le parole della lingua italiana in forma pseudo-sillabica per mezzo di un dispositivo munito di tastiera, e di registrare le pseudo-sillabe su un supporto magnetico per mezzo di un minielaboratore collegato al dispositivo. La tastiera della macchina è suddivisa in due parti simmetriche su cui operano rispettivamente la mano sinistra e la mano destra dello stenotipista.

Ciascuna di tali parti è ulteriormente suddivisa in due zone, chiamate serie, comprendenti rispettivamente sei e quattro tasti sulla parte sinistra e quattro e sei tasti sulla parte destra.

Lo schema della tastiera riflette la composizione strutturale delle pseudo-sillabe: infatti, la generica pseudo-sillaba consta al più di quattro caratteri ed è imprimitibile con una sola battuta. La fig. 2 riporta la codifica dei caratteri per ciascuna serie denotando con la «x» nella i — ma posizione di una colonna la battitura sull'i — mo tasto della serie corrispondente. I caratteri possono essere codificati pigiando uno o più tasti per complessivi ventisei caratteri sulla prima e sulla quarta serie e undici sulla seconda e sulla terza. Si osservi l'analogia tra la prima e la quarta colonna dove figurano caratteri prevalentemente consonantici: il «\$», la «k» e la «y» esprimono le distinzioni fonetiche accettate dal metodo rappresentando, rispettivamente, la «s» dolce, la «c» dura e la «g» dolce. Nella terza colonna sono presenti le vocali, mentre nella seconda compaiono sia alcune vocali, sia alcune consonanti. Esaminando la figura 2) è possibile capire il perché di alcune scomposizioni pseudo-sillabiche: ad esempio, la parola «scrivere» non può essere suddivisa come «skri-ve-re» poiché la composizione «skri» non è imprimitibile con una sola battuta: pertanto, il metodo rinvia la «s» all'ultima sillaba della parola precedente e suddivide l'espressione «di scrivere» come «dis-kri-ver». L'esempio riportato presenta inoltre una eliminazione di vocale; tale accorgimento può essere ripetuto anche più volte nell'ambito di una stessa parola per accrescere la velocità di battitura; infatti, per ogni vocale eliminata il numero di pseudo-sillabe componenti la parola si riduce di una unità, come è possibile verificare nel caso di «attività» (at-tvit). Altre volte, invece, una parola può essere espressa soltanto con la sua parte iniziale, come nel caso di «Camera» (Cam), o con una sua forma caratteristica, come nel caso di «collega» (cleg).

## Introduzione

La registrazione in linea di conferenze, dibattiti, etc, viene attualmente eseguita con l'ausilio di macchine stenografiche che producono testi da sottoporre a successive elaborazioni: tuttavia, le procedure adottate per la trascrizione corrente risultano molto costose e comportano pesanti ritardi nella pubblicazione dei testi. Queste difficoltà sono particolarmente avvertite nell'ambito del Servizio dei Resoconti Parlamentari, dove la trascrizione dei testi stenotipati assume un ruolo centrale per le funzioni documentali. Per facilitare il lavoro del personale addetto alla decrittazione, l'esperienza sulla trascrizione automatica è stata rivolta alla realizzazione di un sistema per l'elaborazione dei testi stenotipati. Il sistema, sviluppato presso l'Istituto di Elaborazione dell'Informazione di Pisa in collaborazione con il Senato della Repubblica, è scritto in Pascal e gira su un Ibm 3031 con risultati alquanto soddisfacenti: esso però è da ritenersi soltanto il primo prodotto di questa esperienza e necessita di ulteriori fasi di studio per giungere ad una versione definitiva.

Il presente articolo illustra alcune caratteristiche del metodo di stenotipia adottato in Italia, descrive la metodologia generale del sistema, discute i risultati ottenuti e termina con prospettive di sviluppi futuri.

\* Consiglio Nazionale delle Ricerche. Istituto di Elaborazione dell'informazione.

