

LA MACCHINA NEL TEMPO

Studi di informatica umanistica
in onore di Tito Orlandi

a cura di

Lorenzo Perilli
Domenico Fiormonte

Le Lettere

In copertina: Tomografia assiale di un frammento del cosiddetto Meccanismo di Anticitera, il più antico calcolatore meccanico della storia, risalente alla Grecia del II secolo a.C. (Da: *The Antikythera Mechanism Research Project*)

Copyright © 2011 by Casa Editrice Le Lettere – Firenze

ISBN 978 88 6087 524 5

www.lelettere.it

Dino Buzzetti

OLTRE IL RAPPRESENTARE:
LE POTENZIALITÀ DEL *MARKUP*

1. Le considerazioni che seguono nascono fondamentalmente da un interrogativo, che sorge dall'indiscusso successo ottenuto dalle norme per la codifica dei testi letterari proposte dalla Text Encoding Initiative (TEI), ora pressoché universalmente adottate, e dall'ingente produzione di dati resisi oramai disponibili in tale formato. A tutt'ora, l'enfasi sembra essere stata posta prevalentemente sulla produzione, piuttosto che sull'impiego dei dati così prodotti, come se si trattasse soltanto di formare un immenso archivio di testi codificati, predisposti per così dire a 'futura memoria,' senza considerare più di tanto il problema dell'uso e delle modalità di impiego di tali risorse. Pare così che una riflessione si imponga sulla sfida ulteriore che si presenta alla comunità degli studiosi di informatica umanistica. L'informatica, infatti, non si limita soltanto ad affrontare il problema della rappresentazione in forma adeguata dell'informazione, ma considera soprattutto il problema delle procedure di elaborazione dei dati codificati in qualsivoglia formato. Ora, se si osservano le cose da un punto di vista così generale, si nota che nell'informatica applicata alle discipline umanistiche, per quanto riguarda il settore degli studi testuali, ci si propone prevalentemente di produrre edizioni digitali di qualità paragonabile a quella delle edizioni convenzionali e che in misura molto minore si insiste sulle maggiori potenzialità che esse offrono ai fini della ricerca. Si rischia così di ignorare la specificità propria della risorsa digitale, che è precisamente quella di poter essere elaborata con procedure computazionali. Se si resta sul piano della rappresentazione, i vantaggi, e il loro valore aggiunto, non sembrano così decisivi da giustificare, se non in casi di particolare interesse, la predilezione della forma digitale di edizione rispetto a quella convenzionale. A quanto pare, la ragione di questo stato di cose risiede proprio nella scarsa attenzione fin qui prestata a ciò che la forma digitale di rappresentazione in quanto tale permette di ottenere sul piano dell'elaborazione del contenuto testuale rappresentato. A que-

sto proposito, ciò su cui si intende qui riflettere è la specifica funzione che il *markup* può svolgere nel passare dalla rappresentazione all'elaborazione del contenuto testuale che il testo digitale riesce a rappresentare.

Il problema è già stato posto e autorevolmente. In un articolo presentato nel gennaio 2003 ad un convegno sul tema *Transforming Disciplines: Computer Science and the Humanities*, John Unsworth scriveva:

Noi dobbiamo (dobbiamo ancora) dimostrare l'utilità di tutti i materiali che abbiamo digitalizzato nel corso degli ultimi dieci anni ed oltre – e dimostrarne l'utilità non solo per quanto riguarda la maggiore facilità di accesso, ma in modo più preciso, per quanto riguarda *ciò che ne possiamo fare* quando li abbiamo a nostra disposizione [...]¹

Si tratta di un invito evidente ad andare 'oltre il rappresentare,' come recita esplicitamente il titolo di uno dei suoi articoli pubblicati successivamente.² Ora, ciò su cui si vuole qui richiamare l'attenzione è il ruolo che il *markup* può svolgere nel rendere operativamente praticabile il passaggio dalla pura rappresentazione del testo all'analisi del suo contenuto informativo e nell'associare ad una determinata variazione della lettera del testo una variazione corrispondente della sua interpretazione e del contenuto ad esso associato. È in grado il *markup* di svolgere questa funzione? Ciò che qui si intende mostrare è che la capacità di operare questa connessione appartiene al *markup* in virtù della sua stessa natura e che solo valutando adeguatamente le caratteristiche intrinseche del *markup* è possibile apprezzarne appieno tutte le potenzialità operative e funzionali.

2. Il fatto che il *markup* non abbia solo una funzione 'dichiarativa,' ma svolga anche una funzione 'performativa' è stato chiaramente riconosciuto da tempo. In un ben noto intervento, Allen Renear

¹ J. Unsworth, *Tool-Time, or 'Haven't we been here already? Ten Years in Humanities Computing*, presentato al convegno "Transforming Disciplines: The Humanities and Computer Science", Washington, DC, 17-18 gennaio 2003, URL <<http://www.carnegie.rice.edu/ppt-unsouth.cfm>> [31.03.2010].

² Id., *Forms of attention: Digital humanities beyond representation*, presentato al terzo convegno del Canadian Symposium on Text Analysis (CaSTA), "The Face of Text: Computer-Assisted Text Analysis in the Humanities", McMaster University, 19-21 novembre 2004, URL <<http://www3.isrl.uiuc.edu/~unsouth/FOA>> [31.03.2010].

ricorda che già Charles Goldfarb, in un suo «articolo fondamentale», aveva dettagliatamente descritto la funzione ‘procedurale’ del *markup*, che consisterebbe nel «richiamare / specificare / prescrivere» una «procedura / effetto / processo / azione di formattazione o di resa grafica» del testo, ed egli stesso riconosce che «il *markup* è dotato di modalità» e che può assumere «almeno in parte» una modalità «performativa», ossia quella propria degli «atti linguistici» teorizzati dal filosofo del linguaggio John Austin.³ Il *markup* non ha dunque solamente una funzione descrittiva, ma è dotato anche di una funzione operativa.

Il passaggio dall’originaria funzione operativa ad una funzione meramente dichiarativa del *markup* è stato effettuato soprattutto per svincolarne l’operatività dalle specifiche piattaforme *hardware* e dallo specifico ambiente *software* a cui era vincolata. Ma i vantaggi ottenuti col passaggio alla modalità dichiarativa del *markup* non dovrebbero farne dimenticare la primaria funzione operativa. Ciò su cui si intende qui principalmente insistere è appunto la modalità performativa del *markup*: valutarne appieno l’insieme delle potenzialità per usi specifici di ricerca rappresenta una delle maggiori sfide di fronte alle quali si trova attualmente la teoria della codifica del testo.

Introdotta originariamente nell’ambito dei sistemi di produzione dei documenti, il *markup* era di natura essenzialmente procedurale e serviva a scopi eminentemente pratici, quali la formattazione, l’impaginazione e la stampa del testo. Il passaggio alla modalità dichiarativa ha reso possibile l’utilizzazione del *markup* per la rappresentazione digitale del testo. Ma anche il recupero della sua modalità operativa può essere reso funzionale all’analisi e alla rappresentazione strutturale del contenuto informativo del testo, andando oltre il suo impiego generico per la visualizzazione e la resa grafica della testualità digitale.

Il punto cruciale consiste nel tener conto in modo adeguato del nesso che lega l’espressione al contenuto del testo,⁴ perché è a partire da questo nesso, dinamico ed operativo, che è possibile considerare, oltre agli aspetti sintattici e semantici, anche l’aspetto pragmatico della testualità e connettere organicamente tutte le dimensioni se-

³ A. Renear, *The descriptive/procedural distinction is flawed*, «Markup Languages: Theory & Practice», 2:4 (2000), 411-420, pp. 412 e 419; cf. C. F. Goldfarb, *A Generalized approach to document markup*, «ACM SIGPLAN Notices», 16:6 (1981), 68-73.

