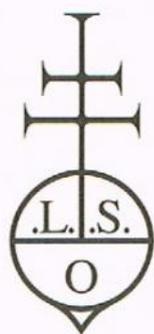


STRUMENTI DI LESSICOGRAFIA LETTERARIA ITALIANA

VOLUME 20

GIUSEPPE SAVOCA

LESSICOGRAFIA
LETTERARIA
E
METODO
CONCORDANZIALE



LEO S. OLSCHKI EDITORE

MM

INFORMATICA E LETTERATURA *

Nell'ambito dei rapporti tra informatica e letteratura non è entrata nell'uso e forse non è ancora stata inventata una paroletta del tipo, ad esempio, di CAD, che è l'acronimo a tutti noto di «Computer Aided Design», e che, appena pronunciata, rende l'idea dei molti servizi che la scienza dei calcolatori può offrire all'architettura e all'ingegneria. A me l'assenza di un termine facile e quasi magico per evocare scenari di lavoro e di interazione uomo-macchina sembra indicativa di una certa difficoltà nell'incontro e nel dialogo tra letteratura e informatica.

Non mi pare sia il caso di affrontare qui la questione terminologica e concettuale di che cosa sia informatica e che cosa letteratura. Ma credo sia utile non dimenticare che in una pertinente definizione di informatica dovrebbero coesistere una componente concreta, tecnologica e una dimensione teorica di tipo logico-scientifico. La letteratura entra in relazione con l'informatica per entrambe le vie, quella tecnico-strumentale e l'altra più propriamente conoscitiva, formale ed epistemologica.

Quanto all'estensione del concetto di letteratura, per ciò che attiene al nostro tema, va osservato che anch'esso è passibile di una prima, sommaria bipartizione tra un versante creativo, concernente la produzione di testi originali (romanzi, versi, ecc.), e un versante analitico qual è quello della riflessione descrittiva e critico-interpretativa sui testi. Nel caso specifico, un argomento di un qualche interesse (ma che sfiorerò solo marginalmente) potrebbe essere quello del posto che l'informatica come motivo contenutistico occupa nei testi letterari contemporanei.

I rapporti tra letteratura e computer sul piano creativo hanno attraversato una fase acutamente innovativa e provocatoria negli anni sessanta, quando alcuni esponenti della neoavanguardia, e particolarmente del Gruppo '63 (cito, per tutti, Nanni Balestrini e Lamberto Pignotti), nella loro rivolta contro il realismo, il sentimentalismo e il linguaggio della tradizione, o di tutto ciò che essi polemicamente identificavano con l'accademia, concepirono progetti di poetiche "tecnologiche" e produssero

* Relazione letta a un convegno organizzato dall'Accademia Nazionale dei Lincei e dalla Fondazione IBM Italia; pubblicata in AA. VV., *Calcolatori e Scienze Umane*, Presentazione di E. Presutti, Prefazione di S. Moscati, Milano, Etaslibri, 1992.

esempi di "poesia elettronica". Quest'ultima si fondava essenzialmente sull'immissione nella memoria del computer di elementi linguistici predeterminati, che venivano poi combinati casualmente in vista della produzione di versi finalizzati alla contestazione di ogni idea alta della poesia. Si vedano, ad esempio, le poesie di Balestrini raccolte in *Ma noi facciamone un'altra* (Milano, Feltrinelli, 1968), e gli scritti teorici di Pignotti in *Istruzioni per l'uso degli ultimi modelli di poesia* (Roma, Lerici, 1968).

Non è un caso che fosse proprio Umberto Eco a elaborare, in relazione a questi esperimenti, il concetto di "poesia virtuale" come risposta polemica alle poetiche classiche che mettevano in primo piano l'autore e il testo. È lo stesso Eco infatti che venti anni dopo nel *Pendolo di Foucault* inserisce, accanto alla voce del narratore, e con effetti di intreccio combinatorio e parodico, frammenti di vari *files* risiedenti nella memoria di un computer.

Tralascio ogni giudizio sui risultati per così dire estetici di simili operazioni e mi fermo un poco sui processi combinatori che dagli anni sessanta in poi affascinarono anche Italo Calvino. È del 1967 una sua conferenza, poi più volte pubblicata, su *Cibernetica e fantasmi*, che ha per sottotitolo *Appunti sulla narrativa come processo combinatorio*. Il giuoco combinatorio, cioè la possibilità che lo scrittore ha di mescolare variamente unità narrative di provenienza e di estensione diverse, è del tutto evidente nel romanzo calviniano del 1979 *Se una notte d'inverno un viaggiatore*.

