

## Dino Buzzetti

### Ambiguità diacritica e markup. Note sull'edizione critica digitale

#### Sommario

0. [Premessa](#)
1. [Adeguatezza dell'edizione digitale](#)
2. [La struttura dell'informazione](#)
3. [Elaborazione dell'informazione strutturata](#)
4. [Il markup e la struttura dell'espressione](#)
5. [La struttura del contenuto e il database](#)
6. [Instabilità e compensazione strutturale](#)
7. [Rappresentazione formale della struttura del testo](#)
8. [Formulazione della legge di compensazione](#)
9. [Per un'analisi delle rappresentazioni strutturali del testo](#)

#### [Bibliografia](#)

---

#### 0. [Premessa](#)

La presentazione (a) dell'edizione-database del *Commentario* di Gentile da Cingoli all'*Isagoge* di Porfirio ( [B IP](#), 145-54) realizzata col DBMS 'kleio' ( [T K](#) ) e (b) del programma DTE, ovvero Digitale TextEdition ( [R DDT](#) ), un'applicazione che ne estende le funzionalità, richiede alcune considerazioni introduttive di ordine teorico.

#### 1. [Adeguatezza dell'edizione digitale](#)

La realizzazione di edizioni in forma di *database* si è imposta come soluzione opportuna per l'edizione di testi prodotti per l'insegnamento nell'università di arti e medicina di Bologna nei secoli XIV e XV ( [BT ICT](#) ; [BPT LM](#) ). In tutti i casi presi in esame, la recensione presenta immancabilmente tradizioni testuali di tipo accentuatamente 'fluido', per ragioni che dipendono con tutta evidenza dalle forme concrete della pratica scolastica. Queste stesse ragioni rendono storicamente preferibile la restituzione dell'intera tradizione testuale alla ricostruzione di un'improbabile versione unica e autentica del testo: così come si può certamente affermare che la diffusione delle dottrine fosse il risultato della produzione dei testi, si può con altrettanta sicurezza affermare che la forma stessa della produzione testuale fosse il risultato dell'impiego effettivo dei testi nella pratica diretta di insegnamento ( [B FA](#), 465). La scelta dell'edizione come *database* è quindi stata determinata dall'esigenza di ottenere una forma di rappresentazione adeguata per tradizioni testuali di tipo marcatamente 'fluido' ( [BR TFDE](#), 14-39). Generalizzando, si può sostenere che solo le esigenze analitiche della ricerca possono fornire, oltre alle motivazioni euristiche, i **criteri di adeguatezza** per un'eventuale rappresentazione digitale del testo. L'edizione digitale si giustifica pienamente solo se può fornire soluzioni a problemi difficilmente affrontabili in modo diverso ( [B RDMT](#), 130).

#### 2. [La struttura dell'informazione](#)

Nel caso in esame, il *database* non realizza solo un'edizione-archivio di tutti i testimoni esistenti di una certa tradizione testuale, ma li organizza in un sistema coerente di informazione strutturata che può fornire risposte precise a specifiche esigenze analitiche. Allo stesso modo, l'inserimento delle **immagini digitali** dei manoscritti non fornisce solo, a soddisfazione dell'occhio, una riproduzione fisica degli originali, ma un diverso tipo di rappresentazione logica dell'informazione testuale, elaborabile con procedure automatiche appropriate. Ogni forma specifica di rappresentazione dell'originale restituisce l'informazione contenuta nel documento come **informazione strutturata** secondo modalità diverse, che rendono possibili operazioni analitiche a volta a volta diverse.

L'immagine di un documento non ne rende superflua, ad esempio, la trascrizione diplomatica. L'informazione grafica ricavabile da un'immagine non è l'informazione testuale ricavabile da una trascrizione. Tipi diversi di rappresentazione dell'informazione permettono tipi diversi di elaborazione. Un'immagine digitale (*bitmap*) può essere considerata come una trascrizione binaria del suo contenuto visivo e permette di analizzarne gli elementi grafici. Una trascrizione diplomatica rappresenta in forma esplicita alcuni elementi strutturali del documento e permette di applicare procedure analitiche che operano su tali elementi. Ogni forma distinta di trascrizione può essere considerata una forma distinta di analisi ( [B IP](#), 145-48). Ciascuna forma di trascrizione trascoglie elementi strutturali diversi e rende praticabili forme di analisi diverse. In generale, la possibilità di applicare determinati procedimenti analitici dipende dall'organizzazione strutturale assegnata all'informazione attraverso la sua forma di rappresentazione.

Per ragioni diverse, sia la ricostruzione filologica, sia l'interpretazione letteraria della produzione testuale richiedono **rappresentazioni strutturali non lineari** del testo. La restituzione critica del testo agisce sulla stratificazione e la molteplicità delle lezioni varianti accumulate nel corso della sua trasmissione materiale; l'interpretazione di qualsivoglia redazione del testo agisce sui modi diversi di intenderne il contenuto e sulla varietà delle sue possibili ricostruzioni strutturali. Il filologo riporta ad una singola rappresentazione strutturale complessa e non lineare l'espressione lineare delle diverse varianti testuali, il critico ricava da

una singola rappresentazione lineare del testo l'insieme complesso e non lineare delle diverse interpretazioni strutturali del suo contenuto ( [B TF](#), 87-89; [B VRI](#), 255). Solo un'organizzazione non lineare dell'informazione può rispondere alle esigenze analitiche della critica testuale e della critica letteraria ( [BR TFDE](#), 36-37). La natura assolutamente astratta di questa generalizzazione non le toglie il carattere di condizione necessaria e fondamentale per l'adeguatezza della rappresentazione digitale del testo.

