

LA TRASCRIZIONE DEI TESTIMONI MANOSCRITTI: METODI DI FILOLOGIA COMPUTAZIONALE

Edoardo Ferrarini *

«O quam gravis est scriptura», si lamenta, nel nono secolo, un anonimo copista della *Lex Romana Burgundionum*, e precisa: «oculos gravat, renes frangit, simul et omnia membra contristat. Tria digita scribunt, totus corpus laborat»¹. Giunto al termine del suo lavoro di trascrizione, egli si abbandona, sottoscrivendolo, a quella che nel latino tornito degli editori degli MGH è una *lamentatio librarii, taedio laboris fatigati*². In essa sono riconoscibili, certo, le formule stereotipate della topica relativa alla pratica scribale come fatica fisica, di tutto il corpo, ma, al di là di questa e simili testimonianze letterarie³, le condizioni spesso disagiati, se non addirittura dolorose, in cui lavoravano gli amanuensi dell'età di mezzo sono fuori discussione. Oggi, nell'era della videoscrittura, copiare un testo non è più una fatica tale da spezzare le reni, piegare il dorso o anchilosare le mani. E tuttavia resta, oggi esattamente come allora, operazione complessa, ricca nelle implicazioni, pesante nelle conseguenze: anche una trascrizione elettronica, dunque, è sempre caratterizzata da una certa *gravitas*, a maggior ragione se riguarda il testo trasmesso da un manoscritto antico, testo che forse non conoscerà neppure il passaggio intermedio della tecnologia della stampa, ma transiterà direttamente dalla pergamena alle memorie labili ed impalpabili dei calcolatori elettronici.

Per non affrontare alla leggera il compito di questa *gravis scriptura*, occorrerà, quindi, porsi, in via preliminare, almeno tre domande: perché è utile una trascrizione digitale di un manoscritto, quando è utile, come dev'essere fatta perché sia utile. A vantaggio della chiarezza espositiva, preferisco qui invertire l'ordine razionale delle questioni, partendo, cioè, dal come, toccando il quando e concludendo sul perché.

* edoardo.ferrarini@univr.it
Università degli Studi di Verona.

¹ Blume 1863, p. 589.

² *Loc. cit.*

³ Cfr. le molte raccolte e studiate da Wattenbach 1896.

1. Il computer e l'uomo (è un'ovvietà) non usano un codice comunicativo condiviso. Il problema della memorizzazione informatica di un qualsiasi testo, dunque, è sempre un problema di codifica, poiché si tratta di tradurre quel testo qualunque in modo che sia leggibile alla macchina, di trasporre l'informazione testuale, come si dice, in *Machine Readable Form* (MRF)⁴. Nella definizione data da Fabio Ciotti, che possiamo prendere a riferimento, «la codifica è la rappresentazione dei testi su un supporto digitale in formato utilizzabile da un elaboratore»⁵. Essa, in altre parole, è la condizione necessaria e sufficiente perché il dato testuale sia manipolabile attraverso i trattamenti automatici ed il modello computazionale proprio dell'informatica in quanto scienza, ossia dell'informatica considerata a prescindere dalle sue macchine⁶. Ad onor del vero, la memorizzazione del testo prevede sempre un doppio ciclo di codifica: il primo nel codice dei caratteri alfanumerici, il secondo che, in base ad una specifica tavola di corrispondenza, denominata *Coded Character Set* (e ve ne sono molteplici), associa i singoli caratteri ad altrettante stringhe ordinate, composte dalla successione dei due numeri del codice binario⁷. Sul secondo ciclo di codifica non ci soffermeremo oltre: poiché si dà corrispondenza biunivoca tra ogni simbolo alfanumerico del set di caratteri scelto e la notazione numerica binaria, questo grado zero della codifica resta, per intendersi, del tutto interno alla macchina. Per quanto concerne il primo ciclo, se ci riferiamo non alla codifica di un testo qualunque, ma di quello recato da un testimone manoscritto di un'opera letteraria (e solo a questo mi riferisco d'ora in avanti), le condizioni necessarie non sono più sufficienti⁸. Una trascrizione dal valore scientifico credo

⁴ Per un'introduzione generale e di carattere manualistico si potrà vedere il capitolo *I dati e la codifica*, in Fadini - Savy 1999, pp. 81-92.

⁵ Ciotti 1997, pp. 55-85 (p. 56). Come introduzione specifica al tema della codifica elettronica dei testi in ambito umanistico potrà essere utile, poi, Id. 1995, pp. 147-78.

⁶ È, infatti, largamente diffusa la confusione tra la scienza informatica, da una parte, che possiamo far risalire, nei suoi momenti fondativi, alla metà del XIX secolo (con la logica algebrica delle proposizioni di George Boole) o addirittura al IX secolo (con il concetto di algoritmo sviluppato in ambito arabo-persiano), e, dall'altra parte, i suoi prodotti, le macchine dell'informatica. Un buon antidoto, in questo senso, si rivela la lettura di Orlandi 1987, pp. 1-38. Sull'importanza dell'aspetto algoritmico e computazionale delle ricerche nel campo dell'informatica umanistica, o *Humanities Computing*, ho creduto d'insistere in Ferrarini 2003, pp. 227-56.

⁷ Adamo 1987, pp. 39-63.

⁸ Il problema è molto ben inquadrato, dal punto di vista teorico, in Orlandi 1998, pp. 349-60.

debba servirsi di un modello di codifica che rispetti almeno queste quattro ulteriori condizioni:

- a) la codifica deve essere documentata, ossia dichiarare esplicitamente la responsabilità della trascrizione ed i criteri seguiti;
- b) deve essere al massimo grado portabile, cioè indipendente dal sistema operativo, svincolata dall'*hardware* e dal *software*, riutilizzabile in processi computazionali diversi;
- c) deve essere esaustiva, quindi tendere ad una rappresentazione la più completa possibile della fonte testuale;
- d) deve essere normalizzata, ossia regolata su norme che rappresentino in qualche modo uno standard riconosciuto dalla comunità scientifica internazionale o da una sua parte significativa.