In sintesi, Calvino in *Cibernetica e fantasmi* ritiene che il «primo narratore della tribù» cominciò esplorando «le possibilità implicite nel proprio linguaggio combinando e permutando le figure e le azioni e gli oggetti su cui queste azioni si potevano esercitare». Riferendosi alle teorie di Propp e di Lévi-Strauss sulle fiabe e sui miti, egli nota che sulla base di un limitato numero di «elementi prefabbricati» si possono ottenere «combinazioni, permutazioni e trasformazioni illimitate».

Passando alle tecniche combinatorie sperimentate nella letteratura scritta, Calvino si chiede se sia prevedibile l'avvento di una macchina che sostituisca il poeta e lo scrittore: «Così come abbiamo già macchine che leggono, macchine che eseguono un'analisi linguistica dei testi letterari, macchine che traducono, macchine che riassumono, così avremo macchine capaci di ideare e comporre poesie e romanzi?».

Paradossalmente egli giudica l'uso fatto in Italia della macchina da parte dell'avanguardia «troppo umano»; e ciò perché «la produzione del disordine» (cioè la «destrutturazione formale» e la «contestazione dei nessi logici abituali») perseguita dai neoavanguardisti rispondeva ancora a un bisogno lirico. La conclusione utopistica di Calvino prefigura l'avvento di una «macchina letteraria», la quale «sentirà essa stessa il bisogno di produrre disordine ma come reazione a una sua precedente produzione di ordine [...] di classicismo».

Il discorso di Calvino è molto articolato e riguarda anche l'incontro tra giuoco combinatorio e concezione psicanalitica della letteratura, l'inconscio sociale e quello individuale. Tuttavia, il nucleo conoscitivo centrale di questa riflessione che va qui

sottolineato attiene all'ipotesi utopica che il creare, come lo scrivere, sia «solo un processo combinatorio tra elementi dati», e che «il risultato poetico sarà l'effetto particolare d'una di queste permutazioni sull'uomo dotato d'una coscienza e d'un inconscio».

Avendo escluso, con la visione di una macchina scrivente che produce testi, la figura dell'autore, Calvino rivaluta la lettura come «momento decisivo della vita letteraria»: «è al lettore che spetta di far sì che la letteratura espliciti la sua forza critica».

Nel romanzo combinatorio per eccellenza che è *Se una notte d'inverno un viaggiatore* c'è una ulteriore riflessione sul rapporto tra creatività e informatica che è sulla linea di quanto discusso in *Cibernetica e fantasmi*. Il narratore infatti si incontra con una giovane che legge i romanzi mediante l'aiuto di un elaboratore; tra l'altro, essa si fa un'idea del loro contenuto dalle liste di frequenza delle parole in essi usate.

Il narratore, con una posizione più morbida, meno utopica, rispetto a quella del conferenziere, in relazione a questi esercizi di lettura assistiti dal computer annota: «Forse anziché un libro potrei scrivere degli elenchi di parole, in ordine alfabetico, una frana di parole isolate in cui si esprima quella verità che ancora non conosco, e dalle quali l'elaboratore, capovolgendo il proprio programma, ricavi il libro, il mio libro».

A questo punto, accettiamo per pura ipotesi giocosa l'utopia o l'ipotesi, pure esse giocose, di Calvino, e vediamo a quali sviluppi paradossali possano portare i procedimenti combinatori messi in opera dalla sua «macchina letteraria». Qualche anno fa è uscito un libro di Stefano Magistretti intitolato *Il poeta elettronico* (Milano, Mondadori, 1984) in cui, partendo da una imprecisata lista di parole, si generano poesie di tre versi sul modello degli *haiku* giapponesi. L'autore avverte che il programma può produrre circa 24.000 miliardi di poesie diverse. Per avere un'idea della sterminatezza di questo *corpus* poetico virtuale si pensi che una vita di cento anni durerebbe circa 3 miliardi e 150 milioni di secondi e che, ammesso che si potesse leggere per tutta la vita una poesia al secondo, ne resterebbero non lette almeno 21.000 miliardi!