⁴ Cfr. L. Hjelmslev, *I fondamenti della teoria del linguaggio* (1943, 1961²), trad. it. Einaudi, Torino, 1968, pp. 52-65.

miotiche del testo in un contesto operativo e computazionale. Ma procediamo con ordine.

3. Riprendiamo il discorso dall'adozione delle tecniche di *markup* sviluppate nell'ambito dei sistemi di produzione dei documenti per la rappresentazione digitale del testo. Il tipo di rappresentazione dell'informazione testuale introdotta in informatica, la stringa o sequenza di caratteri codificati, risulta con tutta evidenza insufficiente a rappresentare esaurientemente tutta l'informazione contenuta in quella che è la forma corrente del testo, vale a dire la pagina stampata o manoscritta. L'introduzione del *markup* era dunque motivata, primariamente, da un'esigenza di completezza e di esaustività della rappresentazione, per ridurre al minimo la perdita di informazione nel passaggio dalla forma convenzionale alla forma digitale del testo.

Ma il testo non è solo la sua «immagine» o rappresentazione,⁵ il testo è anche, e soprattutto, l'insieme delle sue interazioni con chi lo produce e chi lo legge. E la forma di rappresentazione del testo è funzionale agli usi che se ne fanno. Anche la forma digitale del testo non può dunque rinunciare al fondamentale requisito di funzionalità che deve esserle proprio. Dal punto di vista funzionale, essa deve permettere forme di elaborazione computazionale tali da permettere l'espletamento delle operazioni richieste dalle pratiche correnti della critica testuale e della critica letteraria. È possibile usare il *markup* per rappresentare sia le varianti testuali che le varianti interpretative? Si possono trovare forme di collegamento tra i due tipi di rappresentazione e introdurre procedure computazionali per determinare le modalità della loro interconnessione? Ma prima di affrontare direttamente la questione è utile insistere ancora su alcune caratteristiche generali del *markup*.

È legittimo chiedersi in che modo il *markup* sia funzionale all'uso critico del testo. La sua potenzialità funzionale risiede nel suo carattere performativo e operativo. Ciò che esprime al meglio questa caratteristica che gli è propria è la punteggiatura, che è a tutti gli effetti una forma di *markup*:

la punteggiatura non è qualcosa che appartiene solamente al nostro sistema di scrittura; essa è una forma di *markup* del documento che può essere modificata e sostituita da altri tipi di *markup*.⁶

⁵ C. Segre, *Avviamento all'analisi del testo letterario*, Torino, Einaudi, 1985, p. 378.

⁶ J.H. Coombs, A.H. Renear, and S.J. DeRose, *Markup systems and the future of scholarly text processing*, «Communications of the ACM», 30:11 (1987), 933-947, p. 935.

Una variazione nella punteggiatura può essere considerata una variante testuale. Per esempio, le due proposizioni ‘Quando l’atleta saltò, l’ostacolo cadde’ e ‘Quando l’atleta saltò l’ostacolo, cadde’ possono essere considerate come varianti diverse dello stesso testo. Ma una variazione nella punteggiatura non costituisce soltanto una variante testuale – ovvero dell’*espressione* del testo – perché, come qui si vede, il semplice spostamento di una virgola può modificare l’intero significato della proposizione e la virgola può essere considerata anche come un’istruzione per interpretare il testo. Un certo segno di interpunzione – e sempre lo stesso – può essere visto come il risultato di un atto di composizione del testo, così come può essere considerato, altrettanto legittimamente, un’istruzione per un atto, o un’operazione, di interpretazione del testo. Una variazione nella punteggiatura è una modifica del testo e una modifica della sua interpretazione. E una variazione nel *markup* è una variazione della stessa natura. Lo stesso segno di interpunzione produce una variante testuale e una variante interpretativa e il *markup*, allo stesso modo, può essere inteso come uno strumento per trasformare una variante interpretativa in una variante testuale e viceversa. Il *markup* può quindi svolgere, operativamente, la funzione di un dispositivo che collega la rappresentazione dell’espressione e la rappresentazione del contenuto del testo. Il problema è dunque quello di fornire una rappresentazione computabile del contenuto del testo, così come lo è quella della sua espressione e di considerare le possibili forme di interazione tra i due tipi di rappresentazione. Un modo di descrivere questa situazione è quello di dire che il *markup* svolge una funzione diacritica e che, come ogni altro segno diacritico, ha natura riflessiva.

4. Ma prima di trarre da questo fatto le dovute conseguenze e ad evitare che l’esempio addotto possa sembrare artificioso e forzato, vale riportare il caso, emblematico, dell’edizione di una lettera di Spinoza a Tschirnhaus (1674), discusso da Wim Klever, uno dei più accreditati studiosi dell’opera del filosofo olandese.⁷ La questione, di rilevanza non solo filologica, ma eminentemente critica e interpretativa, verte precisamente sull’inserimento e la posizione di due virgole. Il passo esaminato è il seguente:

⁷ Cfr. W. Klever, *Conditioned inertia in the physics of Spinoza and his followers*, URL <<http://www.fogliospinoziano.it/Conditioned%20inertia.pdf>>, pp. 4-6 [02.08.2010].

Ex. gr. Lapis a causa externa ipsum impellente certam motus quantitatem accipit, qua postea cessante [,] causae externae impulsu moveri necessario perget,

Come Klever osserva, nell'edizione degli *Opera quae supersunt* di Spinoza pubblicata nel 1843 da C. H. Bruder, «de due virgole dopo le parole ‘postea’ e ‘impulsu’», che compaiono ancora nell'«edizione critica» del Gebhardt (1924) – in ciò seguita dalla traduzione inglese di Samuel Shirley e da «una dozzina di altre apprezzate e qualificate traduzioni» – sono invece assenti e, secondo Klever, a buon diritto. Infatti, nella sua introduzione, il Bruder dichiarava di seguire il testo dell'*editio princeps*, da cui cercava tuttavia di espungere le imperfezioni (*sublatis vitiis*) e gli errori (*errores detecti et remoti sunt*). Dunque, conclude Klever, «il testo dell'*Opera posthuma* [1677] è corrotto» e le correzioni del Bruder, che ometteva «i segni di interpunzione», sono state successivamente ignorate. La presenza delle «virgole» nelle edizioni successive tradisce «l'interpretazione abituale e preconçetta del fenomeno» dell'inerzia, che non può essere considerata la genuina «concezione scientifica di Spinoza». A giudizio del Klever, «è probabile che il tipografo le abbia introdotte nel testo per la stampa, pensando che ne chiarissero il senso». Klever propone quindi di emendare il testo della lettera, ricollocando le virgole nel modo seguente:

Ex. gr. Lapis a causa externa ipsum impellente certam motus quantitatem accipit, qua postea cessante [,] causae externae impulsu moveri necessario perget,

ossia «trasponendo (mentalmente) la prima delle due virgole nella posizione tra ‘cessante’ e ‘causae’ e cancellando la seconda». Così, mentre la traduzione di Shirley recita

A stone receives from the impulsion of an external cause a fixed quantity of motion whereby it will necessarily continue to move when the impulsion of the external cause has ceased,

la traduzione del testo emendato, proposta da Klever, ne muta radicalmente, come si vede, il significato:

A stone for instance receives from an external cause impelling it a certain degree of motion. After the cessation of this impulsion the stone will necessarily continue its course by the impact of an external cause,

e la rende, a suo parere, più consona al chiaro «rifiuto» spinoziano «di ogni forma di inerzia» incondizionata e più coerente col contesto della lettera a Tschirnhaus e delle argomentazioni che vi svolge, dalle quali sarebbe in realtà addirittura «presupposta». Sulla presenza o l'assenza di una virgola, si gioca così la comprensione o il fraintendimento del pensiero di Spinoza a proposito di un problema fisico e filosofico di capitale importanza. La piena comprensione del ruolo dei segni diacritici – e del *markup*, che ne costituisce, a ben vedere, un caso particolare – si rivela quindi essenziale ed imprescindibile per una pratica corretta della critica testuale in ambiente digitale e ne mostra la diretta connessione con l'analisi critica e interpretativa del contenuto del testo.

