

A. M. CIRESE

E1/RB. Esperimento di elaborazione elettronica IBM
di cento testi di canti popolari della Raccolta Barbi

Relazione

Sommario

1. Progettazione ed esecuzione
 - 1.1 Finalità dell'E1/RB
 - 1.2 Fasi e modalità della progettazione
 - 1.3 Determinazione finale del progetto dell'esperimento
 - 1.4 Operazioni già realizzate e operazione residua ancora in corso

 2. Ragioni e modalità delle singole operazioni previste ed attuate
 - 2.1 Scelta del campione
 - 2.2 Preedizione perforazione, verifica
 - 2.21 Criteri del passaggio dal manoscritto alla scheda perforata
 - 2.22 Perforazione
 - 2.23 Verifica
 - 2.3 Stampa dei testi
 - 2.4 Incipitari
 - 2.5 Concordanze
 - 2.6 Forme elencate in ordine alfabetico e in ordine di frequenza
 - 2.7 Rimari

 3. Conclusione
-

Roma, settembre 1967

1. Progettazione ed esecuzione

=====

1.1 - Finalità dell'E1/RB

L'esperimento che si è convenuto di designare con la sigla E1/RB è stato concepito, progettato e attuato allo scopo di accertare quali siano le effettive possibilità d'impiego delle apparecchiature elettroniche tanto per realizzare gli strumenti repertoriali già in precedenza progettati e in corso di esecuzione manuale (incipitari e rimari), quanto per ottenere elaborazioni più complesse (concordanze; frequenze di singoli termini e di sintagmi, stilemi, moduli; elenchi delle parole che si trovano collocate in reciproca corrispondenza di rima, ecc.) che sono ovviamente di enorme importanza nello studio dei canti popolari, ma che apparivano manualmente irrealizzabili data l'enorme massa dei documenti da sottoporre allo spoglio.

Si è anche considerato che l'esperimento, pur se strettamente delimitato secondo quanto richiedono elementari norme di metodo, poteva fornire indicazioni ed elementi di giudizio circa la utilizzazione degli elaboratori elettronici in altri settori della letteratura popolare

(per es. le fiabe) e più in generale in tutto il complesso delle ricerche demologiche (con particolare riferimento al programma complessivo di una enciclopedia delle tradizioni popolari).

1.2 - Fasi e modalità della progettazione

Il progetto dell'esperimento E1/RB, delineato nelle sue linee scientifiche fondamentali dal prof. Santoli e dallo scrivente, è stato messo a punto in due momenti strettamente interdipendenti. Nel primo, con la preziosa e intelligente collaborazione della dott. sa Paola Raicich Tabet e del sig. Salvatore Barone, si è fissato un elenco preliminare ma razionale e graduato dei desiderata, e cioè delle elaborazioni che era indispensabile, necessario o semplicemente opportuno richiedere alle macchine elettroniche. Nel secondo momento, sempre con l'attivissima cooperazione di Paola Raicich Tabet e di Salvatore Barone i desiderata sono stati esaminati, interpretati, adattati e definiti alla luce delle possibilità teoriche offerte in generale dalle apparecchiature elettroniche ed in rap-

porto alle attrezzature di 'programmi' attualmente disponibili presso il CNUCE di Pisa. In questa fase di progettazione specifica e definitiva (come del resto in tutta la faticosa e pur celere realizzazione dell'esperimento) è stata di decisiva importanza la cooperazione del dott. Antonio Zampolli del Centro studi IBM, che ha messo a completa disposizione la sua ricca esperienza di programmatore nel campo linguistico e in quello della documentazione suggerendo possibilità e aprendo prospettive in precedenza insospettate o appena intraviste.