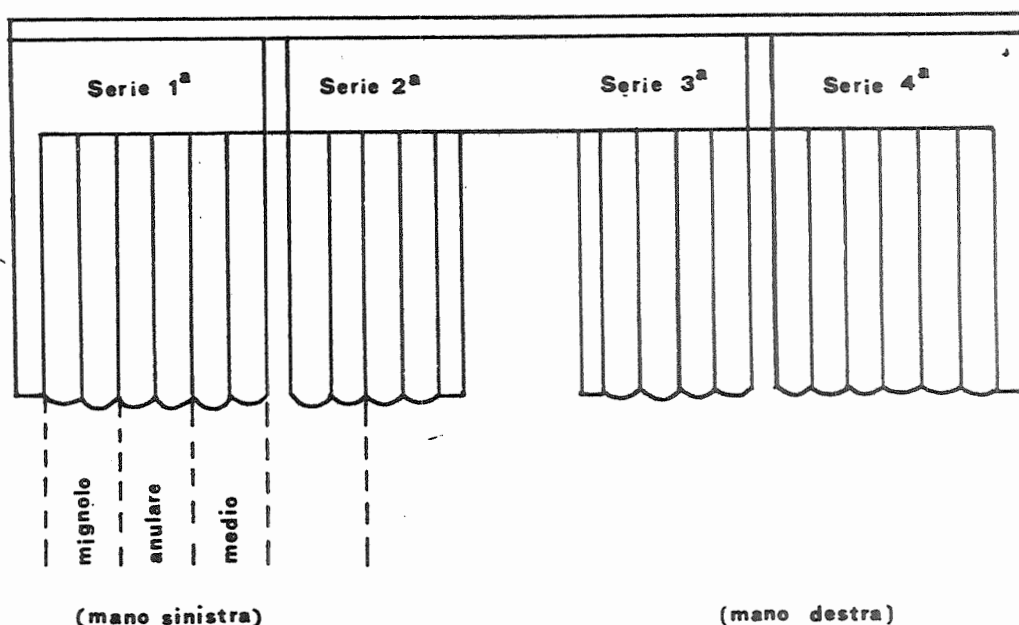


Fig. 1

In generale non esistono regole rigide e una stessa parola può essere suddivisa in più modi dipendenti da fattori diversi, quali l'abilità dello stenografo, la rapidità dell'oratore etc. A questo stato di cose si aggiungono gli inevitabili errori commessi dagli stenografi; ad esempio la difficoltà che si incontra nel premere l'ultimo tasto della quarta serie, spiega la possibile sostituzione della «r» con la «u» in finale di pseudo-sillaba.

La fig. 3, mostra una stringa di pseudo-sillabe costituenti l'ingresso al sistema.

## Metodologia del sistema

Il sistema è stato sviluppato a due livelli: il *primo* si basa sull'analisi delle pseudo-sillabe e mira alla individuazione di intere parole stenotipate, ovvero parole mancanti di vocali e/o desinenze finali; il *secondo* si basa sull'uso di un dizionario e mira a trasformare il trascritto ottenuto al primo livello in un testo della lingua italiana le cui parole però non sono concordate in genere e numero.

### Approccio adottato al primo livello

Per agevolare il processo di analisi e manipolazione delle pseudo-sillabe, si è pensato di scindere la stringa di ingresso al sistema di sottostringhe indipendenti tra loro. I punti di scissione prescelti sono stati il margine sinistro delle pseudo-sillabe iniziali di parola e il margine destro delle pseudo-sillabe finali di parola. Teoricamente, infatti, se riuscissero ad individuare tutte le pseudo-sillabe iniziali di parola (o tutte quelle finali di parola), basterebbe poi procedere all'aggregazione delle pseudo-sillabe all'interno di ogni sottostringa per avere le parole stenotipate complete, e viceversa, se si riuscissero ad individuare tutte le pseudo-sillabe da aggregare alla pseudo-sillaba successiva (o tutte quelle e da aggregare alla pseudo-sillaba precedente), i margini delle parole stenotipate ottenute determinerebbero i punti di scissione della stringa.