Il primo vincolo, quello della documentazione⁹, esclude, a mio avviso, molte delle realizzazioni reperibili oggi in rete (mi riferisco sempre e solo a trascrizioni di manoscritti), peraltro scarse nel numero e riguardanti per lo più lavori in corso d'opera; parecchie di queste non sono accompagnate da alcuna dichiarazione sulle scelte di codifica effettuate dai trascrittori e resta da credere che i manoscritti siano stati codificati direttamente in HTML (*HyperText Markup Language*), il sistema oggi maggiormente in uso per la creazione delle pagine *web*, cioè una delle peggiori scelte che si possano fare in sede di codifica scientifica (in quanto trattasi di marcatura procedurale e non dichiarativa)¹⁰. Molti di questi lavori, inoltre, meriterebbero semplicemente di rientrare nella categoria dei «giochi di prestigio», che molte volte ha sollecitato l'ironia di Tito Orlandi¹¹.

La condizione della portabilità trova attuazione all'interno dello stesso sistema fra applicazioni diverse e, *ad extra*, fra sistemi operativi diversi, nonché fra le tecnologie oggi in uso e quelle a venire (onde evitare l'obsolescenza o, peggio, la perdita del-

⁹ Ciotti 1997, pp. 62-3.

¹⁰ *Infra*, § 2.

¹¹ Lo studioso, a proposito di molti presunti studi di informatica umanistica, recentemente così si esprimeva: «Per lo più contengono la descrizione di lavori in corso d'opera [...], nei quali i problemi di carattere teorico sono schiacciati sullo sfondo di un ambiente di pura tecnica [...]. È raro trovare riferimenti alle soluzioni informatiche precedenti, utilizzate per problemi analoghi [...]. Tutto questo contribuisce a rendere un cattivo servizio alla disciplina cui gli studi si riferiscono e giustificano lo scetticismo», e poco più oltre: «gli scettici sono spesso i migliori studiosi, che [...] riconoscono a fiuto la differenza fra uno studio serio ed una specie di gioco di prestigio»; Orlandi 2002, pp. 105-9 (pp. 107-8).

le informazioni)¹². Se portata al massimo grado (come credo sia giusto fare) ha delle implicazioni di non poco conto: esclude, infatti, la possibilità di servirsi di un comune *word processor*, cioè di un pacchetto applicativo per la videoscrittura, che, basato su una codifica proprietaria, inserirà all'interno del testo caratteri di controllo invisibili all'utente, arrivando ad un livello tale di manipolazione da rendere il testo inservibile per un impiego al di fuori del programma stesso. L'attenzione alla portabilità obbliga, poi, alla scelta di un set di caratteri il meno esteso possibile; il più adatto allo scopo resta il vecchio codice ASCII (*American Standard Code for Information Interchange*), noto anche come ISO 646, letto allo stesso modo da ogni piattaforma *hardware* (compreso l'ormai vetusto DOS)¹³. Esso contempla unicamente le lettere minuscole e maiuscole, le cifre arabe e pochi segni diacritici: dunque, niente lettere accentate, niente cediglie, niente parentesi quadre o graffe. Ci si illude ancora, in qualche occasione, di mettersi al riparo da problemi di compatibilità usando solo i cosiddetti 'caratteri da tastiera', ma si dimentica che la tastiera è una semplice unità di *input* dei dati, oltretutto con caratteristiche diverse da nazione a nazione e da produttore a produttore, e non un set di caratteri valido.

Tendere ad una rappresentazione la più completa possibile della fonte, cioè del manoscritto, nel rispetto del requisito di esaustività di una codifica scientifica¹⁴, obbliga anzitutto alla consapevolezza che il nostro manoscritto rappresenta in sé un sistema complesso, anzi, che esso è all'intersezione di più sistemi linguistici e ideologici. Ricorrendo alla nota teoria di Cesare Segre, possiamo definirlo come un diasistema, un compromesso tra il sistema del testo e quello del copista¹⁵; le cose si complicano, inoltre, se consideriamo che, nel procedere della tradizione, i sistemi dei copisti si sono sovrapposti più volte l'uno all'altro e noi stessi, nel momento in cui prendiamo possesso di quel testo, ad esempio trascrivendone un testimone, contribuiamo inconsciamente a generare un nuovo diasistema. E poiché il compromesso fra questi sistemi è globale, quanto

¹² Ciotti 1997, pp. 63-4.

¹³ La presente raccomandazione, ormai quasi esclusivamente imposta dall'attenzione ad una piena retro-compatibilità, è destinata ad essere presto superata dal convergere delle tecnologie attuali, compreso l'XML, verso la nuova tavola di caratteri *Unicode*, riconosciuta anch'essa dall'*International Organization for Standardization* (ISO).

¹⁴ Ciotti 1997, pp. 61-2.

¹⁵ Segre 1979, pp. 45-9.

all'estensione, ed inestricabile, quanto agli effetti, in sede di codifica occorrerà, come insegna Tito Orlandi, «riconoscere ed elencare tutti gli elementi che si trovano nel manoscritto, che concorrono a fornire in quel manoscritto la rappresentazione del testo»¹⁶. Come farlo, avendo a disposizione, per quanto detto prima, nient'altro che pochi caratteri? La strategia migliore è quella di servirsi di un *generic markup language*, cioè di una codifica che utilizzi determinate combinazioni di segni fra quelli propri del codice adottato, per rappresentare tutta una serie di ulteriori caratteristiche del testo (come una 'e' caudata o una parola aggiunta in interlinea, per cominciare a fare qualche esempio). Un vero e proprio apparato interpretativo, dunque, costituito da segni complessi o supersegni (i marcatori), un metatesto chiaramente individuabile e distinto dalla serie dei caratteri alfabetici che compongono il testo, nonché automaticamente eliminabile (deve trattarsi, dunque, di un intervento reversibile); un apparato interpretativo in grado di esplicitare e rappresentare tutte le informazioni rilevanti contenute nel testo (*non omnes textus, sed textus toti*). Però non tutte le informazioni, ché non sarebbe possibile pena un indescrivibile accumulo di segni, ma tutte le informazioni rilevanti, tutti i tratti significativi, in base ad un'operazione di classificazione e selezione in funzione del nostro determinato punto di vista. All'atto stesso della codifica, infatti, deve essere chiaro, a grandi linee, il tipo di trattamento cui il testo è destinato. Felice, al proposito, la sintesi di Dino Buzzetti: «Le forme di rappresentazione informatica che si scelgono inizialmente condizionano il tipo di operazioni analitiche che si vogliono fare sull'informazione rappresentata»¹⁷; in altre parole, non si danno procedimenti computazionali su quegli elementi del testo che non siano stati considerati rilevanti all'atto della codifica.