In realtà, il semplice immaginare che qualunque testo letterario esistente possa essere riprodotto da un'operazione combinatoria eseguita dal computer porta a prefigurare operazioni umanamente incontrollabili e del tutto irrealistiche. La macchina può produrre per caso una o più combinazioni linguisticamente accettabili, ma solo a patto di produrre un'altra quantità pressoché illimitata di testi del tutto assurdi.

Se noi mettiamo a caso nella memoria del computer le quattro parollette dell'ungarettiano «M'illumino d'immenso», ed eseguiamo un semplice programma di permutazione, per ottenere le quattro parole nella successione voluta da Ungaretti dobbiamo scartare ventitré delle ventiquattro redazioni diverse che la macchina produrrà. Com'è noto anche a chi, come me, non è un matematico, il numero di permutazioni ottenibili con un qualunque numero di elementi è dato dal fattoriale dello

stesso numero (quindi, nel caso, $4 \times 3 \times 2 \times 1 = 24$). Se, ad esempio, prendiamo il primo verso di un sonetto foscoliano come *A Zacinto*, che è composto di sette parole («Né più mai toccherò le sacre sponde»), le permutazioni saranno 5.040. Quante permutazioni ci vorranno per ottenere le circa 100.000 parole della *Divina Commedia* nell'ordine voluto da Dante? Per averne una pallida idea si rifletta sul fatto che le 21 lettere dell'alfabeto italiano danno luogo (secondo un calcolo da altri già eseguito) a un numero di permutazioni composto da 20 cifre (51090942171709440000), e che se una macchina scrivesse ognuna di queste permutazioni in un secondo occorrerebbero più di 1.620 miliardi di anni per scriverle tutte.

È dunque inimmaginabile un'operazione combinatoria eseguita dal computer in vista della ricreazione di un'opera letteraria come prodotto umano. Diverso sarebbe il caso di una letteratura gioco prodotta dal computer, al modo di come si parla di una computer art, nel cui ambito tuttavia, proprio sul terreno della creatività, si è realizzato spesso un incontro felice tra uomo e macchina. Ma questo sarebbe un altro discorso. A me premeva soltanto dimostrare l'assurdità del paradosso o dell'utopia di una letteratura prodotta dal computer a misura d'uomo.

Prima di passare al secondo punto di questo mio intervento (e cioè al tema delle innovazioni che l'informatica può favorire nella critica letteraria), conviene accennare al fatto che la tecnologia elettronica sta modificando, e sempre più modificherà, la produzione letteraria nel senso materiale di attività scrittoria da parte di un soggetto che è l'autore di un qualunque testo. Già molti scrittori confessano di scrivere le loro opere direttamente al personal, tra l'altro perché, com'è noto, intervenire su un testo direttamente sulla pagina che appare a video è più facile e più economico che non correggere o riscrivere un manoscritto o un dattiloscritto. Ed è già venuto il tempo in cui tutti gli editori pretendono che gli si consegnino non più testi su supporto cartaceo, ma esclusivamente in dischetti usciti dal calcolatore. Recentemente sul «Corriere della sera» (20 agosto 1991) Luciano De Crescenzo, ad incoraggiamento degli scrittori ancora diffidenti nei confronti del computer, dava una simpatica definizione del personal come di «quell'aggeggio elettronico che fa sembrare un piacere perfino la correzione». De Crescenzo aggiungeva: «C'è chi [...] ha la pagina ideale posta alla prima stesura, e chi invece se la ritrova solo dopo il ventesimo tentativo. Per questi ultimi il Personal Computer è una specie di lampada di Aladino che li tramuta in scrittori, sempre però, sia chiaro, che abbiano qualcosa da dire».

È evidente che in questa prospettiva di scrittura elettronica stanno cambiando le condizioni della stesura di un'opera che, in linea di ipotesi, esisterà solo in una redazione finale unica (delle correzioni a video infatti si perderà traccia). In conseguenza delle mutate tecniche di scrittura cambierà anche la filologia testuale e, forse, si evolverà verso forme impensate anche la critica letteraria.