In pratica, però, ci sono delle pseudo-sillabe che assumono ruoli diversi in base al contesto linguistico occupato per cui nessuna delle suddette condizioni è completamente conseguibile. Così, l'orientamento adottato è stato quello di ricercare le pseudo-sillabe soddisfacenti ad una qualsiasi delle condizioni citate e di elaborare le due seguenti operazioni di *segmentazione* e *aggregazione* in modo opportuno:

La *segmentazione* consiste nello scindere la stringa di partenza sul margine sinistro delle pseudo-sillabe iniziali di parola; mentre l'*aggregazione* consiste nell'affiancare le pseudo-sillabe che non risultano iniziali di parola a quelle che le precedono (risp. le pseudo-sillabe che non risultano finali di parola a quelle che le seguono).

Tali operazioni sono governate da un insieme di regole che permettono di individuare le pseudo-sillabe di interesse.

### Regole

La determinazione delle regole si è basata sui seguenti fattori:

- caratteristiche del metodo Michela;
- analisi delle pseudo-sillabe;
- relazioni che introducono tra una pseudo-sillaba e quelle del proprio contesto.

Per ciascun fattore si riporta una regola di rilievo e la relativa giustificazione.

**Fattore a):** «le pseudo-sillabe inizianti per vocale individuano una scissione della stringa sul loro margine sinistro».

#### Eccezione

«una pseudo-sillaba iniziante per i preceduta e seguita da vocale, può non essere iniziale di parola».

Es. a iu tar

La regola citata trova riscontro nel fatto che il metodo Michela generalmente riduce il numero delle sillabe costituenti una parola o con l'eliminazione di alcune vocali, o con l'eliminazione di desinenze

f	x....	r	x....	u	x..	f	....x
s	.x...	s	.x..	i	.x..	s	....x.
sc	..x...	i	..x.	e	..x.	sc	...x..
\$	...x..	u	...x	a	...x	\$	..x...
p	....x.					p	.x....
n	....x					n	x....
h	x.x...	l	x.x.	-	.x.x	h	...x.x
v	.xx...	vf	.xx.	o	.xx.	v	...xx.
j	x...x..	m	x..x	è	x..x	j	..x..x
z	.x.x..	n	.x.x	eu	x.x.	z	..x.x.
t	x...x.	pb	..xx	u	xx..	t	.x...x
c	.x..x.					c	.x..x.
k	..x.x.					k	.x.x..
y	...xx.					y	.xx...
gn	x...x					gn	x...x
gl	.x...x					gl	x...x.
u	..x..x					u	x..x..
i	...x..			oi	xx.x	i	x.x...
b	x.x.x.	td	x.xx	ui	xxx.	b	.x.x.x
d	.xx.x.	kg	.xxx			d	.x.xx.
g	x..xx.					g	.xx..x
m	.x.xx.					m	.xx.x.
r	x.x..x					r	x..x.x
l	.xx..x					l	x..xx.
nt	x..x.x					nt	x..x.x
x	.x.x.x					x	x.x.x.

Fig. 2 - Codifica dei caratteri

finali per cui le pseudo-sillabe in questione sono iniziali di parola; mentre la sua eccezione è giustificata dall'impossibilità di imprimere un trittongo come «aiu» mediante una battuta soltanto.

**Fattore b):** «Le pseudo-sillabe iniziali di parola (risp. finali di parola) che non risultano essere parole stenotipate sono aggregabili alla pseudo-sillaba successiva (risp. precedente)»

Es: ab, ac, ul

Invero, le pseudo-sillabe che oltre ad essere iniziali di parola, non abbreviano alcuna parola della lingua italiana non possono che essere affiancate alle pseudo-sillabe successive; per cui «ab», ad esempio, sarà iniziale di una parola come «abside», «abnegare», etc.

**Fattore c):** L'analisi delle pseudo-sillabe in relazione al proprio contesto, ha permesso di definire le operazioni di aggregazione e segmentazione in modo più generale di quanto visto finora. A tale scopo, sono stati introdotti i concetti di *aggregazione forte* e *aggregazione debole* (risp. *segmentazione forte* e *segmentazione debole*) e i concetti di *aggregazione condizionata* e *segmentazione condizionata*.

L'*aggregazione forte* sussiste quando due pseudo-sillabe sono sicuramente parte di una stessa parola, come accade per la pseudo-sillaba «ab», mentre l'*aggregazione debole* sussiste quando esiste una piccola probabilità che le due sillabe possano appartenere a differenti parole; ad esempio, se una pseudo-sillaba finisce per «y» e la sua successiva inizia ancora per «y», è probabile che esse si congiungano in una doppia «g», in quanto sia le parole terminanti per «gi», «ge», etc., sia quelle inizianti allo stesso modo sono poco frequenti nella lingua italiana.