Per quanto riguarda l'ultima condizione che ho posto per un esito dal valore scientifico (l'ho chiamata normalizzazione)¹⁸, lo standard in maggior misura condiviso dalla comunità scientifica internazionale interessata al trattamento digitale dei dati umanistici è quello espresso nelle *Guidelines for Electronic Text Encoding and Interchange*¹⁹, attualmente giunte alla versione

¹⁶ Orlandi 1997, pp. 159-72 (p. 169).

¹⁷ Così Buzzetti interveniva in uno degli spazi di discussione del Convegno Leonardi 1998, pp. 55-6.

¹⁸ Ciotti 1997, pp. 64-5.

¹⁹ *Guidelines* 2002. Per maggiori dettagli su motivazioni, prospettive e storia di quest'importante iniziativa scientifica internazionale, cfr. *Text Encoding Initiative* 1995.

P4²⁰, uno schema di codifica che utilizza l'*Extensible Markup Language* (XML)²¹.

Il mio intendimento, a questo punto, è stato quello di accertare se per ogni elemento rilevante, rintracciabile in un manoscritto ai fini della sua memorizzazione elettronica, esistesse o meno una strategia di codifica corrispondente nello schema TEI. Laddove, poi, mi sono trovato di fronte a possibilità alternative di codifica di uno stesso fenomeno²², ho cercato di operare la scelta di volta in volta più opportuna. Da una parte, per questa verifica, il riferimento obbligato è stato quello all'*additional tag set*, al *set* opzionale di marcatori, descritto nel capitolo 18 delle *Guidelines (Transcription of Primary Sources)*²³; dall'altra, ho elaborato un'originale griglia di classificazione che comprende quelli che sono, a mio avviso, gli «Elementi rilevanti del diasistema testuale rappresentato da un testimone manoscritto di un'opera letteraria ai fini della sua codifica digitale». Nel costruire questo schema, certo passibile di correzioni, integrazioni e modifiche, mi sono stati d'aiuto, *in primis*, gli *Appunti per una semiotica della trascrizione nella procedura ecdotica computazionale* di Raul Mordenti²⁴, l'unico studio, a mia conoscenza, che distingue correttamente il momento dell'individuazione degli elementi segnici rilevanti da quello della precettistica sulla maniera di rendere tali elementi in sede di edizione. In questo secondo contesto e, dunque, totalmente dipendenti dalla tecnologia della stampa (pensata come unica forma possibile di edizione), si muovono, invece, le norme per la trascrizione di testi letterari dettate in diverse occasioni da François Masai, il padre nobile dell'edizione diploma-

²⁰ Una nuova revisione generale delle *Guidelines* TEI, denominata P5, è tuttora in corso, ma prossima alla pubblicazione definitiva. Tra le novità più rilevanti, l'introduzione di un *set* di marcatori specifico per le descrizioni dei manoscritti.

²¹ Il manuale di riferimento del linguaggio è quello di Goldfarb - Prescod 2003⁵. Dal momento che XML deriva da SGML (*Standard Generalized Markup Language*), è con esso compatibile e ne rappresenta uno dei più fortunati esiti, non ha perso utilità, dello stesso Goldfarb 1992²; raccomandabili, soprattutto per l'orientamento più operativo, le guide di Bryan 1988, e di Herwijnen 1994². Della codifica di manoscritti mediante SGML si è occupato Robinson 1994.

²² È fra gli obiettivi principali delle *Guidelines* TEI, d'altronde, quello di offrire «alternative encodings for the same text features», *Guidelines* 2002, p. 7.

²³ *Ibid.*, pp. 453-79. Un'es. molto documentato della sua applicazione lo presenta Hansen 2002, pp. 72-82.

²⁴ Mordenti 1987, pp. 85-124; ora anche in Id. 2001, pp. 53-82.

tica (1950)²⁵, da D'Arco Silvio Avalle (1972)²⁶ ed, infine, da Francesco Mazzoni (1998)²⁷. Maggiori sollecitazioni, onestamente, mi sono venute dalla ricca precettistica elaborata con continuità negli anni dai diplomatisti per l'edizione delle fonti documentarie (con continuità, ma anche con «vero e proprio accanimento», ha scritto qualcuno di loro)²⁸: i nomi sono quelli di Pratesi²⁹, Tognetti³⁰, Bartoli Langeli³¹; aggiungo la *Diplomatique médiévale* uscita nella collana di Brepols *L'atelier du médiéviste*³².

²⁵ Masai 1950, pp. 177-93.

²⁶ D'Arco S. Avalle 1978².

²⁷ Mazzoni 1998, pp. 145-72; mi riferisco, in particolare, all'appendice: *Principi di trascrizione del testo e delle rubriche della Commedia. Regole per la codifica di compendi, lacune, correzioni, aggiunte, revisioni, ecc.*, pp. 154-9.

²⁸ Bartoli Langeli 1993, pp. 116-31 (p. 116).

²⁹ Pratesi 1957, pp. 312-33; ora anche in Id. 1992, pp. 7-31.

³⁰ Tognetti 1982.

³¹ Bartoli Langeli 1993, pp. 128-31.

³² Guyotjeannin 1995²; mi riferisco, in particolare, alla sezione 9.B: *Règles et conseils d'édition*, pp. 402-11.