5. Ora, un segno diacritico fa parte, da un canto, del nostro sistema di scrittura e può essere considerato come un'espressione del nostro linguaggio oggetto, ma poiché, d'altro canto, ha natura riflessiva, descrive una certa caratteristica del testo e quindi assume un valore metalinguistico. Questo tipo di 'ambiguità diacritica' è precisamente ciò che assegna al *markup* la sua funzione o capacità operativa. Una stessa espressione del linguaggio oggetto può essere intesa in due modi diversi: (a) direttamente come una variante testuale, oppure (b) come un'asserzione autoriflessiva del secondo ordine, equivalente ad un'asserzione metalinguistica sul contenuto del testo. Ed è proprio quando viene inteso in quest'ultimo senso che il *markup* muta le sue prerogative linguistiche: esso cessa di svolgere una mera funzione espressiva ed assume una funzione sostanzialmente performativa, ossia perde il suo carattere primariamente rappresentativo e acquisisce una natura eminentemente operativa. Considerato in questo modo, il *markup* può essere inteso come una regola di inferenza. È stato sostenuto, e in modo affatto convincente, che

per descrivere il significato del *markup* di un documento è sufficiente generare l'insieme delle inferenze riguardanti quel documento che sono autorizzate dal *markup*,

sicché

il lettore può usare legittimamente il *markup* per trarre inferenze sulla struttura e sulle proprietà del testo.⁸

Di conseguenza, possiamo considerare il *markup* come una regola per trarre inferenze a proposito del contenuto assegnato ad una determinata espressione del testo. Intendendo il *markup* in questo modo, una variante testuale – perché tale è da intendersi il *markup* considerato come espressione del linguaggio oggetto – diventa una regola per generare una variante interpretativa e, da questo punto di vista, sembra essere del tutto legittimo comprendere una descrizione formale e computabile di un valore interpretativo nella rappresentazione, ovvero nell'edizione del testo – in effetti, a ben vedere, ciò equivale, né più né meno, ad una annotazione aggiunta ad una pagina del Web semantico. Un'edizione assume un carattere funzionale quando viene dotata di un apparato e con la descrizione del contenuto informativo del testo non facciamo altro, in realtà, che arricchire l'apparato di una sua edizione. La rappresentazione del testo dotata di *markup* può essere usata come un dispositivo computazionale per produrre una modificazione nella descrizione formale del suo contenuto e una rappresentazione strutturata del contenuto testuale può essere usata come un dispositivo computazionale di natura interpretativa per produrre una modificazione o 'deformazione' della sua espressione e della sua rappresentazione codificata col *markup*. Non è fuori luogo, a questo proposito, ricordare che una «deformazione del testo»⁹ è stata autorevolmente descritta da Jerome McGann come «un'operazione critica» (36) e «interpretativa» (46) e che, in effetti, «possiamo utilmente considerare ogni forma di critica e di interpretazione come una deformazione» (46) e concepire una rappresentazione dinamica e operativa del testo come una specie di dispositivo deformante.

6. Ma è possibile cercare di procedere in modo meno astratto e concettuale e proseguire con considerazioni più concrete, citando ricerche effettive ed esempi tangibili? In primo luogo, possiamo riferirci ad un articolo sullo Henry the Third Fine Rolls Project pre-

⁸ C.M. Sperberg McQueen, C. Huitfeldt, A. Renear, *Meaning and interpretation of markup*, «Markup Languages: Theory & Practice», 2:3 (2000), 215-234, pp. 231 e 216.

⁹ L. Samuels and J. McGann, *Deformation and interpretation*, «New Literary History», 30:1 (1999), 25-56, p. 38.

sentato al congresso Digital Humanities 2007 di Urbana / Champaign.¹⁰ In altre pubblicazioni riguardanti lo stesso progetto, gli autori insistono sull'opportunità di «collegare testi TEI e ontologie»¹¹ e presentano un tentativo riuscito di «popolare» un'ontologia¹² «con dati estratti dal *markup*» (320). Con il suo «reticolo di associazioni» il modello concettuale fornito dall'ontologia risulta «complementare al *markup*», ma suggerisce anche «soluzioni alternative e correzioni» alla codifica in atto» (317). I vantaggi che si ottengono modellando «il mondo del testo»,¹³ affermano gli autori, sono evidenti; infatti

nel costruire l'ontologia, il ricercatore ha una visione dei materiali studiati diversa da quella che ha nel codificare il testo lineare in XML. Il ricercatore può così proporre nuove interpretazioni (come identificare due personaggi o distinguerli l'uno dall'altro) che possono essere agevolmente rappresentate mediante l'ontologia.

Non pare tuttavia ancora risolto il problema inverso, quello di arricchire il *markup* a partire dall'ontologia. Infatti, a fronte di tutto questo, i ricercatori impegnati nel progetto osservano che le nuove interpretazioni ricavate dall'ontologia «debbono essere importate coerentemente nei documenti XML di origine»,¹⁴ ma che «c'è ancora bisogno di perfezionare [tale] procedimento inverso, in modo che il *markup* possa trarre vantaggio dal processo di costruzione dell'ontologia». ¹⁵ A loro modo di vedere, la difficoltà risiede nel fatto che

la costruzione di un'ontologia basata su documenti codificati in XML solleva un problema di sincronizzazione tra i documenti d'origine e

¹⁰ A. Ciula, P. Spence, J. M. Vieira, G. Poupeau, *Expressing complex associations in medieval historical documents: The Henry III Fine Rolls project*, in «Digital Humanities 2007: Conference Abstracts», Urbana-Champaign, Graduate School of Library and Information Science, University of Illinois, 2007, 34b-37b.

¹¹ A. Ciula and J. M. Vieira, *Complementing and extending TEI documents with an ontology: Henry III Fine Rolls project*, TEI Members Meeting 2008, Abstract, URL <<http://www.cch.kcl.ac.uk/cocoon/tei2008/programme/abstracts/abstract-161.html>>.

¹² A. Ciula, P. Spence, J. M. Vieira, *Expressing complex associations in medieval historical documents: The Henry III Fine Rolls project*, «Literary and Linguistic Computing», 23:3 (2008), 311-25, p. 314.

¹³ Ciula and Vieira, *Complementing and extending TEI documents with an ontology*, URL <<http://www.cch.kcl.ac.uk/cocoon/tei2008/programme/abstracts/abstract-161.html>>.

¹⁴ Vieira and Ciula, *Implementing an RDF/OWL Ontology on Henry the III Fine Rolls*, [paper presented at OWLED 2007, Innsbruck, June 6-7], URL <http://www.webont.org/owled/2007/PapersPDF/submission_6.pdf>, p. 9.

¹⁵ Ciula et al., *Expressing complex associations in medieval historical documents*, p. 320.