Occorre anche aggiungere che né la progettazione né l'esecuzione dell'esperimento sarebbero state possibili senza il patrimonio di documenti, di criteri e di prospettive acquisito (anche a livello di esperienze e competenze personali) attraverso il lavoro intellettuale e manuale di inventario e ordinamento dei manoscritti della Raccolta Barbi, di costruzione degli incipitari, dei rima-ri e dello schedario base dei termini significativi: in una parola attraverso tutte le ricerche e le operazioni finora effettuate in attuazione del programma approvato e finanziato dal Consiglio Nazionale delle Ricerche. Notevoli a-

evoluzioni hanno fornito anche il modello offerto dalla edizione della Divina Commedia elaborato e stampato nel VII centenario dantesco in occasione della inaugurazione del CNUCE e del Sistema elettronico IBM 7090 a Pisa, e l'esistenza presso il CNUCE di Pisa di alcuni programmi elaborati per altre operazioni su materiale linguistico, tra cui quelle per il Vocabolario dell'Accademia della Crusca. Le particolari caratteristiche della materia da sottoporre alla elaborazione, e le specifiche finalità della nostra ricerca, hanno tuttavia richiesto l'approntamento di altri quattro programmi elettronici speciali, curati dal dott. Zampolli, alla cui Nota tecnica è qui opportuno fare rinvio preliminare e generale (cfr. All. II c, pp. II-XI).

1.3 - Determinazione finale del progetto dell'esperimento

Il lavoro di progettazione si è concluso in data 15 luglio 1967 con le seguenti determinazioni:

"L'esperimento E1/RB ha lo scopo di verificare, in via preliminare e dimostrativa, le possibilità teoriche e pratiche offerte dai calcolatori elettronici IBM

per le operazioni di documentazione e di elaborazione dei dati nel campo specifico della poesia popolare e, in prospettiva, in quello complessivo della ricerca demologica.

L'esperimento, giovandosi della autorizzazione concessa dal prof. Vittorio Santoli nella sua qualità di Presidente del Comitato per la Raccolta Barbi, utilizzerà cento testi manoscritti della Raccolta stessa e dopo averli trasferiti su schede perforate realizzerà le seguenti operazioni:

- a. Riproduzione in caratteri di stampa, con indicazione della collocazione di ciascun testo;
- b. Incipitario e cioè elenco di tutti i versi dei cento componenti, disposti secondo l'ordine alfabetico degli inizi di verso o capoversi, e con segno distintivo per i versi di testa di ciascun componimento;
- c. Concordanze, e cioè lista alfabeticamente ordinata di tutte le parole contenute nei cento componimenti, con riproduzione per ciascuna parola dei versi in cui essa ricorre;
- d. Forme in ordine alfabetico, e cioè lista di tutte le parole contenute nei cento componenti con indicazione della loro frequenza e cioè del numero di volte in cui ciascuna di esse ricorre;
- e. Forme in ordine di frequenza decrescente, e cioè lista di tutte le parole contenute nei cento componenti, ordinate secondo il rango, ossia secondo l'ordine decrescente della loro frequenza;
- f. Rimario, e cioè elenco di tutti i versi dei cento componimenti, disposti secondo l'ordine alfabetico delle rime e cioè della terminazione dei versi a partire dall'ultima vocale accentata;

- g. Elenco delle 'rime perfette', delle 'assonanze regolari' e delle parole comunque collocate in reciproca relazione di 'proposta e risposta di rima';
- h. Gruppi ternari, e cioè lista di tutti i versi in cui ricorrono particolari raggruppamenti di parole (tre, senza calcolare le "parole grammaticali") con riproduzione integrale dei versi in cui la ricorrenza si verifica.
- i. Riduzione fotomeccanica e stampa dei tabulati forniti dagli elaboratori elettronici.

L'esecuzione dell'esperimento verrà curata in stretta collaborazione dal dott. Antonio Zampolli per conto del CNUCE (approntamento dei nastri-programma e cura di tutte le operazioni elettroniche) e dalla dott. ssa Paola Reicich Tabet e dal sig. Salvatore Barone per conto del gruppo di ricerca per il Corpo sistematico delle tradizioni popolari (approntamento dei testi base. loro trasferimento sulle schede perforate, messa a punto dei criteri di riproduzione e di esecuzione delle diverse elaborazioni).