Occorre chiaramente ribadire che allo stato non esiste una critica che possa dirsi *tout court* informatica, così come, ad esempio, si parla di una critica psicanalitica o sociologica o stilistica. Ma è, credo, da affermare con forza che se la critica letteraria

vuole aspirare ad avere un futuro adeguato ai tempi, essa deve disporsi ad accogliere sempre di più strumenti e metodi innovativi che solo l'informatica oggi può offrirle.

Sui rapporti tra critica letteraria e computer si è scritto abbastanza e in concreto (cioè sul terreno dell'analisi e dell'interpretazione) si è fatto pure qualcosa, anche se ancora molto poco. Il dato elementare e fondamentale da cui partire, per una rapida messa sul tappeto della problematica in oggetto, è quello che ogni opera letteraria è fatta di parole combinate insieme secondo regole immanenti a un determinato sistema linguistico. Il che comporta che ogni operazione di lettura di un'opera (cioè appunto l'esercizio della critica) non può che essere di natura linguistica. Parlare, in termini saussuriani, è passare dall'asse associativo (o paradigmatico) al sintagma, e cioè, secondo la riformulazione di Jakobson, dall'asse delle selezioni all'asse delle combinazioni. Un parlante, e quindi anche uno scrittore, sceglie delle unità linguistiche in un determinato codice e le attualizza nel discorso, nella pagina scritta. Il destinatario di un messaggio, ascoltatore o lettore, per capire deve compiere il percorso inverso, passare cioè dalla catena sintagmatica del discorso alla comprensione dei paradigmi, delle strutture virtuali di cui esso è realizzazione particolare.

Se la critica si apparenta alla linguistica, e se gli elementi linguistici (sul versante del significante, del significato, ma anche della frequenza) sono dati formalizzabili, appare conseguente che l'informatica, in quanto scienza che tratta informazioni, può e deve avere qualcosa o molto da dire anche nella lettura critica di un testo. La critica è una forma di conoscenza basata su dati, su conoscenze particolari. Il computer, secondo una definizione che Renato Ghiotto ne ha dato in un romanzo caratterizzato da una rappresentazione terrificante della civiltà informatica (*I vetri*), è un «contenitore di conoscenze (e non di conoscenza)». Come tale esso è, in linea teorica, un alleato indispensabile per chiunque abbia a che fare con informazioni di qualsiasi natura. Il critico che voglia servirsi del computer dovrà avere una prospettiva che direi galileiana, se Galilei ha scritto: «Io stimo più il trovar un vero benché di cosa leggiera che il disputar lungamente delle massime questioni senza conseguir verità nissuna».

La critica letteraria nelle sue aspirazioni di fondo ha sempre mirato al livello dei massimi sistemi, e questo forse non è un male. Credo tuttavia che oggi un critico consapevole delle sfide che alla ricerca e alla conoscenza pone l'informatica ci guadagni molto per le sue analisi a partire concretamente dalle piccole verità, dagli elementi concreti e misurabili che ad esso può offrire il trattamento automatico dei testi.

È molto probabile che la rivoluzione informatica abbia sulla letteratura in generale effetti a lungo andare più vasti e più radicali di quelli registrati con l'avvento della stampa o con l'invenzione della macchina da scrivere. Ovviamente, è impensabile che il critico letterario si trasformi in programmatore, ma non appare assurdo ipotizzare che egli possa diventare un utente evoluto del computer, che sappia cioè porre problemi e richiedere adeguate soluzioni allo specialista.

Intanto è da rilevare che più del letterato, diciamo così, creatore, è il critico che dall'uso del computer può conseguire notevoli vantaggi strumentali. La costituzione di banche dati di testi, di sussidi bibliografici, di repertori ecc., e la possibilità di consultare archivi enormi in tempi brevi danno ovviamente al critico l'opportunità di fondare analisi e ipotesi su materiali altrimenti difficilmente maneggiabili con la stessa rapidità, sicurezza e vantaggio da parte di chi non disponga di simili aiuti. Questo sul piano economico-quantitativo; ma è prevedibile che anche a livello qualitativo e conoscitivo possa verificarsi un arricchimento e forse un mutamento nel lavoro critico e nella sua dimensione epistemologica.