L'*aggregazione* (risp. la *segmentazione*) *condizionata* sussiste quando l'affiancamento (risp. la scissione) è permesso soltanto se si verificano opportune condizioni; ad esempio, la pseudo-sillaba «in» può essere utilizzata o come preposizione, o come prefisso. Nel primo caso, essa è scritta sempre come

«in», nel secondo caso essa è scritta come «im», se la sillaba successiva inizia per «b», «m» e «p» (es. imbrunire), e come «in» altrimenti.

Ne deriva la seguente regola:

«Se la pseudo-sillaba successiva ad *in* inizia per «b», «m», o «p», la *in* individua una scissione sul suo margine destro, altrimenti nulla si può dire».

L'implementazione delle regole individuate utilizza una semplicissima tabella contenente le pseudo-sillabe e le opportune specificazioni. Allo stato attuale essa è ancora incompleta poiché le pseudo-sillabe considerate e inquadrate nelle regole sono soltanto 1/3 di quelle totali; inoltre, poiché l'aggregazione è applicabile in modo ricorsivo, la tabella può contenere anche aggregati di pseudo-sillabe.

#### Vantaggi offerti dall'approccio

L'approccio adottato al primo livello si presta al conseguimento dei seguenti obiettivi:

- utilizzo del sistema per la trascrizione on-line di testi provenienti da ambienti linguistici diversi (ambiente politico, medico, etc);
- implementazione del sistema su un microcomputer.

A tali scopi, si osservi che:

- 1) le regole individuate si riferiscono alle pseudo-sillabe in quanto tali e prescindono da contesti linguistici particolari;

Fig. 3

IN TANT IL REN DIM DEL UN E DEL AL TRO NON E'U GUAL A CHEL LO DEL TI  
TOL DI CUI DIS CU TIAM E POI QUAL O PE RAZ COM FIAM? DI CIAM AL RIS PAU  
MIA TOU SIK COM SCE GLI I TI TII CHE TI DAN DI FTU' IY ME DIA TAM MA SIK  
COM NON SAP FIAM DOV VAN A FI NIU TI PRE MIAM E TI TAS SIAM DI MEN NON  
MI SEM BRA CHE VEN GA UN BEL SE GNAL DAL LO STAT IN QUES TO CAS IL HIS  
PAU MIA TOR IN VEM DEV ES SEV RES PON SAB DLA SCEL TA

- 2) la formazione di parole stenotate permette di non inserire nel dizionario i termini che risultano già essere parole della lingua italiana;
- 3) il partizionamento della stringa iniziale in sottostringhe indipendenti permette il parallelismo tra i due approcci proposti per il sistema ed ostacola l'eventuale propagazione di errori lungo il testo.

**Misura della leggibilità**

Per valutare i risultati ottenuti già al primo livello del sistema, si è definita una misura della leggibilità. La valutazione della leggibilità è una questione molto complessa poiché legata a vari fattori, alcuni dei quali anche di carattere personale: tuttavia, dal punto di vista del trascritto ottenuto, si può ritenere leggibile un testo le cui pseudo-sillabe sono state correttamente aggregate e le cui parole sono state correttamente separate, e non leggibile un testo in cui le aggregazioni e le segmentazioni non sono state eseguite, o sono state eseguite in modo errato.

Per esprimere matematicamente questo concetto, si denoti con S un testo stenotipato, con T1 la versione trascritta e con T la versione che si otterrebbe se si apportassero correttamente tutte le aggregazioni e le segmentazioni. Siano inoltre Na ed Ns, rispettivamente, le aggregazioni e le segmentazioni richieste per ottenere T; ed na, ea ed ua (risp. ns, es, ed us,) le aggregazioni corrette, non corrette e non risolte (risp. le separazioni corrette, non corrette e non risolte) richieste per ottenere T1.