L'esperimento verrà compiuto in tempo utile per la presentazione dei risultati al Consiglio Nazionale delle Ricerche, la cui scadenza è stata fissata al 30 settembre 1967."

1.4 - Operazioni già realizzate e operazione residua ancora in corso.

Nonostante la brevità del tempo a disposizione, e la complessità del lavoro, l'esperimento E1/RB è stato effettivamente realizzato in tutte le operazioni previste,

con la sola eccezione di quella concernente i gruppi ternari, che è tuttora in corso e di cui si riferirà non appena verrà completata. Il merito di tale risultato, che non si esita a definire assai più che soddisfacente va attribuito, oltre che alla razionalità del progetto, alla liberalità della Direzione scientifica e amministrativa del CNUCE e della IBM di Pisa, e in particolare al dott. Renato Pennacchi e prof. Guido Torrigiani, che hanno consentito la cooperazione di programmatori e tecnici e la utilizzazione dei tempi-macchina; al dott. Antonio Zampolli che alle sue competenze ha aggiunto un entusiasmo che ha consentito di superare difficoltà altrimenti preclusive; alla dott. ssa Raicich e al sig. Barone che hanno dedicato all'impresa non solo tutte le competenze solidamente acquisite nel corso dei precedenti lavori, ma anche tempi ed energie assai superiori a quelli cui la loro collaborazione con il gruppo di ricerca li impegnava; alla Direzione e al personale della Editrice Tecnico-scientifica di Pisa che hanno curato con rapidità e competenza la fotocopia dei manoscritti, la riduzione fotomeccanica dei tabulati e la stampa di tutto il materiale dell'esperimento.

2. Ragioni e modalità delle singole operazioni previste e attuate

=====

2.1 - Scelta del campione

La Raccolta Barbi è stata scelta come fonte da cui trarre i testi da sottoporre all'esperimento in ragione del fatto che offre testi manoscritti: era infatti opportuno verificare sin dall'inizio le difficoltà del loro trasferimento su schede perforate, e fornire elementi per stabilire i criteri e le modalità di tale operazione (cfr. 2.21).

Il numero dei testi è stato fissato a 100 perché tale numero costituiva un campione sufficientemente ampio per un primo esperimento, e tuttavia non superava le possibilità di elaborazione consentite in una fase iniziale e con tempo limitato.

La scelta si è limitata ai canti lirici per non aggravare l'esperimento con testi di eccessiva lunghezza e per avere campioni omogenei.

I singoli testi sono stati prelevati distribuendo la scelta su forme metriche diverse (rispetti, stornelli ecc.) e su componimenti di diversa lunghezza (da 3 a 17 versi). E' stato tenuto conto particolare dell'importanza che le va

rianti rivestono nel campo della poesia popolare: di qui la presenza nel campione di più lezioni (fino a 4) del medesimo testo.

Nella scelta non si è data preferenza a componimenti abitualmente considerati come "regolari" metricamente o "pregevoli" formalmente o esteticamente.

Per consentire il confronto tra i documenti di partenza e le successive operazioni, i 100 testi prescelti come campione sono stati riprodotti in fotocopia e vengono allegati alla presente relazione (v. All. II b).

2.2 - Preedizione, perforazione, verifica

Si ritiene opportuno accogliere il termine di 'preedizione' che nella terminologia specialistica delle elaborazioni elettroniche in campo linguistico designa tutte le operazioni che vengono manualmente condotte sui testi prima del loro trasferimento sulle schede perforate (per ogni particolare tecnico su queste operazioni cfr. All. II c, pp. IX-X).

Nel caso specifico del presente esperimento le ope-

razioni della preedizione vera e propria si sono strettamente unificate con quelle della perforazione e della verifica, in ragione del fatto che tutte le operazioni sono state eseguite dalle stesse persone. Qui dunque si dà conto cumulativamente delle tre fasi.