**ELEMENTI RILEVANTI DEL DIASISTEMA TESTUALE RAPPRESENTATO
DA UN TESTIMONE MANOSCRITTO DI UN'OPERA LETTERARIA
AI FINI DELLA SUA CODIFICA DIGITALE:**

1. SEGNI DI SCRITTURA

1.1 Segni alfabetici e numerali

1.2 Segni abbreviativi

1.2.1 Tachigrafie

1.2.2 Compendi³³

1.3 Segni non alfabetici e non abbreviativi

1.3.1 Figure

1.3.2 Segni misti di scrittura e figura³⁴

1.4 Segni di interpunzione

1.5 Segni di correzione della prima scrittura³⁵

1.5.1 Interventi di soppressione

1.5.1.1 *Rasure*

1.5.1.2 *Espunzioni*

1.5.1.3 *Depennamenti*

1.5.1.4 *Biffature*

1.5.2 Interventi di integrazione

³³ Lo scioglimento dei compendi, in sede di trascrizione, è sempre causa di grossi crucci (già per il copista medievale). Consigliava di non porsi soverchi problemi l'amabile storia raccontata da Aulo Gellio (*Noctes Atticae* 10, 1, 7): volendo Pompeo consacrare un'epigrafe a sé stesso nel tempio della Vittoria, dal momento che gli richiedevano se incidere *consul tertio* o *consul tertium*, per indicare la sua carica, e le persone più colte della città, interpellate sulla questione, si dividevano in due partiti più o meno equivalenti, egli si risolse, infine, di domandare consiglio a Cicerone. Questi, però, gli rispose più da politico, che da grammatico. Non volendo inimicarsi nessuna di quelle persone ragguardevoli, infatti, consigliò Pompeo di abbreviare: ... *ut neque «tertium» neque «tertio» scriberetur, sed ad secundum usque «t» fierent litterae, ut, verbo non perscripto, res quidem demonstraretur, sed dictio tamen ambigua verbi lateret*. Trovo, poi, molto persuasiva, al riguardo, l'osservazione di Bartoli Langeli, per cui: «Il fatto è che scrivendo *īperium*, lo scrivente non vuole significare né *imperium* né *inperium*; vuole significare *īperium*»; Id. 1993, p. 129 (punto b).

³⁴ È il caso, ad es., dei capilettera.

³⁵ Al di là del fatto che siano tracciati dalla prima mano o da mani di copisti successivi.

1.5.2.1 *Aggiunte interlineari*³⁶

1.5.2.2 *Aggiunte marginali*

1.5.3 Interventi di sostituzione

1.5.3.1 *Interventi di soppressione e integrazione*

1.5.3.2 *Riscritture*³⁷

1.5.3.3 *Interventi sull'«ordo verborum»*

1.5.4 Soppressione di interventi

1.6 Segni glossematici³⁸

1.6.1 Glosse

1.6.2 Segni di esecuzione³⁹

1.6.3 Sottolineature e riquadrature

2. RILEVANZE CODICologiche

2.1 Descrizione del manoscritto

2.2 Disposizione della scrittura

2.3 Identificazione e successione delle mani di scrittura

2.4 Guasti meccanici

2.5 Spaziature anomale e intenzionali⁴⁰

³⁶ «La distinzione fra una lettera da considerare parte di una linea da una da considerare in interlinea dipende dalla collocazione spaziale al di sopra del rigo. Ma molte lettere (per la fretta dello scrivente, per anomalie del materiale scrittorio, ecc.) sono in effetti al di sopra del rigo senza per questo poter essere considerate in interlinea. Diremo dunque che una lettera si considera in interlinea quando è (oltretutto) evidente l'intenzione dello scrivente di porla appunto in interlinea», Orlandi 1987, p. 35.

³⁷ Si definiscono tali gli interventi che, per effettuare la sostituzione, utilizzano un qualche tratto della lettera o della parola da sostituire.

³⁸ Questi sono distinti dai restanti segni di scrittura non tanto perché tracciati da mani diverse dalla prima, bensì perché non ascrivibili al testo vero e proprio, ma costituenti una sorta di commento al testo. Sono segni glossematici: le glosse vere e proprie (a margine o in interlinea), i segni di esecuzione, le sottolineature e le riquadrature di porzioni di testo (spesso derivanti da lettura personale).

³⁹ Devo l'integrazione del punto 1.6.2, assente nella versione dello schema presentata durante i lavori del Convegno, alla competenza e alla cortesia di Paolo Chiesa.

⁴⁰ Non si tratta (è il caso di precisarlo) delle lacune, l'esistenza delle quali può emergere soltanto da un confronto con altri testimoni o con la tradizione indiretta.

3. RILEVANZE DEL CONTENUTO

3.1 Antroponimi

3.2 Toponimi

3.3 Date ed eventi databili

3.4 Titoli

3.5 Brani riportati

3.5.1 Discorsi diretti

3.5.2 Citazioni esplicite e/o letterali

3.6 Parole in lingua diversa

3.7 Struttura strofica

La corrispondenza fra gli elementi rilevanti di un manoscritto, così rubricati, ed i metodi di codifica compatibili con le *Guidelines* TEI, è mostrata in una serie di diapositive contenute nel CD-ROM allegato al presente volume e non ho ritenuto né semplice, né utile riprodurre qui il contenuto. Una delle ultime diapositive mostra, poi, un esempio di come si possano inserire, già nel *file* della trascrizione, eventuali emendamenti volti a correggere il testo tradito. Tutti gli esempi che ricorrono nelle *slides* sono illustrati tramite riproduzioni fotografiche di parti del manoscritto n. 325 della Biblioteca Città di Arezzo (saec. XIV in.)⁴¹, di cui altrove mi sono già occupato⁴². A conclusione del percorso, si troverà uno *specimen* concreto di *file* XML, codificato secondo i criteri esposti: si tratta delle prime ventiquattro strofe del ritmo *Heu, heu, mundi vita*, riportato ai fogli 34v.-36v. del manoscritto

⁴¹ Mi corre l'obbligo qui di ringraziare pubblicamente la Direzione della Biblioteca Città di Arezzo per il permesso accordatomi di fotografare parte del codice. Queste le descrizioni del manoscritto, nell'ordine in cui sono apparse: *Inventari dei manoscritti delle biblioteche d'Italia*, fondati e diretti da Mazzatinti, 1896, p. 228, n. 325 (la sommaria descrizione dovuta a G.F. Gamurrini, allora bibliotecario ad Arezzo, assegna il codice al XV secolo); Magheri Cataluccio - Fossa 1979, pp. 224 (n. A 224 II), 387, 467, 486 (considera il codice del XIV secolo e vi riconosce i canoni dello *scriptorium* camaldolese di Fontebono); Cenci 1988, pp. 50-71 (finalmente si considera la sottoscrizione dell'ultimo copista, *Donatus de Muriano*, che riporta la data 1330, o forse 1335); *Catalogo di manoscritti filosofici* 2000, pp. 43-4, n. 25 (con più di una svista).