La misura della leggibilità è definita nel modo seguente:

$$L(T) = \frac{1}{2} \left( \frac{na}{Na} \left( 1 + \frac{us}{Ns} \right) + \frac{ns}{Ns} - \frac{ea}{Ns} - \frac{es}{Na} \right)$$

Per cogliere il significato di questa formula si considerino le seguenti situazioni al limite:

- a) si apportano tutte le segmentazioni e tutte le aggregazioni in modo corretto;

IN TANT \*\* IL \*\* REN DIM DEL \*\* UN \*\* E DEL \*\* AL TRO \*\* NON \*\* E' \*\* UGUAL  
 \*\* A CHEL LO DEL \*\* TI TOL DI \*\* CUI \*\* DISCUTIAM \*\* E \*\* POI \*\* QUAL \*\* O  
 \*\*PE \*\* RAZ \*\* COMPIAM \*\* ? \*\* DICIAM \*\* AL RISP AU MIA TOR \*\* SIK COM  
 SCEGLI \*\* I TI TII CHE \*\* TI DAN DI PIU \*\* IMMEDIATAM MA SIK COM \*\* NON  
 \*\* SAPPIAM \*\* DOV VAN \*\* A FI NIU TI PREMIAM \*\* E TI TASSIAM \*\* DI MEN \*\*  
 NON \*\* MI SEMBRA CHE \*\* VEN GA \*\* UN \*\* BEL \*\* SEGNA DAL LO \*\* STAT \*\*  
 IN QUES TO CA\$ \*\* IL \*\* RISP AU MIA TOR \*\* IN VEM \*\* DEV \*\* ESSER RESPON  
 SAB \*\* DLA SCEL TA \*\*

Fig. 4

Fig. 5

INTANTO IL RENDIMENT DEL UN E DEL ALTRO NON E' UGUAL A CHEL LO DEL  
 TITOLO DI CUI DIS CUTIAM E POI QUALE OPERAZION COMPIAM? DICIAM AL  
 RISP AU MIA TOU SICCOME SCEGLI I TI TII CHE TI DAN DI PIU' IY ME DIATAM  
 MA SICCOME NON SAPPIAM DOVE VAN A FINIU TI PREMIAM E TI TASSIAM DI  
 MEN NON MI SMBRA CHE VENGA UN BEL SEGNA DALLO STAT IN QUEST CAS IL  
 HISP AU MIA TOR IN VEM DEVE ESSER RESPONSABIL DELLA SCELTA

- b) non si apporta alcuna segmentazione ed alcuna aggregazione;
- c) si aggregano tutte le pseudo-sillabe in una sola lunga parola;
- d) si apportano tutte le aggregazioni in modo corretto e nessuna segmentazione;
- e) si apportano tutte le segmentazioni in modo corretto e nessuna aggregazione;
- f) si apportano tutte le segmentazioni al posto delle aggregazioni e viceversa.

Relativamente al caso a), risulta: na = Na; ns = Ns L(T) = 1;

Relativamente al caso b), risulta: na = ns = es = ea = 0 L(T) = 0;

Relativamente al caso c), risulta: na = Na; es = Ns L(T) = 0

Relativamente al caso d), risulta: na = Na; us = Ns L(T) = 1;

Relativamente al caso e), risulta: ns = Ns; ua = Na L(T) = 1/2;

Relativamente al caso f), risulta: ea = Ns; ea = Na L(T) = -1.

In base a tale formula, la leggibilità di un testo appartiene all'intervallo (0,1) ed assume valori negativi in corrispondenza di situazioni patologiche. La stessa formula può essere utilizzata anche per calcolare la leggibilità al secondo livello, quando si utilizza cioè il dizionario.

La fig. 4 fornisce una idea visiva dei risultati ottenuti al primo livello. La valutazione numerica è stata poi calcolata come valore medio delle leggibilità di varie zone di un testo, e si aggira intorno al 47% con valori compresi tra il 39% e il 54%. Essa è chiaramente collegata al numero di pseudo-sillabe inquadrate nelle regole ed è un buon risultato poiché curando ulteriormente il processo di analisi si possono raggiungere risultati variabili intorno al 60%. Naturalmente, la leggibilità effettiva del testo viene poi migliorata ispezionando il dizionario.