Probabilmente, il punto decisivo dell'impatto dell'informatica sulla critica letteraria sarà dato dalla possibilità di registrare e classificare esaustivamente un determinato ambito di elementi testuali e dalla conseguente quantificazione dei fenomeni rilevati. L'interpretazione avverrà di conseguenza sulla base dei dati, ma sarà essa stessa stimolata e indirizzata da ciò che di inedito e di invisibile a prima vista riveleranno i dati stessi. La rivoluzione informatica come caratteristica più evidente delle sue procedure propone la sua capacità di computare, e quindi la sua idoneità a rappresentare e descrivere le strutture quantitative e matematiche di una determinata realtà com'è quella di un testo. La linguistica matematica non è disciplina nata dopo l'avvento del computer, ma è innegabile che solo oggi essa può raggiungere livelli di estensione e di approfondimento capaci di abbracciare universi testuali molto ampi e vari.

L'esistenza di un livello quantitativo dei fatti linguistici non comporta per il critico l'obbligo di trasformarsi ad ogni costo in matematico, ma impone che egli sia consapevole anche di questo aspetto dei dati testuali che vuole comprendere. In questa materia non esiste poi, com'è noto, nessuna grammatica statistica che in astratto ci dica qualcosa sul significato e sull'interpretazione di rilievi quantitativi e frequenze presi isolatamente. Perché i dati ci dicano qualcosa di nuovo e di fondato, è essenziale classificare e quantificare il maggior numero di elementi testuali al fine di arrivare a una carta di identità quantitativa di opere, autori, figure, fenomeni stilistici, ecc. Solo dal confronto tra dati relativi ad autori e testi diversi possono venire indicazioni sulle specificità individuali.

Nel trattamento informatico dei testi letterari la possibilità e la scelta di quantificare certi caratteri e fenomeni dovrebbe inserirsi fra le ipotesi euristiche di partenza, inventate dal ricercatore come mosse di un lavoro e di un giuoco nuovi con i testi. In realtà una critica letteraria assistita dall'informatica non può che partire sempre dal livello testuale, al quale essa si può anche fermare senza giungere alla sfera delle grandi questioni.

Vista *sub specie linguistica*, e schematizzando un poco, la critica letteraria si può esercitare almeno su tre piani, e cioè al livello del significante grafico e fonico, a quello del lessico e a quello della sintassi. Ognuno di questi tre livelli ha ovviamente a che fare, sia pure in misura diversa, da una parte con la statistica, dall'altra con la semantica.

Accenno molto brevemente a qualcuna delle numerose opportunità di descrizione delle strutture significanti di un testo offerte dall'informatica. Si pensi, ad esempio, alla possibilità di computo dei singoli grafemi e fonemi al fine, tra l'altro, di costruire una lista di frequenza dell'uso medio delle lettere dell'alfabeto e dei fonemi con cui confrontare gli usi di un singolo autore. Ricerche di questo genere potrebbero tra l'altro dare un fondamento oggettivo al settore di indagine che va alla ricerca del cosiddetto simbolismo fonetico, nel quale convergono posizioni di creatori di lingua come Mallarmé, e di linguisti e critici come Fonagy e Kristeva.

La registrazione quantitativa di questi elementi potrebbe anche portare qualche luce su un tema affascinante e controverso com'è quello degli anagrammi o, in termini saussuriani, delle parole sotto le parole. Ma, a questo livello, sono da ricordare altri fenomeni come la rima, l'assonanza, l'intonazione, il ritmo, il metro, la punteggiatura, ecc., che sono caratteri tutti formalizzabili e, quindi, catalogabili e quantificabili con procedure informatiche.

Più difficile, senza dubbio, è operare forme di trattamento automatico al livello complesso della sintassi, per la quale tuttavia non mancano studi e procedimenti analitici di sicuro interesse concettuale anche se difficilmente usufruibili, allo stato, dal critico di formazione umanistica tradizionale. Tra i lavori e le interpretazioni particolarmente istruttivi, grazie alla semplicità delle ipotesi di base, penso, ad esempio, ad uno studio di Étienne Brunet sulla *Phrase de Zola* fondato essenzialmente su un solo carattere esterno, quale quello della lunghezza della frase rilevato «dal numero di parole grafiche registrate nell'intervallo di due punteggiature forti». Con uno strumento analitico di estrema semplicità Brunet riesce a disegnare un diagramma interno all'opera di Zola e a individuare una sua specificità sintattica nel confronto con altri prosatori francesi.