⁴² Ferrarini 2005, pp. 310-28.

in questione⁴³. Basterà applicare più fogli di stile a detto documento elettronico, per ottenere diverse rese a video: nel primo esempio di trasformazione, si noteranno gli scioglimenti delle abbreviazioni in corsivo (tra parentesi se dubbi), gli interventi di soppressione tra parentesi uncinate, le barre verticali ad indicare la fine delle righe nel codice; nel secondo esempio, invece, sono rispettate le linee di scrittura del manoscritto e le barre verticali passano ad indicare la fine dei versi del ritmo; nell'ultimo esempio, infine, il più elaborato graficamente, i diversi fenomeni sono evidenziati mediante il ricorso a colori differenti⁴⁴. La trasformazione del *file* XML in HTML è possibile attraverso diverse tecnologie, tra cui l'XSL (*Extensible Stylesheet Language*), di cui ho scelto in questo caso di avvalermi⁴⁵.

⁴³ Chevalier 1892-1921, nn. 5270, 7730, 7732; Walther 1969², n. 7752; Bloomfield 1979, n. 2329. Il ritmo, una interminabile e lamentosa apostrofe indirizzata alla *vita mundi*, dovette godere di molta fortuna, almeno fino al XV secolo, se ho potuto censirne ben 27 testimoni manoscritti; questi riportano l'opera come *Versus de contemptu mundi* o *de fragilitate humanae vitae*, o ancora con il titolo di *Carmen rithmicum de vita humana*. Conferma questa fortuna anche il francescano Salimbene de Adam, che trova il modo di inserirne una versione nelle pagine della sua *Cronica*, cfr. Scalia 1998-1999, pp. 872-8 [460c-461b]. Tra le molte edizioni del testo, spesso parziali e sempre fondate su uno o due testimoni, mi limito a segnalare quella di Rudolf 1957, pp. 36-7, 422-3. Queste, in sintesi, le attribuzioni tentate dai copisti o dagli studiosi moderni: Gregorio Magno, Bernardo di Chiaravalle, Bernardo di Cluny, Ugo Primate d'Orléans, il cardinale Latino Malebranca, Iacopone da Todi, Pietro Gonnella. Il fatto che almeno un paio dei manoscritti che ho rintracciato risalgano, prestando fede alle informazioni dei cataloghi, all'XI secolo, esclude, di necessità, la più parte di queste attribuzioni.

⁴⁴ Dev'essere chiaro, in ogni caso, che il problema della visualizzazione o *display* del testo codificato riveste un'importanza del tutto secondaria (si tratta unicamente di una modalità di *output*, alternativa o concorrente a quella della stampa). La trascrizione digitale è destinata, invece, primariamente alla computazione, cioè all'elaborazione che può essere fatta di questi dati mediante processi computazionali (cfr. *infra*, § 3). In questo senso, dunque, è decisamente fuorviante opporre la trascrizione elettronica all'immagine digitale dello stesso testo, che è ovviamente insuperabile quanto a fedeltà all'originale. Non è questo, infatti, l'obiettivo perseguito, ma quello di una codifica *Machine Readable Form*, un modo, in sostanza, per rendere in grado le macchine di 'leggere' questi stessi testi ed elaborarli secondo le nostre necessità, ma secondo le loro modalità.

⁴⁵ La presente ricerca è stata condotta nell'ambito delle attività del Laboratorio di Informatica Umanistica (LAB.I.U.M), attivo presso il Dipartimento di Linguistica, Letteratura e Scienze della Comunicazione dell'Università degli Studi di Verona e diretto da Antonio De Prisco. Ringrazio, inoltre, il collega Stefano Minozzi, con cui ho proficuamente discusso svariati aspetti del presente lavoro.

2. Sin qui il discorso è rimasto centrato sui metodi, sul come debba essere condotta la trascrizione digitale di un manoscritto, perché possa risultare utile; ora affrontiamo più rapidamente le altre due questioni rimaste aperte. Innanzitutto: quando è utile una trascrizione di questo tipo?

Sempre e senza dubbio alcuno, credo, nel caso che la filologia chiama *codex unicus*, ossia in presenza di tradizioni a testimone unico. In questa condizione, infatti, la metodologia che ho descritto permette di integrare perfettamente tutti e quattro i modelli fondamentali di edizione⁴⁶: quella cosiddetta meccanica, che si configura come un facsimile o una riproduzione tecnica (in genere fotografica) dei manoscritti; l'edizione diplomatica *stricto sensu*⁴⁷; quella diplomatico-interpretativa, in cui a testo figura la forma realmente attestata, anche se chiaramente erronea, ed in apparato l'emendamento dell'editore⁴⁸; ed, infine, l'edizione critica, che è possibile anche per un testo trådito da un unico manoscritto, correggendolo non tanto *ope codicum* (operazione evidentemente impossibile), quanto *ope ingenii*, e restituendolo in una *facies* più vicina all'archetipo o alla supposta forma originale. Nell'esempio che ho proposto, le fotografie digitali del manoscritto, la marcatura strettamente diplomatica o imitativa di ogni suo segno rilevante, la memorizzazione tanto della forma realmente attestata, quanto di eventuali emendamenti dell'editore, tutto è racchiuso, in termini di informazione, nel *file* contenente il frutto della codifica scientifica (cioè documentata, portabile, esaustiva e normalizzata) e tutto può derivare da esso, nei termini di differenti *output*. La distinzione netta tra il *markup* dichiarativo (queste lettere della parola sono abbreviate) ed il *markup* procedurale (visualizza in questo modo le lettere frutto di scioglimento), la stabilità del primo e l'intercambiabilità del secondo, permettono di superare l'opposizione fra le pratiche di trascrizione più imitative o conservative e quelle più innovative o interpretative. Permettono,

⁴⁶ Sui vari tipi di edizione possibili e sulle loro reciproche differenze resta molto chiara la spiegazione nel testo, ormai classico, di Balduino 1989³, pp. 35-53. Per la definizione della natura e delle caratteristiche dell'edizione diplomatica, si veda anche il più recente Chiesa 2002, pp. 167-9.