Pontologia stessa. Mentre l'integrazione dell'informazione nell'ontologia a partire dai documenti XML è immediata e può essere realizzata in modo automatico, nel procedimento inverso si possono incontrare dei problemi.¹⁶

A quanto pare, ciò che preoccupa i responsabili del progetto è la coerenza tra le fonti documentarie e le interpretazioni ricavate dall'ontologia. E ciò è assolutamente doveroso quando si tratta di fonti storiche. È tuttavia importante osservare che, in generale, un testo non deve necessariamente essere coerente. Una data interpretazione assegnata ad un'espressione del testo può entrare in conflitto con un'interpretazione diversa suggerita da altri passi dello stesso testo.

Gli esempi non mancano. Negli *Eroici furori* di Giordano Bruno, è stato possibile mostrare che le relazioni concettuali che si stabiliscono tra determinate nozioni mutano ed evolvono diacronicamente nel corso dei diversi dialoghi che compongono l'opera.¹⁷ Così, per esempio, in tre diversi dialoghi, I.1, I.4 e II.5, rispettivamente all'inizio, al centro e alla fine dell'opera, la stessa «unità concettuale (intelletto)» presenta «tre varianti interpretative» e «la relazione tra amore e intelletto» si articola diversamente nel modo seguente: «l'amore illumina l'intelletto (I.1); l'amore è definito metaforicamente come la facoltà della vista (I.4), conducendo così alla visione del sommo bene (II.5), che è il vero oggetto dell'amore (I.4)».¹⁸ Jerome McGann, da parte sua, analizzando il sonetto *As Kingfishers Catch Fire* del poeta Gerard Manley Hopkins (1844-1889), mostra che «la ripetizione della parola [as] nel secondo verso», giocando sulle diverse accezioni del termine, «introduce una grammatica completamente nuova» e «ne enfatizza le possibilità di variazione» interpretativa. Insistendo poi sui numerosi fenomeni di questo tipo presenti nel componimento, osserva che «i poeti si comportano continuamente in questo

¹⁶ J. M. Vieira and A. Ciula, *Implementing an RDF/OWL ontology on Henry the III Fine Rolls*, p. 9.

¹⁷ S. Bassi, F. Dell'Orletta, D. Esposito, A. Lenci, *Computational linguistics meets philosophy: A latent semantic analysis of Giordano Bruno's texts*, in «Rinascimento», 46 (2006), pp. 631-47.

¹⁸ A. Isolani, C. Lorito, Ch. Genovesi, D. Marotta, M. Matteoli, C. Tozzini, *Topic maps and MVD for the representations of interpretative variants*, in «Digital Humanities 2009: Conference Abstracts», Proceedings of the 21th ALLC and ACH Joint International Conference (University of Maryland, College Park, June 22-25, 2009), College Park, The Maryland Institute for Technology in the Humanities (MITH), 2009, 8b 11a, p. 9b.

modo» e che in ciò consiste «l'essenza stessa della testualità poetica». ¹⁹

Sicché una rappresentazione adeguata del contenuto e delle possibili interpretazioni del testo richiede una soluzione per la sovrapposizione di modelli interpretativi contrastanti, proprio come una rappresentazione adeguata della lettera, ovvero dell'espressione del testo, richiede una soluzione adeguata per la sovrapposizione di varianti testuali diverse. L'editore e in ugual misura il lettore del testo debbono essere lasciati liberi di scegliere tra le diverse varianti, testuali e interpretative, mentre procedono sequenzialmente, passo dopo passo, nell'esame dello sviluppo lineare del testo. Perché, dunque, le «difficoltà (*implications*) tecniche» del tentativo di «mantenere sincronizzate le due risorse», il testo codificato e l'ontologia, vengono considerate «spaventose» dai ricercatori del Fine Rolls Project? ²⁰ Ci troviamo di fronte, in questo caso, ad un problema di natura storica o ad un problema eminentemente tecnico?

Le difficoltà che essi mettono in luce sembrano più di natura storica che di natura esclusivamente tecnologica. Non si vede infatti perché, in generale, si debbano far dipendere le soluzioni computazionali riguardanti l'interrelazione tra le varianti testuali e le varianti interpretative del testo dalla necessità – dettata, nel caso in esame, dalla natura storica della ricerca – di imporre un vincolo di coerenza totale al procedimento di «integrazione dell'informazione» che si ricava rispettivamente dai documenti e dal modello concettuale del loro contenuto. Gli autori stessi ammettono «correzioni» (317) – ossia, in ultima analisi, 'deformazioni' interpretative – dei documenti, suggerite dall'ontologia che ne rappresenta il contenuto. Sicché il vero problema pare consistere più propriamente nel sincronizzare le variazioni e nell'essere in grado di generarle e di verificarle con metodi computazionali. E dobbiamo tenere ben presente che la variazione, così come viene qui considerata, riguarda in primo luogo il *markup* e la struttura del modello concettuale e che solo in un secondo momento può essere considerata come una variazione del contenuto o dell'espressione del testo.

¹⁹ J. McGann, *Visible and invisible books: Hermetic images in n-dimensional space*, «New Literary History», 32:2 (2001), 283-300, pp. 290-91.

²⁰ Ciula et al., *Expressing complex associations in medieval historical documents*, p. 321.

7. E di nuovo, nel fare riferimento ad un altro esempio concreto, mette conto insistere proprio sulla variazione. Il caso in esame, presentato al congresso DH 2008 di Oulu da Øyvind Eide e Christian-Emil Ore,²¹ consiste in un tentativo riuscito di importare un'ontologia in un sistema di codifica, ossia, più precisamente, l'ontologia del Conceptual Reference Model del Comitato Internazionale per la Documentazione dell'International Council of Museums (ICOM), il cosiddetto CIDOC-CRM, nello schema di codifica proposto dalle *Guidelines* della Text Encoding Initiative (TEI).²² Non sono molti, a quanto è dato di sapere, i tentativi del genere, mentre sono molto più frequenti i casi in cui ci si sforza di estrarre una descrizione semantica o un modello concettuale dal *markup* inserito nel testo. Tra questi ultimi, basti ricordare la biblioteca digitale del New Zealand Electronic Text Centre (NZETC), dove, da documenti codificati secondo le direttive TEI, sono stati estratti i descrittori della mappa topica utilizzata per la ricerca e la navigazione nel sito.²³

Nel loro articolo, Eide e Ore giungono alla seguente conclusione:

La nostra scelta del CRM Core come base di verifica ha mostrato che con piccole aggiunte o aggiustamenti agli elementi della versione P5 delle *Guidelines* della TEI è possibile esprimere pacchetti CRM Core con elementi TEI.²⁴

Eccoci di nuovo riportati all'osservazione che l'integrazione di un modello concettuale in uno schema di codifica (*markup*) comporta qualche modifica delle sue regole generali. A quanto pare, quindi, ci troviamo qui di fronte ad un problema strutturale:

²¹ Ø. Eide and Ch E. Ore, *TEI and cultural heritage ontologies*, in L.L. Opas Hänninen et al. (eds.), «Digital Humanities 2008: Conference Abstracts», Proceedings of the 20th ALLC and ACH Joint International Conference (Oulu, 24-29 June 2008), Oulu, University of Oulu, 2008, pp. 115a-117a.

²² «Il CRM del CIDOC è un'ontologia formale creata per agevolare l'integrazione, la trasmissione (mediation) e l'interscambio di informazioni eterogenee riguardanti i beni culturali. [...] Nel 2005 il CRM fu riformulato come una semplice DTD XML, chiamata CRM Core, per permettere la codifica (*markup*) di metadati multimediali. Un pacchetto XML CRM Core può contenere informazioni su un certo elemento di una qualsiasi classe del CRM e su come esso può venire collegato ad altri oggetti attraverso eventi e proprietà» (ibid., p. 115a b).