2.21 - Criteri del passaggio dal manoscritto alla scheda perforata: Il criterio generale che si è ritenuto opportuno adottare è stato quello della assoluta fedeltà al manoscritto (con la sola eccezione della punteggiatura: v. oltre). Per esempio, nel testo 0268 (p. 5 della Fotocopia, All. II b, e p. 1 dei Tabulati, All. II c) il ME del primo verso è stato trascritto fedelmente, nonostante sia un evidente errore di mano.

Con ciò non si intende sostenere che la fedeltà fotografica e meccanica sia il criterio da adottare in linea definitiva. Ma in assenza di norme precise e soprattutto in mancanza di criteri di discriminazione tra interventi leciti e interventi arbitrari, è parso opportuno attenersi rigidamente all'originale. Starà all'esame più approfondito del problema di determinare le norme definitive in pro

posito, anche per ciò che riguarda le fonti a stampa.

Per contro si è ritenuto indispensabile eliminare i segni di punteggiatura, e ciò per due ragioni che appaiono assai valide. La prima è che nei testi di tradizione orale la punteggiatura costituisce sempre un intervento soggettivo del trascrittore o dell'editore con larghissimi margini di arbitrio e di variabilità anche all'interno di un'unica raccolta e perfino di un unico testo. La seconda è che tali segni comportano nella perforazione e in tutte le operazioni successive un aggravio di lavoro un ingombro e un ulteriore rischio di errori che appaiono sproporzionati rispetto alla utilità effettiva dei segni (che nelle elaborazioni è nulla o quasi) e rispetto al loro carattere soggettivo e arbitrario. Pertanto si propone che il criterio ora indicato venga adottato in via definitiva per tutti i lavori successivi.

Il campione prescelto non presentava problemi di segni diacritici speciali. La questione tuttavia dovrà essere studiata con attenzione per i testi appartenenti ad aree dialettali diverse da quella toscana.

Poiché l'esperimento era dedicato esclusivamente ai

problemi della elaborazione dei testi si è ovviamente omissa ogni indicazione esterna sui luoghi di raccolta, gli informatori, i raccoglitori ecc. Ma è naturale che tali indicazioni verranno incluse nei lavori definitivi; e poiché è evidente che l'utilizzazione dei calcolatori elettronici offre enormi facilitazioni nel campo del repertorio e degli indici per località, fonte ecc., è già allo studio il complesso degli accorgimenti tecnici che occorrerà adottare per ottenere tali risultati.

2.22 - Perforazione. La scelta di una fonte manoscritta come la Raccolta Barbi ha permesso di accertare che in casi analoghi (e cioè con fonti manoscritte) non è possibile affidare il passaggio dei testi sulle schede perforate a personale che non sia particolarmente qualificato alla lettura (e spesso decifrazione) di grafie non sempre chiare e di testi che possono allontanarsi molto dall'italiano comune.

Nel caso specifico la difficoltà è stata abbastanza agevolmente superata perché le persone che già avevano

curato l'ordinamento dei manoscritti della Raccolta Barb
(e quindi avevano conoscenza ed esperienza del materia-
le) hanno rapidamente acquisito le abilità tecniche neces-
sarie per operare con le macchine perforatrici ed hanno
eseguito direttamente la perforazione. Ma è evidente che
esiste un più generale problema di addestramento di un
maggior numero di collaboratori qualificati all'uso delle
macchine perforatrici.

Per i testi a stampa (anche dialettali e con abbon-
danza di segni diacritici) il problema non esiste perché
il personale tecnico dei centri specializzati nella esecu-
zione di schede perforate ha acquisito capacità di lettura
meccanica che assicurano anche nel nostro campo (come
già in quello delle più complesse formule matematiche)
la fedeltà più assoluta all'originale.

2.23 - Verifica. Nel presente esperimento il controllo
dei testi perforati e la correzione degli errori sono sta-
ti eseguiti con i metodi correnti, e cioè attraverso la
lettura dei tabulati. E' tuttavia opportuno che nei prossim

mi esperimenti si adottino in via definitiva i sistemi di verifica abitualmente usati per il controllo delle schede perforate (cfr. All. II c, p. X).