⁴⁷ Nel dibattito all'interno della comunità scientifica sul valore dell'edizione diplomatica, è apparso decisivo l'intervento di Masai 1950, le cui indicazioni metodologiche furono subito messe in pratica nella preparazione di due importanti edizioni, condotte con metodologia diplomatica ed apparse a breve distanza l'una dall'altra: *Regula Magistri* (si veda, però, anche la recensione, non del tutto convinta, di Bieler 1955, in «Speculum») e Delaissé 1956.

⁴⁸ Balduino 1989³, p. 41, e Chiesa 2002, p. 168-9.

altresì, di superare le divergenze, troppo a lungo durate, sull'uso di questo o di quell'altro tipo di parentesi per indicare questo o quell'altro fenomeno⁴⁹: si tratta, in fondo, di un semplice cambio di foglio di stile.

E quando abbiamo a che fare con tradizioni a testimone plurimo, la vera palestra della critica testuale, è utile trascrivere secondo questo metodo tutti i testimoni manoscritti di un'opera letteraria? Tra il censimento delle testimonianze dirette ed indirette e la *collatio codicum*, i manuali di critica del testo non contemplano la necessità di trascrivere alcunché. Qualora proprio non vi fossero edizioni precedenti dell'opera, beninteso, si dovrà necessariamente trascrivere un codice, scelto con motivazioni in larga parte arbitrarie, per servirsene come esemplare di collazione; ma in tutti gli altri casi, è consigliato il ricorso ad una delle edizioni precedenti, per rendere più comodo il lavoro. A quel punto, secondo il Balduino, «basterà segnare i risultati della *collatio* sui margini» dell'edizione a stampa, oppure si potrà «interfogliare» quest'ultima⁵⁰; il West, più rispettoso dei libri, confessava, invece, di utilizzare abitualmente allo scopo un taccuino distinto (*separate notebook*)⁵¹. Se anche uno ed un solo testimone è da trascrivere (con quali norme, poi, non mi risulta sia mai stato detto), l'attenzione prima va comunque alla costruzione di quel repertorio delle varianti che formerà la base di lavoro per la *recensio*. Voglio dire con questo una cosa semplicissima, ma mai sufficientemente sottolineata: è l'informatica che ci fa intravedere la possibilità di procedere a trascrizioni complete di tutti i testimoni, è l'informatica che le pretende per poter poi applicare su questi dati dei processi computazionali.

3. L'informatica ci indica una possibilità: la domanda cruciale è perché valga la pena di coglierla⁵². Le obiezioni all'uso del computer nella critica del testo, infatti, non hanno mai messo in discussione, a ben guardare, la possibilità di ottenere qualche risultato attraverso metodiche computazionali; hanno, invece, più volte ripetuto che la

⁴⁹ Ricordo, per fare un solo es., l'opposizione tra la pratica dei filologi e quella dei diplomatisti nell'uso delle parentesi uncinata, riconosciuta addirittura come «divertente» da Bartoli Langeli 1993, p. 123.

⁵⁰ Balduino 1989³, pp. 133-4 (nota 30).

⁵¹ West 1991, pp. 67-8.

⁵² Ai molti aspetti del rapporto fra metodologie e soluzioni informatiche, da una parte, e problemi della critica del testo, dall'altra, sono dedicati gli studi di Perilli 1995 e Mordenti 1987, con ampia bibliografia generale alle pp. 151-72.

loro introduzione nella pratica editoriale è molto costosa, molto più gravosa del lavoro manuale ormai collaudato, perciò assolutamente antieconomica.

Comincia D'Arco Silvio Avalle, molto tempo prima dell'introduzione dell'*home computer*, il quale vede come svantaggiosi i casi in cui «le varianti di forma sono talmente numerose da obbligare la macchina a trascrivere praticamente quasi tutto il testo dei singoli manoscritti»⁵³. Continua il West, parlando di *very considerable trouble*: «L'incomodo veramente considerevole implicato dal sottoporle [le varianti dei testimoni] ad un elaboratore non pare giustificato»⁵⁴. E ancora, Balduino: «Nella disarmante laboriosità di programmazione e utilizzazione sembrano da ricercare le ragioni per cui l'automazione della *collatio* è ancora ferma alla fase sperimentale»⁵⁵. Rincarà la dose, se possibile, Pratesi, riferendosi alla digitalizzazione delle fonti documentarie: «Risposte soddisfacenti da un trattamento improntato all'informatica si potrebbero conseguire soltanto con una memorizzazione dei documenti *in extenso*: ma una soluzione di questo genere sarebbe estremamente gravosa in fase di registrazione ed eccessivamente complicata in fase di consultazione, ossia del tutto negativa in termini 'economici'»⁵⁶. Da ultimo, in ordine di tempo, si aggiunge Paolo Chiesa, per il quale ad un certo punto «l'intervento umano si rende nuovamente necessario; esso è anzi così cospicuo da rendere antieconomico l'impiego della macchina»⁵⁷.

Di fronte a questi argomenti, che poi si riducono ad una sola pesante obiezione, quale senso può avere la proposta di un modello di codifica che è tutto tranne che leggero, economico e poco gravoso? Provo a rispondere e noto che, nonostante la bibliografia sull'automazione nella critica del testo risalga al 1947⁵⁸, a tutt'oggi non possiamo registrare nessuna proposta davvero soddisfacente, nessun sostanziale progresso. Il motivo? Non potrebbe essere che si

⁵³ D'Arco S. Avalle 1978², pp. 24-5.

⁵⁴ West 1991, p. 73.

⁵⁵ Balduino 1989³, p. 136.

⁵⁶ Pratesi 1977, pp. 187-90; ora anche in Id. 1992, pp. 45-8 (p. 48).

⁵⁷ Chiesa 2002, p. 50. La cit., ovviamente, è riferita all'uso del computer nella fase di collazione dei testimoni, ma, per avere un'idea più completa dell'opinione dell'autore sui rapporti tra informatica e critica testuale, non andranno trascurate le pp. 178-9 della *Conclusione*, da cui emerge un approccio molto riflessivo e moderatamente critico, ma senza chiusure preconette.

⁵⁸ Mi riferisco al lavoro, davvero pionieristico, di Hinman 1947, pp. 99-106. L'ultima proposta che conosco, in ordine di tempo, è l'algoritmo di collazione presentato da Howe - Spencer 2004, pp. 253-70, con ampia bibliografia sugli studi precedenti alle pp. 267-70.