²³ URL <<http://www.nzetc.org/>>. Cfr. Conal Tuohy, *Topic maps and TEI: Using topic maps as a tool for presenting TEI documents*, «New Zealand Electronic Text Centre Papers», URI = <<http://hdl.handle.net/10063/160>>.

²⁴ Eide and Ore, *TEI and cultural heritage ontologies*, p. 116b.

l'integrazione di informazione tratta dal *markup* in una rappresentazione semantica complessiva e, reciprocamente, l'integrazione di informazione ricavata dalla rappresentazione semantica nello schema di codifica induce una variazione nella struttura che riceve i nuovi contenuti importati. È allora possibile considerare questo fenomeno come una funzione e porsi il problema di trovarne una formulazione generale? Per il momento, prima di riflettere più a fondo su questo problema, conviene richiamare un'altra osservazione di Eide e Ore sui risultati del loro lavoro: importare l'ontologia CIDOC CRM nella TEI, essi affermano,

rapresenta un cambiamento notevole dei nostri orientamenti precedenti (Ore 2006), che assumevano l'ontologia come un modulo esterno alla TEI. (116b)

Dovremmo quindi cercare di importare le nostre rappresentazioni semantiche nel nostro schema di codifica e trasferire a loro volta le strutture di codifica nelle nostre rappresentazioni semantiche, o dovremmo tenere distinti i due tipi di rappresentazione? E ancora, abbiamo a che fare solo con una questione di integrazione, o non abbiamo piuttosto a che fare con un problema di reciproca dipendenza e interrelazione?

8. Prima di affrontare tali questioni è utile chiedersi, più in generale, che lezione possiamo trarre dagli esempi sopra descritti. Per rispondere a questa domanda possiamo osservare che non c'è niente di male nell'applicare il cosiddetto modello OHCO (Ordered Hierarchy of Content Objects) per i documenti codificati, purché si riconosca che qui 'contenuto' significa una mera stringa di caratteri, ovvero Parsed Character Data (PC DATA).²⁵ Si deve infatti ammettere che è più che lecito assegnare una certa struttura ad una stringa di caratteri, perché a ben vedere questa operazione è assolutamente necessaria per assegnare un modello strutturale e per applicare procedure computazionali all' 'immagine' del testo, come direbbe il criti-

²⁵ «Parsed data is made up of characters, some of which form character data, and some of which form markup. [...] All text that is not markup constitutes the character data of the document». (Extensible Markup Language (XML) 1.0 (Fifth Edition), W3C Recommendation 26 November 2008, URL <http://www.w3.org/TR/REC_xml/>).

co letterario,²⁶ oppure l' 'espressione' del testo, come direbbe il linguista,²⁷ ovvero la sua rappresentazione in caratteri ASCII, come direbbe l'informatico.²⁸ Il *markup* in realtà è 'struttura' ed assegna una struttura alla stringa di caratteri.²⁹ E possiamo considerare ed elaborare la struttura assegnata da forme di *markup* di tipo XML ad una stringa di caratteri come una struttura di dati costituita da una gerarchia ordinata di oggetti di contenuto (OHCO). Ma sappiamo anche che elaborando la struttura OHCO non elaboriamo il contenuto, ossia l'informazione veicolata dal testo. Il modello OHCO è un modello strutturale del documento, cioè dell' 'immagine,' o dell' 'espressione' del testo; ma non è un modello strutturale del testo e del suo contenuto informativo.

Il tipo di rappresentazione che offre il modello OHCO costituisce solo una struttura di dati e non è dotato di una semantica operativa da un punto di vista computazionale, qual è fornita solo dalla definizione di un modello di dati;³⁰ ma anche fornendo tale definizione, ciò che si ottiene elaborando una struttura OHCO è funzionale solo al trattamento dell'espressione del testo e delle sue varianti

²⁶ Segre, *Avviamento all'analisi del testo letterario*, p. 378; cfr. supra, nota 5.

²⁷ Hjelmslev, *I fondamenti della teoria del linguaggio*, pp. 52-65; cfr. supra, nota 4.

²⁸ Da un punto di vista informatico, il testo consiste in "informazione codificata come caratteri o come sequenze di caratteri", benché tale definizione si discosti da quella del testo inteso come "materiale letterario qual è originariamente scritto da un autore" (A. C. Day, *Text processing*, Cambridge, Cambridge University Press, 1984, p. 1).

²⁹ «Il markup [...] è rappresentazione della struttura e, al tempo stesso, è direttamente struttura» (D. Buzzetti, *Rappresentazione digitale e modello del testo*, in *Il ruolo del modello nella scienza e nel sapere*, Atti del Convegno (Roma, Accademia Nazionale dei Lincei, 27-28 ottobre 1998), Roma, Accademia Nazionale dei Lincei, 1999, (Contributi del Centro Linceo Interdisciplinare "Beniamino Segre", N. 100), 127-161, p. 152). Cfr. D. R. Raymond, F. W. Tompa and D. Wood, *Markup reconsidered*, paper presented at the First International Workshop on Principles of Document Processing, Washington DC, October 22-23, 1992, URL <<http://softbase.uwaterloo.ca/~draymon/papers/markup.ps>>, Abstract e p. 4.

³⁰ «SGML non è un modello di dati perché non definisce nessun operatore» (D. Raymond, F. Tompa and D. Wood, *From data representation to data model: Meta semantic issues in the evolution of SGML*, in "Computer Standards and Interfaces", 10 (1995), 25a 36b, p. 28a). Si ritiene opportuno distinguere tra le nozioni di 'struttura di dati' e 'modello di dati.' Un modello di dati è qui inteso come una struttura di dati insieme con la definizione degli operatori ad essa applicati. Cfr. A. Le Hors and R. S. Sutor, *Glossary*, in "Document Object Model (DOM) Level 3 Core Specification", Version 1.0, W3C Recommendation 07 April 2004, URL <<http://www.w3.org/TR/DOM-Level-3-Core/glossary.html>>: «A data model is a collection of descriptions of data structures and their contained fields, together with the operations or functions that manipulate them».

e se non si assegna, oltre alla struttura, un modello di dati adeguato al contenuto semantico nel suo insieme, non è possibile elaborare in modo soddisfacente l'informazione trasmessa dal testo.

Come si fa allora a dotare una rappresentazione del testo di funzionalità che agiscano anche sulle varianti interpretative? Si può affrontare il problema considerando, da un lato, una rappresentazione strutturale dell'*espressione* del testo, che è possibile ottenere col *markup* e, dall'altro, una rappresentazione strutturale del contenuto del testo; si tratta poi di trovare soluzioni operative adeguate per mettere in relazione i due tipi di rappresentazione organicamente e sistematicamente. A questo proposito, occorre tuttavia osservare che, in ogni caso, cambiando un elemento dell'una o dell'altra rappresentazione, la struttura complessiva di tale rappresentazione muta. In altri termini, entrambe le rappresentazioni sono caratterizzate da una struttura di tipo olistico o gestaltico. Quindi, una rappresentazione strutturale di natura formale formulata in termini esclusivamente insiemistici non riesce a tener conto in modo adeguato di un aspetto essenziale del fenomeno che intende descrivere. Se trascuriamo questa circostanza, ci esponiamo inevitabilmente agli inconvenienti dell'intelligenza artificiale 'forte' e dei sistemi di rappresentazione della conoscenza che le sono propri. Le difficoltà di questi metodi si avvertono chiaramente nella cosiddetta 'massima del formalista' enunciata da John Haugeland:

se si pensa alla sintassi, la semantica penserà a se stessa.³¹

L'intuizione fondamentale che giustifica un'affermazione di questo tipo postula un isomorfismo strutturale, ovvero una corrispondenza biunivoca tra le componenti di natura sintattica e quelle di natura semantica del testo. Questa circostanza si realizza pienamente solo per un linguaggio formale la cui forma logica corrisponda esattamente alla sua struttura semantica. Secondo Donald Davidson,

assegnare una forma logica ad un enunciato significa [...] descriverlo in modo tale da farlo rientrare nell'ambito di una teoria semantica.³²

³¹ J. Haugeland, *Artificial intelligence: The very idea*, Cambridge, Mass., MIT Press, 1985, p. 106.