2.3 - Stampa dei testi

I testi dei 100 componimenti sono stati stampati secondo il loro ordine di successione nella Raccolta Barbi, e recano l'indicazione del numero che li contraddistingue nel gruppo B, e cioè nella serie dei Canti lirici. Per semplificare i programmi di questo primo esperimento, si è convenuto di ridurre l'indicazione di tale numero alle sole ultime quattro cifre e di omettere l'indicazione della lettera B. In sede definitiva potranno adottarsi altre determinazioni.

I testi sono stati stampati in sole lettere maiuscole per ragioni tecniche contingenti. Nei lavori successivi potranno impiegarsi invece caratteri con la normale distinzione tra maiuscole e minuscole (cfr. All. II c p. II, nota).

Ove occorresse, i testi e tutti gli altri risultati

della elaborazione elettronica potranno essere stampati su schede di speciale formato.

2.4 - Incipitari

Poiché si sono incontrate alcune difficoltà contingenti nella distinzione tra i versi di testa ed i successivi, l'elenco unico previsto nel progetto è stato sdoppiato in due elenchi distinti: quello contenente tutti i versi, quale che sia la loro collocazione nel componimento, e quello dei soli versi di testa. Nei lavori successivi si avrà modo di redigere elenchi unici, con adeguate distinzioni per i versi di testa.

Tuttavia appare opportuno segnalare fin da questo momento che occorrerà adottare una più precisa terminologia per distinguere gli elenchi di tutti i versi da quelli limitati ai versi di testa. In questa fase si è fatto ricorso allo espediente di chiamare Incipitario I l'elenco di tutti i versi e Incipitario II quello dei versi di testa.

Tra i due possibili criteri di ordinamento alfabetico - quello per parole e quello per lettere - è stato qui preferito il primo, a differenza di quanto si è fatto in altre occa-

sioni. La scelta è dovuta al fatto che l'ordine per parole consentiva di evidenziare meglio i raggruppamenti. Ma resta ancora da decidere se questa sia la soluzione migliore. In ogni caso la macchina è in grado di fornire i versi secondo l'uno o l'altro ordine, senza bisogno di accorgimenti speciali.

Per ragioni tecniche del tutto contingenti, e presto superabili le parole precedute da una aferesi precedono nell'ordine alfabetico la lettera A (cfr. All. II c, p. II, nota). L'inconveniente sarà eliminato nel prossimo lavoro.

2.5 - Concordanze

I criteri con cui è stato realizzato e presentato l'elenco delle concordanze sono esposti e discussi nella Nota tecnica del dott. Zampolli (All. II c, pp. III-V).

Qui occorre sottolineare che il problema si presenta con particolari caratteristiche nel nostro campo, data la pluralità linguistica (dialettale) dei testi da sottoporre alla elaborazione, cui si aggiunge una sovrabbondanza di 'grafie'. In altri termini, oltre alle omografie, avremo da af-

frontare quelle che possiamo chiamare le "allo-" o "etero-grafie" (parole di identico valore traslitterate con criteri diversi dai diversi editori) | le "etero-glossie" (e cioè parole di identico significato ma di significante diverso, perché appartenenti a dialetti differenti: per es. 'vorrei' e 'vur-ria').

Per le nostre ricerche è assolutamente indispensabile che l'elaboratore elettronico fornisca anche queste concordanze "allografe" ed "eteroglosse". Sono perciò già allo studio dei criteri di lemmatizzazione che si spera possano riuscire a risolvere felicemente la questione.

Anche se in questo primo esperimento si sono elencate tutte le concordanze, comprese quelle che per la loro abbondanza riescono non significative ed ingombrano inutilmente i tabulati, nel lavoro successivo ci si propone di ridurre gli elenchi alle forme essenziali.

2.6 - Forme elencate in ordine alfabetico e in ordine di frequenza decrescente

Non vi sono osservazioni particolari da aggiungere a quanto è detto nella ricordata Nota tecnica (All. II c, pp. V-VII).