è cercato di collazionare tutti i manoscritti (*omnes textus*), ma non i manoscritti tutti interi (*textus toti*)? Che non si sono forniti ai vari algoritmi di collazione sviluppati di volta in volta tutti i dati di cui questi avevano bisogno per arrivare a risultati significativi? Che, forse, non basta semplicemente copiare i testimoni, ma che c'è bisogno di descriverli ed interpretarli con un'adeguata codifica?

Potrebbe essere, ma a pensarci bene la risposta più giusta l'ha data ancora una volta padre Roberto Busa: «Tenete conto che l'informatica non è per fare le stesse ricerche di prima, con gli stessi metodi di prima, ma solo più velocemente e magari con meno lavoro umano»⁵⁹.

Bibliografia

- Adamo 1987 = G. Adamo, *La codifica come rappresentazione. Trasmissione e trattamento dell'informazione nell'elaborazione automatica di dati in ambito umanistico*, in *Studi di codifica e trattamento automatico di testi*, a cura di G. Gigliozzi, Roma 1987, pp. 39-63.
- Balduino 1989³ = A. Balduino, *Manuale di filologia italiana*, Firenze 1989³.
- Bartoli Langeli 1993 = A. Bartoli Langeli, *L'edizione dei testi documentari. Riflessioni sulla filologia diplomatica*, in *L'edizione di testi mediolatini. Problemi metodi prospettive*. Testi della VIII Settimana Residenziale di studi medievali (Carini, 24-28 ottobre 1988), Palermo 1993, pp. 116-31.
- Bieler 1955 = *Regula Magistri* (rec.), L. Bieler, in «*Speculum*», XXX (1955), pp. 690-2.
- Bloomfield 1979 = M.W. Bloomfield - B.G. Guyot - D.R. Howard - T.B. Kabbealo, *Incipits of Latin Works on the Virtues and Vices, 1100-1500 A.D.*, Cambridge MA 1979.
- Blume 1863 = *Lex Romana Burgundionum*, F. Blume (ed.), MGH *Leges* III, Hannoverae 1863 (rist. anast.: Hannover 1993).
- Bryan 1988 = M. Bryan, *SGML. An Author's Guide to the Standard Generalized Markup Language*, Wokingham-Reading-New York 1988.
- Busa 1997 = R. Busa, *Introduzione ai lavori in Umanesimo & Informatica. Le nuove frontiere della ricerca e della didattica nel campo degli studi letterari*. Atti del Convegno (Trento, 24-25 maggio 1996), a cura di D. Gruber - P. Pauletto, Pesaro 1997, pp. 9-13.
- Catalogo di manoscritti filosofici* 2000 = *Catalogo di manoscritti filosofici nelle biblioteche italiane*, X: Arezzo, Borgomanero, Novara, Palermo, Pavia, Sansepolcro, Siena, Stresa, Unione Accademica Nazionale, Firenze 2000.

⁵⁹ Busa 1997, pp. 9-13 (p. 9).

- Cenci 1988 = C. Cenci, *Lettera De bono animae di fr. Gosmaro da Verona al B. Rainaldo, Arcivescovo di Ravenna*, in «Archivum Franciscanum Historicum», LXXXI (1988), pp. 50-71.
- Chevalier 1892-1921 = U. Chevalier, *Repertorium hymnologicum*, Louvain 1892-1921.
- Chiesa 2002 = P. Chiesa, *Elementi di critica testuale*, Bologna 2002.
- Ciotti 1995 = F. Ciotti, *Testi elettronici e banche dati testuali: problemi teorici e tecnologie*, in «Schede Umanistiche», n.s. IX, 2 (1995), pp. 147-78.
- Ciotti 1997 = F. Ciotti, *Cosa è la codifica informatica dei testi?*, in *Umanesimo & Informatica. Le nuove frontiere della ricerca e della didattica nel campo degli studi letterari*. Atti del Convegno (Trento, 24-25 maggio 1996), a cura di D. Gruber - P. Pauletto, Pesaro 1997, pp. 55-85.
- D'Arco S. Avalle 1972 = D'Arco S. Avalle, *Norme pratiche per la trascrizione dei manoscritti letterari centro-meridionali dei primi secoli (sezione poetica)*, Torino 1972.
- D'Arco S. Avalle 1978² = D'Arco S. Avalle, *Principi di critica testuale*, Padova 1978².
- Delaissé 1956 = L.M.J. Delaissé (ed.), *Le manuscrit autographe de Thomas a Kempis et l'Imitation de Jésus Christ: examen archéologique et édition diplomatique du Bruxellensis 5855-61*, Bruxelles 1956.
- Fadini - Savy 1999 = B. Fadini - C. Savy, *Informatica per le scienze umane*, Milano 1999.
- Ferrarini 2003 = E. Ferrarini, *L'informatica umanistica oggi*, in «Orpheus», n.s. XXIV (2003), pp. 227-56.
- Ferrarini 2005 = E. Ferrarini, *La Littera de bono animae di Gosmaro da Verona. Appunti per un'edizione mediante l'ausilio delle tecnologie digitali*, in «Vichiana», n.s. VII (2005), pp. 310-28.
- Goldfarb 1992² = C.F. Goldfarb, *The SGML Handbook*, Oxford 1992².
- Goldfarb - Prescod 1999 = C.F. Goldfarb - P. Prescod, *The XML Handbook*, Upper Saddle River NJ 2003⁵ [trad. it.: *XML*, Milano 1999].
- Guidelines 2002 = *Guidelines for Electronic Text Encoding and Interchange (TEI P4). XML-compatible edition*, a cura di C.M. Sperberg-McQueen - L. Burnard, Oxford-Providence-Charlottesville-Bergen 2002.
- Guyotjeannin 1995² = O. Guyotjeannin - J. Pycke - B.-M. Tock, *Diplomatique médiévale*, Turnhout 1995².
- Hansen 2002 = A.M. Hansen, *Text encoding of manuscripts: Danish prayer books from the 16th century*, in «Le Médiéviste et l'Ordinateur», XLI (2002), pp. 72-82.
- Herwijnen 1994² = E. van Herwijnen, *Practical SGML*, Boston-Dordrecht-London 1994².
- Hinman 1947 = C. Hinman, *Mechanized collation. A preliminary report*, in «Papers of the Bibliographical Society of America», XLI (1947), pp. 99-106.
- Howe - Spencer 2004 = C.J. Howe - M. Spencer, *Article: Collating Texts Using Progressive Multiple Alignment*, in «Computers and the Humanities», XXXVIII (2004), pp. 253-70.
- Leonardi 1998 = *Testi, manoscritti, ipertesti. Compatibilità informatica e*