Dove tuttavia a me sembra che l'informatica possa dare gli apporti più interessanti al critico letterario è al livello propriamente lessicale (e quindi semantico) del testo. È in quest'ambito che allo stato della questione si può perseguire un incontro tra critica e computer particolarmente produttivo. Se è vero infatti che la letteratura è fatta di parole, nessun critico potrà a giusto titolo ignorare uno strumento come il computer che lavora con le parole, sia pure ridotte a combinazioni di due cifre. È forse proprio la parola ciò che dell'opera letteraria più ignoriamo, ed è certo alla conoscenza della lingua dei testi che l'informatica applicata alla letteratura potrà rendere i servizi più utili.

Ovviamente esistono pressoché per ogni autore di rilievo, ma anche per molti minori, studi specifici sul loro lessico, ma ci mancano in genere repertori, indici e concordanze che possano darci risposte rapide ed esaustive sulla presenza e la storia di questo o quel lemma in un autore, un testo, un'area cronologica, ecc.

La verità nuda e cruda, per chi studia letteratura italiana, è che la lingua letteraria italiana ci è in gran parte ignota a causa della mancanza di adeguati strumenti di documentazione lessicale. Io credo che uno dei punti su cui riflettere, da parte dei

critici italiani, sia la presa di coscienza che il livello testuale che più ci sfugge è proprio quello immediato e macroscopico della parola nella sua realtà lessematica individuale e nelle sue connessioni storico-genetiche.

Ne segue, nella mia valutazione dei fatti e nella mia prospettiva di lavoro, la necessità di una convergenza tra lessicografia, intesa come registrazione esaustiva del lessico di autori e periodi, e semantica, vista come operazione di interpretazione critica fondata sui dati lessicali. Una concordanza, stampata, in dischetto o virtuale (come possibilità di concordare un testo residente nella memoria del computer), mi pare strumento di cui il critico del presente-futuro non potrebbe fare a meno.

Forse non a tutti è noto l'aiuto prezioso che alcuni genetisti riescono a trarre oggi dallo studio dei cognomi registrati negli elenchi telefonici, in vista della ricostruzione delle migrazioni di popolazioni e della indagine sulla distribuzione dei geni nel territorio. Tempo fa, un professore di liceo di fronte a un volume di concordanze ha detto: «sembra un elenco telefonico». Il paragone mi pare corretto e istruttivo.

Immaginiamo per un attimo che cosa sarebbe la nostra vita senza le guide telefoniche. E proviamo anche a immaginare come potrebbe cambiare il lavoro del critico letterario se egli, invece che scavalcare, come spesso fa, la parola, credendo così di potere cogliere, quasi per divinazione, le ideologie, la storia, la psiche o altro, fosse invece costretto a fare sempre i conti con il concreto del linguaggio in cui vivono le ideologie, la storia e tutto il resto.

Ebbene, è auspicabile, e decisivo per il futuro della critica italiana, che si possa disporre di una serie di guide alle parole della letteratura che, così come gli elenchi telefonici unificano l'Italia, coprano gran parte se non tutto il territorio della nostra storia letteraria.

Non è, ovviamente, il caso di sviluppare ulteriormente la metafora, ma non va dimenticato che quando si parla di servizi telefonici si sottintende la possibilità di conoscere un numero o un indirizzo anche senza avere l'elenco stampato a portata di mano, ma semplicemente telefonando. Per analogia, credo che non sarebbe difficile disegnare in Italia, a servizio della ricerca letteraria e linguistica, una serie di nodi collegati in una rete telematica che contenga archivi di testi e di parole, e cioè un tesoro elettronico della lingua e letteratura italiana consultabile da qualunque utente e ricercatore.

È partendo da questo patrimonio e da questo punto che si dovrebbe porre la questione del significato e delle prospettive del lavoro del critico nella società informatica. Lascio ad altri questo compito, limitandomi solo a segnalare i pericoli che uno smisurato accrescersi dei dati può comportare. Il critico informatico rischia di essere bombardato e sommerso dalle informazioni che gli fornirà il computer. Per salvarsi egli dovrà necessariamente elaborare strategie di dominio della complessità testuale finora impensabili. Dovrà sempre tenere ben fermo, credo, il principio che è solo la lettura, cioè la comunicazione con i testi, il punto di partenza e di arrivo della sua ricerca del senso.