2.7 - Rimari

Per il normale elenco dei versi disposti secondo l'ordine alfabetico delle terminazioni a partire dall'ultima vocale accentata (Rimario I) si rinvia all'All. II c, p. VII.

Viceversa occorrono alcune osservazioni per l'elenco che si è convenuto di chiamare provvisoriamente Rimario II, giacché si tratta di un tentativo nuovo che può produrre risultati particolarmente interessanti ma che richiede di essere studiato in modo approfondito.

Viste le opportunità offerte dagli elaborati, si è ritenuto utile effettuare un primo tentativo di reperimento (e conseguentemente di statistica) delle parole che sono collocate in reciproca corrispondenza di "proposta e risposta di rima". In effetti, quando parliamo di "rima" nei canti popolari, facciamo costante riferimento alle norme codificate e

riconosciute nella poesia culta, ed anche se allarghiamo la nozione di "omofonia" fino a ricomprendervi, oltre alla rima "perfetta" anche la "assonanza" e la "consonanza atona", tuttavia tendiamo ad escludere da quella nozione certe assonanze-limite come 'superbia-erba', 'balia-mamma', e consideriamo come "guasti" relazioni del tipo 'bello-stelle'. Già Silvio Pieri, quasi novanta anni fa, in base a criteri linguistici e in qualche misura statistici, trovava che dovevano considerarsi come rimanti certe parole che viceversa non appaiono tali secondo le norme abituali. Se infatti una particolare relazione (per esempio 'bello-stelle') che non è considerata "rima", né perfetta né imperfetta, ricorresse con particolare frequenza, dovremmo modificare la nostra concezione della omofonia e ammettere che quella relazione, pur se estranea al sistema ufficialmente riconosciuto, è invece "normale" nel sistema della poesia popolare tradizionale. In tal caso diverrebbero "regolari" anche certi componimenti che invece sono considerati "irregolari" o "guasti".

Si tratta ovviamente di ipotesi; ma poiché il problema esiste, e non da oggi, ed oggi esistono anche i mez-

zi tecnici per approntare la materia di studio si ritiene utile il sottoporle a verifica.

I dati forniti dal primo esperimento possono costituire una base di partenza. Occorrerà perciò studiarli con attenzione. Ma per intanto già può dirsi che l'esperimento, per dar frutto, deve essere ampliato in più direzioni.

Occorre innanzi tutto moltiplicare i componimenti da elaborare. E' inoltre necessario affrontare decisamente i componimenti totalmente "irregolari" di cui non si è potuto tener conto in questa prima prova (cfr. All. II c, p. VIII). Bisognerà anche includere nell'esame un altro fenomeno per ora trascurato (cfr. l. c.) e cioè il rapporto di consonanza atona controalterna che esiste in componimenti con numero pari di versi e quello che lega il secondo verso degli stornelli al primo e al terzo. Occorrerà infine ottenere una presentazione diversa degli elementi ed accompagnarli con indicazioni di frequenza.

Ma il punto più delicato della operazione consisterà senza dubbio nella determinazione dei tipi di schema da considerare come punti di riferimento per estrarre le cor

rispondenze di rima. Non è qui possibile affrontare la discussione dell'argomento che dovrà essere trattato a parte e con la dovuta ampiezza. Ma si è ritenuto opportuno farne cenno, a dimostrazione del frutto che l'utilizzazione degli elaboratori elettronici può portare non solo sul terreno puramente repertoriale ma anche su quello di ricerche e sistemazioni concettuali più complesse.

3. Conclusioni

Da quanto si è esposto risulta chiaro che l'utilizzazione degli elaboratori elettronici in materia di poesia popolare non solo è possibile ma è altamente efficace. Si ritiene pertanto necessario procedere ulteriormente sulla stessa strada, estendendo la sperimentazione anche ad altri tipi di documenti e passando al più presto alla realizzazione di imprese non più sperimentali.

Alberto M. Cirese

[poligrafato]