- letteratura medievale*. Atti del Convegno Internazionale (Firenze, Certosa del Galluzzo, 31 maggio-1 giugno 1996), a cura di L. Leonardi, Firenze 1998.
- Magheri Cataluccio - Fossa 1979 = M.E. Magheri Cataluccio - A.U. Fossa, *Biblioteca e cultura a Camaldoli*, Roma 1979.
- Masai 1950 = F. Masai, *Principes et conventions de l'édition diplomatique*, in «Scriptorium», IV (1950), pp. 177-93.
- Mazzatinti 1896 = G. Mazzatinti, *Inventari dei manoscritti delle biblioteche d'Italia*, vol. 6, Forlì 1896.
- Mazzoni 1998 = F. Mazzoni, *Per un censimento dei codici della Commedia*, in *Testi, manoscritti, ipertesti. Compatibilità informatica e letteratura medievale*. Atti del Convegno Internazionale (Firenze, Certosa del Galluzzo, 31 maggio-1 giugno 1996), a cura di L. Leonardi, Firenze 1998, pp. 145-72.
- Mordenti 1987 = R. Mordenti, *Appunti per una semiotica della trascrizione nella procedura ecdotica computazionale*, in *Studi di codifica e trattamento automatico di testi*, a cura di G. Gigliozzi, Roma 1987, pp. 85-124.
- Mordenti 2001 = R. Mordenti, *Informatica e critica dei testi*, Roma 2001.
- Orlandi 1987 = T. Orlandi, *Informatica umanistica. Riflessioni storiche e metodologiche, con due esempi*, in *Studi di codifica e trattamento automatico di testi*, a cura di G. Gigliozzi, Roma 1987, pp. 1-38.
- Orlandi 1997 = T. Orlandi, *Il testo critico e il supporto magnetico*, in *Discipline umanistiche e informatica. Il problema della formalizzazione*. Ciclo di Seminari, febbraio-giugno 1994, a cura di T. Orlandi, Roma 1997, pp. 159-72.
- Orlandi 1998 = T. Orlandi, *Teoria e prassi della codifica dei manoscritti*, in *Gli Zibaldoni di Boccaccio. Memoria, scrittura, riscrittura*. Atti del Seminario internazionale di Firenze-Certaldo, 26-28 aprile 1996, a cura di C. Cazalé Bérard - M. Picone, Firenze 1998, pp. 349-60.
- Orlandi 2002 = T. Orlandi, *A che punto siamo con l'informatica umanistica?*, in *Giuseppe Gigliozzi: la fondazione dell'informatica applicata ai testi letterari*, a cura di R. Mordenti [«Testo & Senso», IV/V (2001-2002)], Roma 2002, pp. 105-9.
- Perilli 1995 = L. Perilli, *Filologia computazionale*, Roma 1995.
- Pratesi 1957 = A. Pratesi, *Una questione di metodo: l'edizione delle fonti documentarie*, in «Rassegna degli Archivi di Stato», XVII (1957), pp. 312-33.
- Pratesi 1977 = A. Pratesi, *Limiti e difficoltà dell'uso dell'informatica per lo studio della forma diplomatica e giuridica dei documenti medievali*, in *Informatique et histoire médiévale*. Communications et débats de la Table Ronde CNRS, organisée par l'École française de Rome et l'Institut d'Histoire Médiévale de l'Université de Pise (Rome, 20-22 mai 1975), a cura di L. Fossier - A. Vauchez - C. Violante, Roma 1977, pp. 187-90.
- Pratesi 1992 = A. Pratesi, *Tra carte e notai. Saggi di diplomatica dal 1951 al 1991*, Roma 1992.
- Regula Magistri = Regula Magistri. La Règle du Maître. Édition diplomati-*

- que des manuscrits latins 12205 et 12634 de Paris*, H. Vanderhoven - F. Masai - P.B. Corbett (eds.), Bruxelles 1953.
- Robinson 1994 = P. Robinson, *The Transcription of Primary Textual Sources Using SGML*, Oxford 1994.
- Rudolf 1957 = R. Rudolf, *Ars moriendi. Von der Kunst des heilsamen Lebens und Sterbens*, Köln-Graz 1957.
- Scalia 1998-1999 = Salimbene de Adam, *Cronica*, G. Scalia (ed.), Turnhout (*Corpus Christianorum. Continuatio Mediaevalis* CXXV-CXXV A), 1998-1999.
- Segre 1979 = C. Segre, *Les transcriptions en tant que diasystèmes*, in *La pratique des ordinateurs dans la critique des textes*. Actes du Colloque international (Paris, 29-31 mars 1978), a cura di J. Irigoin - G.P. Zarri, CNRS, Paris, 1979, pp. 45-9.
- Text Encoding Initiative* 1995 = *Text Encoding Initiative. Background and Context*, a cura di N. Ide - J. Véronis [«Computers and the Humanities», XXIX, 1-3 (1995)], Boston-Dordrecht-London 1995.
- Tognetti 1982 = G. Tognetti, *Criteri per la trascrizione di testi medievali latini e italiani* («Quaderni della Rassegna degli Archivi di Stato», LI), Roma 1982.
- Walther 1969² = H. Walther, *Initia carminum ac versuum Medii Aevi posterioris latinorum*, Göttingen 1969².
- Wattenbach 1896³ = W. Wattenbach, *Das Schriftwesen im Mittelalter*, Leipzig 1896³ (rist. anast. Graz 1958).
- West 1991 = M.L. West, *Critica del testo e tecnica dell'edizione*, trad. it. a cura di G. Di Maria, Palermo 1991 [ediz. orig.: *Textual Criticism and Editorial Technique (applicable to Greek and Latin texts)*, Stuttgart 1973].