Ma questo non è il caso del linguaggio naturale, le cui strutture semantiche e sintattiche sono aperte e indeterminate, una circostanza che non impedisce però alla sua rappresentazione digitale di essere sottoposta ad elaborazione automatica. Tutti sanno che un testo letterario in forma digitale può essere elaborato automaticamente senza imporre nessuna restrizione alla sintassi e alla semantica della lingua in cui è scritto e che, d'altra parte, una rappresentazione che ostacolasse la piena espressività di un'opera letteraria sarebbe di scarsa utilità allo studioso umanista. L'indeterminatezza è una caratteristica essenziale del linguaggio ordinario e del linguaggio letterario. L'indeterminatezza del testo si manifesta con grande evidenza nella relazione di compensazione tra la *sinonimia* e la *polisemia* delle espressioni, che si può descrivere brevemente in questo modo: ci sono molti modi diversi per esprimere la stessa cosa, proprio come la stessa espressione può essere intesa in molti modi diversi.

E questo è l'aspetto veramente degno di nota della rappresentazione digitale del testo, il fatto cioè che può essere elaborata automaticamente, senza dovere necessariamente sottostare ai vincoli formali di un linguaggio simbolico artificiale, oppure, se vogliamo, a quelli di un sistema di rappresentazione della conoscenza proprio dell'intelligenza artificiale. Nel nostro caso, ciò che deve essere elaborato in modo formale è la struttura del tipo di dato che rappresenta l'espressione del testo – la sequenza lineare dei caratteri – cioè la struttura che il *markup* assegna alla stringa, non la struttura del linguaggio che la stringa dei caratteri rappresenta. E anche per la rappresentazione del contenuto del testo si possono fare considerazioni analoghe. Ciò che deve essere elaborato automaticamente è la struttura del tipo di dato che esprime la descrizione semantica dell'informazione testuale. Il problema consiste quindi, come si è detto, nel mettere in relazione tra loro, sistematicamente ed operativamente, le strutture dei due tipi di dato e per rappresentare dinamicamente l'interazione tra le variazioni letterali e le variazioni interpretative del testo, possiamo trarre vantaggio proprio dall'indeterminatezza del testo e dall'ambivalenza del *markup*.

Ora, per mettere in relazione in modo sistematico e operativo rappresentazioni strutturate dei dati testuali – ovvero «strutture sintattiche di *markup*» – da un lato, e rappresentazioni strutturate del

³² D. Davidson, *Essays on actions and events*, Oxford, Oxford University Press, 2001, p. 144.

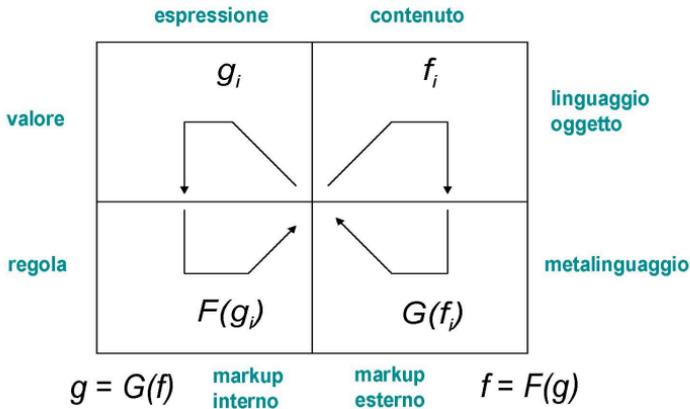
contenuto informativo corrispondente – ovvero «oggetti, proprietà e relazioni» – dall'altro, possiamo fare affidamento proprio sul fondamentale principio di indeterminazione che vige per il linguaggio naturale. Se ci riferiamo a rappresentazioni strutturate di entrambi i generi, questo principio si può enunciare affermando, con Dubin e Birnbaum, «che lo stesso *markup* può comunicare significati diversi in contesti diversi», e «che il *markup* può comunicare lo stesso significato in modi diversi usando una sintassi molto diversa». ³³ Di conseguenza, concludono Dubin e Birnbaum, una relazione strutturale tra i due tipi di rappresentazione può essere espressa per mezzo di «una nuova codifica dei documenti con un *markup* più ampio e dettagliato» da una parte, oppure «può essere espressa in modo serializzato nella forma di una descrizione RDF o di una mappa topica» dall'altra (8). Ma i termini di questa alternativa rappresentano rispettivamente i valori di una funzione che connette variazioni letterali con variazioni interpretative, da un lato, e i valori della funzione inversa, dall'altro.

9. Possiamo cercare di illustrare questi fenomeni per mezzo di un diagramma (Fig. 1) che esprime dimensioni diverse del *markup*. ³⁴ Come si è detto, possiamo avere diverse rappresentazioni e distinzioni strutturali che si riferiscono rispettivamente all'*espressione* e al *contenuto* del testo. Inoltre, un elemento o espressione del linguaggio di *markup* può essere intesa sia come un'istruzione per un'operazione di ristrutturazione, o più brevemente come una *regola*, sia come una variante introdotta a seguito di un'operazione di ristrutturazione, di cui costituisce il risultato o il *valore*. Infine, un'espressione del linguaggio di *markup* può essere considerata come appartenente al *linguaggio oggetto* del testo, oppure ad un *metalinguaggio* che lo descrive. Tutte queste dimensioni possono essere rappresentate nel diagramma. Per finire, possiamo richiamare il fatto che le strutture espresse dal *mark-*

³³ D. Dubin and D. Birnbaum, *Interpretation beyond markup*, in B. T. Usdin (ed.), «Proceedings of the Extreme Markup Languages 2004 Conference» (Montreal, Quebec, August 2004), p. 1, URL <<https://www.ideals.illinois.edu/bitstream/handle/2142/11838/EML2004Dubin01.pdf?sequence=2>> [02. 08.2010].

³⁴ Le osservazioni che seguono circa il diagramma qui proposto riprendono nella sostanza le considerazioni svolte nel saggio D. Buzzetti, *Digital editions and text processing*, in M. Deegan and K. Sutherland (eds.), *Text editing, print, and the digital world*, Aldershot, Ashgate, 2009, 45-62, alle pp. 57-60.

up possono essere «inserite (*embedded*)» e «presenti nei dati»,³⁵ oppure rappresentate a parte e costituire una «struttura non inserita (*non-embedded structure*)» nei dati; di conseguenza, possiamo distinguere tra struttura o *markup* interno e struttura o *markup* esterno, «il cosiddetto *markup* non in linea (*out-of-line markup*)» (4). Anche queste due dimensioni possono essere incluse nel nostro diagramma, a completamento dello spazio pluridimensionale che esso rappresenta.