**Il dizionario**

Il dizionario utilizzato dal sistema consta di circa 2000 ingressi costituiti da una parola stenotipata e dalla corrispondente parola in lingua italiana. I termini introdotti rientrano nell'ambito di quelli più comunemente usati nei dibattiti parlamentari e mancano delle desinenze finali.

Es: EKNOMIC -> ECONOMIC

L'ispezione nel dizionario si applica ad ogni sottostringa prodotta al precedente livello e procede parallelamente alle operazioni di segmentazione ed aggregazione.

Per trasformare il trascritto ottenuto al primo livello in un testo della lingua italiana si fa uso della tecnica del match più lungo. Tale approccio non è chiaramente in grado di risolvere correttamente tutti i casi ambigui, né il dizionario è sufficientemente ampio da contenere tutte le parole utilizzate nei dibattiti parlamentari; tuttavia, la necessità di ridurre le dimensioni del dizionario in vista di una futura implementazione su microcomputer, ha reso accettabile la non risolvibilità di alcune situazioni e il possibile inserimento di errori nel trascritto. I linguisti risolvono infatti le principali ambiguità analizzando il pattern costruito dalle classi sintattiche delle parole vicine ai termini stenografici. Ciò richiede però l'inserimento di tutte le classi sintattiche delle pa-

role nel dizionario senza tra l'altro poter eliminare quelle ottenute in modo diretto.

La fig. 5 fornisce una idea visiva dei risultati ottenuti dalla versione attuale del sistema. La valutazione numerica della leggibilità del testo di uscita si aggira intorno all'85% con valori compresi tra il 72% e il 92%; essa, però, può essere ulteriormente migliorata. A tale scopo, si sta progettando un dizionario suddiviso in due parti: una contenente le radici delle parole della lingua italiana, e una contenente suffissi. Poiché questi ultimi sono altamente ripetitivi, la suddivisione prevista permetterà di contenere il dizionario in dimensioni ragionevoli; non solo, ma nella maggior parte dei casi sarà possibile associare ai suffissi anche le classi sintattiche e questo permetterà di risolvere sia alcune ambiguità, sia alcuni problemi di concordanza.

## Conclusioni

La trascrizione automatica dei testi stenotipati è un problema rilevante da un punto di vista *pratico*, poiché il suo uso nei dibattiti, simposi, etc. può migliorare notevolmente la pratica corrente dell'attuale trascrizione manuale; da un punto di vista *scientifico*, per lo stimolo che l'applicazione stessa è in grado di fornire alla proposta di nuovi approcci risolutivi e da un punto di vista *delle applicazioni*

*future*, poiché esso si pone come fase intermedia al riconoscimento del parlato.

Queste motivazioni hanno portato allo sviluppo di sistemi di trascrizione automatica per diverse lingue. Il presente articolo descrive una esperienza di trascrizione automatica per la lingua italiana basata sulla macchina stenografica «Michela». Gli approcci e le tecniche utilizzate sono state quasi sempre dettate da finalità pratiche; tuttavia, la metodologia adottata risulta semplice ed elegante poiché basata su concetti altamente formalizzabili.

## Ringraziamenti

Ringrazio il prof. R. Sprugnoli e la dott. M.R. Laganà del Dipartimento di scienze dell'informazione di Pisa e il dott. C. Pinzani e l'ing. F. Guelfi del Senato della Repubblica Italiana per la collaborazione offertami durante il mio lavoro.

## Riferimenti

- [1] A.M. Deroualt, B. Merialdo, J.L. Stehle, «Une expérience de transcription automatique stenotypie-français»; *Techniques et science informatiques* vol. 2, n. 5, 1983.
- [2] A.C. Downton, «Speech Transcription as a Communication Aid for the Deaf», *Care. Health and Development*, n. 5, 41-48, 1979.
- [3] A.C. Downton, A.f. Newell, «An Assessment of Palantype Transcription as an Aid for the Deaf», *Int. J. Man-Machines Studies*, n. 11, 666-680, 1979.
- [4] M.R. Laganà, E. Locuratolo, «An approach to the automatic transcription of stenotyped texts», *Atti del Congresso Aica '85*, vol. 1, 161-177.