– Figura 1 –

Questo diagramma può essere usato per rappresentare i fenomeni di instabilità testuale e può offrire un modello dinamico della mobilità del testo. Esso mostra che l'ambiguità diacritica del *markup*, lungi dal costituire «un ostacolo all'elaborazione automatica dell'informazione testuale», può essere utile proprio a questo scopo, perché ci permette di fornire «una rappresentazione formale della dinamica testuale». ³⁶ Per illustrare il diagramma, possiamo fare ancora riferimento alla punteggiatura, che come si è detto può essere considerata a tutti gli effetti una forma di *markup*. Ciò che un editore moderno fa, inserendo, per esempio, segni di punteggiatura in un testo medievale, non è altro, in realtà, che aggiungere *markup*. Supponiamo che g_i sia il risultato di un'operazione di questo genere,

³⁵ Raymond et al., *Markup reconsidered*, p. 3.

³⁶ D. Buzzetti and J. McGann, *Critical editing in a digital horizon*, in *Electronic textual editing*, edited by L. Burnard, K. O'Brien O'Keefe, and J. Unsworth, New York, The Modern Language Association of America, 2006, 51-71, p. 64.

come l'inserimento di una virgola o, se vogliamo, di un marcatore qualunque nella stringa di caratteri che rappresenta il testo. Il segno g_j è il *valore* dell'operazione di marcatura e costituisce un'espressione del *markup interno* al testo. Non appena tracciato, il segno diventa parte dell'*espressione* del testo e un costrutto del suo *linguaggio oggetto*. Possiamo quindi collocarlo nel settore in alto a sinistra del nostro diagramma. In quanto nuovo elemento dell'espressione del testo, esso costituisce una variante testuale.

Ma il nostro segno può anche essere considerato come una *regola* per un'operazione di mappatura (*mapping*) dell'espressione sintattica della quale fa parte nel dominio di ciò che il testo significa. L'operazione di mappatura è una funzione di g_j che possiamo indicare come $F(g_j)$. Considerato come una funzione, il segno assume prerogative linguistiche diverse e diviene un'asserzione di ordine superiore del linguaggio oggetto. Un'asserzione di questo tipo ha portata metalinguistica e può essere convertita nell'asserzione ad essa logicamente equivalente espressa in un vero e proprio *metalinguaggio*. Possiamo così collocarla nel quadrante in basso a sinistra del nostro diagramma.

L'operazione di mappatura $F(g_j)$, rappresentata dall'espressione di *markup* considerata come una regola, assegna una certa struttura al contenuto del testo. Sia $f_i = F(g_j)$ il risultato della mappatura, o il *valore* della funzione. Come valore dell'operazione di mappatura, f_i designa un elemento strutturale del dominio di ciò che il testo significa e fa parte della struttura esterna descritta dal testo. Non appartiene quindi all'espressione, o alla struttura interna del testo, e non ne fa parte, ma deve invece essere considerato come un elemento del *markup esterno* (o *stand-off markup*). Poiché consiste in un'assegnazione di struttura al dominio di ciò che è significato dal testo, esso costituisce una variante interpretativa del suo contenuto. Esso può anche essere designato da espressioni appropriate di un linguaggio per la modellazione dei dati (*data modelling language*) – quali, per esempio, ERA, RDF, UML, ecc. – in grado di assegnare un modello formale adeguato al contenuto del testo e tale da costituire un *linguaggio oggetto* conveniente per la designazione degli elementi strutturali che lo compongono. L'elemento considerato f_i , appartenente alla struttura esterna del testo, deve così essere collocato nel quadrante superiore sinistro del diagramma.

A sua volta il costrutto strutturale f_i che appartiene al contenuto del testo può essere considerato come una *regola*, o un'istruzione, per

un'operazione di mappatura e di ristrutturazione del dominio dell'espressione del testo. Possiamo designare questa funzione come $g_j = G(f_i)$. Anche in questo caso, l'espressione che designa l'elemento strutturale esterno f_i , considerato come una funzione, assume prerogative linguistiche diverse e diventa un'asserzione del secondo ordine del linguaggio oggetto in cui viene espresso il *markup* esterno, ovvero del linguaggio di modellazione dei dati usato per descrivere il dominio di ciò che è significato dal testo. Tale espressione, poiché assume f_i come suo argomento, ha una portata metalinguistica e può essere convertita in un costrutto equivalente espresso in un opportuno *metalinguaggio* riferito al testo, o più precisamente al suo contenuto. Possiamo quindi collocare l'espressione funzionale $g_j = G(f_i)$ nel quadrante inferiore destro del diagramma. L'espressione funzionale $G(f_i)$ ha g_j come suo valore e così il ciclo si chiude. Costrutti del *markup* esterno f_i , che dipendono dalla mappatura di costrutti del *markup* interno g_j , possono essere ri-mappati su costrutti del *markup* interno g_j .

Il *ciclo* che abbiamo descritto può essere considerato come lo svolgersi e il dispiegarsi dell'instabilità testuale e dell'indeterminatezza del testo. L'indeterminatezza del testo può essere descritta come una relazione di non-identità del testo con se stesso. Per esprimerci in modo più formale, possiamo dire che la non-identità del testo con se stesso è logicamente equivalente a un endomorfismo, ovvero a una mappatura del testo su se stesso. Con una formula:

$$(A = A \text{ se e solo se } A \neq A) \Leftrightarrow A \xrightarrow{\text{markup}} A.^{37}$$

Il *markup* può quindi essere considerato come l'espressione della funzione autoriflessiva che collega elementi semiotici del testo ad altri elementi semiotici dello stesso testo.

La relazione di non-identità del testo con se stesso è una relazione vincolata. La struttura del testo è stata definita come «l'assieme delle relazioni latenti» tra le sue parti,³⁸ ovvero tra tutti i suoi costrutti strutturali virtuali; e se poi si concepisce la struttura come «la totalità di tutte le sue varianti testuali e interpretative», è possibile anche considerare il testo come una «unità virtuale identica a se

37

Buzzetti, *Rappresentazione digitale e modello del testo*, p. 156.

38

Segre, *Avviamento all'analisi del testo letterario*, p. 44.

stessa». Entro questa totalità, o unità virtuale, possiamo tracciare una prima distinzione tra l'espressione e il contenuto del testo e «la variazione di ciascuna delle due parti componenti dipende dall'invarianza della sua rispettiva controparte». ³⁹ La relazione del testo con se stesso è quindi una relazione uno-molti ed entrambe le mappature possono assumere più di un valore.

Tali funzioni di mappatura possono essere svolte dal *markup* che esprime in forma esplicita caratteristiche strutturali implicite nel testo. Ma il *markup* è ambivalente e può inoltre essere tanto interno quanto esterno al testo. Il *markup* può quindi convertire varianti testuali in varianti interpretative e viceversa: infatti, può mappare componenti strutturali dell'espressione in componenti strutturali del contenuto, oppure può mappare costrutti strutturali del contenuto in costrutti strutturali dell'espressione. Poiché tanto le espressioni del *markup* interno, quanto quelle del *markup* esterno possono assumere natura performativa oltre che dichiarativa, l'inversione della mappatura tra costrutti semantici e costrutti sintattici, ovvero dell'espressione e del contenuto del testo, può essere descritta esplicitamente e il *markup* può fornire una rappresentazione dei vincoli dinamici di compensazione tra i costrutti strutturali dell'espressione e del contenuto del testo. La mobilità interna del testo può così essere descritta come «un ciclo continuo di azioni di compensazione tra stati di determinazione e di indeterminazione dell'espressione e del contenuto del testo» (66).

10. Con ciò abbiamo cercato di delineare uno schema formale per la struttura semiotica della rappresentazione digitale del testo. Si possono senz'altro proporre schemi più adeguati, ma lo scopo era soprattutto quello di far riflettere sulla complessità dei fenomeni da analizzare. Pare infatti del tutto improbabile che si possa assicurare un'efficace e adeguata elaborazione dell'informazione testuale che non operi sul fondamento di una rappresentazione funzionale esauritiva di tutti i suoi elementi e i suoi vincoli di natura semiotica. L'applicabilità del modello qui proposto può tuttavia essere esemplificata facendo riferimento a realizzazioni concrete presentate rispettivamente ai congressi Digital Humanities 2008 e 2009.⁴⁰

³⁹ Buzzetti and McGann, *Critical editing in a digital horizon*, p. 64.

⁴⁰ Cfr. D. Buzzetti and M. Rehbein, *Towards a model for dynamic text editions*, in L.L. Opaš Hänninen et al. (eds.), «Digital Humanities 2008: Conference Abstracts», Proce-

Il primo esempio riguarda l'edizione di un testo «a più livelli», ossia modificato a più riprese, realizzata da Malte Rehbein.⁴¹ Il problema consiste nel determinare la successione delle modificazioni del testo ed è affrontato raccogliendo tutte le informazioni necessarie a stabilire l'ordine degli interventi di revisione in un *database* esterno al testo. La rappresentazione dell'informazione riguardante «l'espressione del testo» e le «conoscenze esterne, di natura procedurale e contestuale» sono mantenute separate (79a). Le diverse varianti testuali sono rappresentate mediante il *markup* inserito nel testo, mentre l'informazione esterna, che specifica «il ruolo svolto da una certa porzione del testo» (79b) nel contesto degli avvenimenti storici che lo riguardano, è rappresentata in un *database*. Il collegamento tra l'informazione esterna riguardante il motivo delle revisioni e l'informazione interna riguardante l'espressione del testo permette di ricostruirne «lo sviluppo evolutivo» (78b). L'applicabilità del modello dinamico proposto è mostrata dal fatto che «i cambiamenti nel *markup* modificano le registrazioni del *database*» e che, reciprocamente, i cambiamenti nelle «registrazioni del *database* producono nuovo *markup* per la visualizzazione del testo».⁴²

Il secondo esempio riguarda lo studio dello «sviluppo diacronico dei concetti», nel *De gli eroici furori* di Giordano Bruno, intrapreso dal gruppo di ricerca del centro Signum della Scuola Normale Superiore di Pisa.⁴³ Per la rappresentazione semantica delle variazioni nelle relazioni tra i concetti «sono state scelte le mappe topiche, che hanno caratteristiche simili a quelle del processo naturale di interpretazione del testo». La rappresentazione del testo, in *file* «codificati in formato TEI», e delle mappe topiche, «codificate in XTM», entrambi linguaggi XML, permette al motore di ricerca di presentare «in un'unica visualizzazione» le «occorrenze del testo e le 'costellazioni

dings of the 20th ALLC and ACH Joint International Conference (Oulu, 24-29 June 2008), Oulu, University of Oulu, 2008, pp. 78-81; D. Buzzetti, *How to build a textual data model?*, in D. Buzzetti et al., *For a dynamic model of textual variation: What do we need?*, in «Digital Humanities 2009: Conference Abstracts», Proceedings of the 21th ALLC and ACH Joint International Conference (University of Maryland, College Park, June 22-25, 2009), College Park, The Maryland Institute for Technology in the Humanities (MITH), 2009, 8b 14a, pp. 12b 14a.

⁴¹ Buzzetti and Rehbein, *Towards a model for dynamic text editions*, p. 78b.

⁴² D. Buzzetti, *How to build a textual data model?* in Buzzetti et al., *For a dynamic model of textual variation: What do we need?*, p. 13a.

⁴³ A. Isolani, et al., *Topic maps and MVD for the representations of interpretative variants*, p. 9b.

concettuali?» ad esse riferite (10a). Anche in questo caso, l'applicabilità del modello dinamico qui proposto è dimostrata dalla circostanza che «un cambiamento nella mappa topica genera una riorganizzazione delle relazioni intratestuali e che, al contrario, una nuova organizzazione delle relazioni testuali genera una diversa mappa topica».44

Un ulteriore passo in avanti è costituito dalla possibilità di ottenere una rappresentazione formale omogenea, tanto delle variazioni testuali, quanto delle variazioni interpretative. Il ricorso ai grafi MVD introdotti da Desmond Schmidt⁴⁵ permette di risolvere il problema della sovrapposizione di strutture gerarchiche concomitanti, irrisolto dal modello OHCO, e di ottenere una rappresentazione unitaria e complessiva, una sorta di «somma logica»,⁴⁶ non solo delle varianti testuali, ma anche delle varianti interpretative, che possono anch'esse venire rappresentate in forma serializzata e dunque, in ultima analisi, in forma testuale. Nel caso delle mappe topiche, per esempio, «questo compito è reso possibile dal fatto che le diverse interpretazioni sono descritte facendo ricorso allo standard XTM che, in quanto linguaggio XML, è anch'esso un testo».47 Si apre così la possibilità di mettere in relazione un determinato percorso all'interno del grafo delle varianti testuali con percorsi compatibili all'interno del grafo delle varianti interpretative, e viceversa, unificando in un unico modello sia la mappatura delle strutture del *markup* nelle strutture della rappresentazione semantica, sia la mappatura inversa. La somma logica delle varianti interpretative e la somma logica delle varianti testuali, rappresentate ciascuna da un grafo MVD, determinano il campo delle variazioni possibili e le relazioni uno-molti tra le singole varianti dell'un tipo e le diverse varianti compatibili del tipo opposto e corrispondente.

⁴⁴ Buzzetti, *How to build a textual data model?*, p. 13a.

⁴⁵ D. Schmidt and R. Colomb, *A data structure for representing multi version texts online*, in «International Journal of Human Computer Studies», 67:6 (2009), 497-514. «La struttura di dati MVD [Multi Version Document] è costituita da un grafo orientato con un nodo iniziale e un nodo finale»; in tale struttura, «ciascuno strato del testo e ciascuna descrizione interpretativa è rappresentata da un percorso diverso all'interno del grafo» (Buzzetti, *How to build a textual data model?*, p. 13a).

⁴⁶ M. Thaller, *Historical information science: Is there such a thing? New comments on an old idea*, in *Discipline umanistiche e informatica*, a cura di T. Orlandi, Roma, Accademia Nazionale dei Lincei, 1993, p. 64.

⁴⁷ A. Isolani, et al., *Topic maps and MVD for the representations of interpretative variants*, p. 10a.

Anche a questo proposito, il riferimento ai casi citati intende fornire soltanto l'indicazione di un percorso possibile e già praticato per ottenere una rappresentazione formale dei fenomeni di instabilità e di indeterminazione del testo e rendere così possibile l'elaborazione dell'informazione testuale trasmessa da un'edizione digitale che cerchi di sviluppare appieno le proprie funzionalità computazionali e non si limiti, con evidente vantaggio, alla pura visualizzazione e rappresentazione del testo. Nella sostanza, tuttavia, si è cercato di mostrare che l'indeterminatezza dei fenomeni testuali e l'ambivalenza delle espressioni diacritiche possono costituire non una anomalia da eliminare, ma una risorsa utile e preziosa per superare i gravi limiti di molti procedimenti proposti da forme oramai irrimediabilmente datate di intelligenza artificiale.