Tuttavia il filologo opera su varianti testuali, ossia su elementi strutturali dell'**espressione** del testo, mentre il critico letterario opera su varianti interpretative, ossia su elementi strutturali del suo **contenuto**. Occorre quindi considerare in che modo la struttura assegnata all'espressione può essere messa in relazione con la struttura assegnata al contenuto.

### 3. [Elaborazione dell'informazione strutturata](#)

La funzionalità della rappresentazione digitale del testo dipende dalla struttura assegnata all'informazione e dalle operazioni ad essa applicabili. L'informazione testuale viene rappresentata mediante sequenze lineari, o stringhe, di caratteri codificati in forma binaria. Il **markup**, ossia l'inserimento di marcatori (*tags*), permette di assegnare una struttura alla rappresentazione del testo ( [RTW MR](#), 1-4 a ), distinguendo, nella sequenza dei caratteri codificati, parti diverse con funzioni diverse; a loro volta, gli elementi costitutivi dell'informazione testuale possono essere organizzati in un **database**, che ne descrive e ne analizza le relazioni strutturali. La struttura assegnata all'informazione testuale, rappresentata dalla sequenza dei caratteri codificati, dipende dal sistema di *markup*; le operazioni eseguibili sugli elementi di tale struttura dipendono dal sistema di gestione della base di dati (il Data Base Management System, o DBMS). Per verificare l'adeguatezza di un'edizione digitale è quindi necessario considerare in che modo il sistema di *markup* e il sistema di gestione della base di dati si rapportano alla rappresentazione digitale del testo.

Ora, il *markup* deve necessariamente agire sull'espressione del testo, mentre il *database* ne deve altrettanto necessariamente prescindere, per poter applicare un modello operativo assolutamente formale ed astratto agli elementi costitutivi del suo contenuto. Occorre quindi evitare che si produca qualsiasi discrasia tra la struttura del contenuto e la struttura dell'espressione digitale del testo; occorre, in altri termini, che il **modello semantico** del contenuto del testo sia applicabile alla **struttura sintattica** della sua rappresentazione lineare.

Questa condizione non è soddisfatta, in generale, dai sistemi di marcatura conformi allo standard SGML (Standard Generalized Markup Language). La grammatica formale che regola l'uso dei marcatori ammette soltanto una segmentazione sequenziale del testo e ammette tra i diversi elementi solo relazioni di subordinazione gerarchica in unità più comprensive, anch'esse ordinate sequenzialmente e in successione - in sostanza, permette di rappresentare la struttura del testo solo come una **struttura gerarchica ad albero** i cui **elementi**, così come i nodi a cui afferiscono, siano **ordinati linearmente**. In un componimento poetico, per esempio, non è possibile rappresentare contemporaneamente una segmentazione per versi e una segmentazione per enunciati, o per altri costrutti grammaticali che si estendono oltre i limiti di uno stesso verso: la sintassi dei marcatori non permette di contrassegnare i diversi elementi strutturali di un *enjambement*. A tali limitazioni della capacità espressiva si aggiunge una peculiarità dalle conseguenze ancor più gravi e compromissorie. I sistemi SGML assumono come *constraint language*, ossia come linguaggio formale per definire vincoli operativi sugli elementi della struttura testuale, la *context-free grammar* definita dalla DTD (Document Type Definition) del documento SGML, ossia la sintassi formale che disciplina l'uso dei suoi marcatori ( [SH CDH](#), 30 a ). Con ciò le restrizioni della sintassi sono immotivatamente estese alla semantica: la possibilità di assegnare solo relazioni lineari e gerarchiche alla struttura dell'espressione porta all'impossibilità di operare sulle relazioni non lineari della struttura del contenuto. I documenti in formato SGML possono valersi di una procedura automatica per il controllo della congruenza sintattica, ma non permettono di definire procedure operative applicabili alle relazioni testuali di natura non lineare ( [SH CDH](#), 41).

### 4. [Il markup e la struttura dell'espressione](#)

Ma quale relazione sussiste tra la struttura dell'espressione e la struttura del contenuto? Quale connessione sussiste tra la sintassi e la semantica della rappresentazione digitale del testo? L'analisi dello status del *markup* e del *database* in relazione alla forma digitale del testo permette di delineare una semiotica della sua rappresentazione informatica e di porre i primi elementi di una **teoria del testo digitale**.

Il *markup* è essenzialmente notazionale ( [B RDMT](#), 146) e agisce direttamente sull'espressione del testo; anzi, può essere considerato come una parte dell'espressione stessa del testo, rispetto alla quale svolge una funzione propriamente **diacritica**. Il *markup* è al tempo stesso inserito e separabile dal testo, assegna una struttura all'espressione del testo ed è esso stesso struttura appartenente al testo ( [RTW MR](#), 1-4 b ). Può, come la punteggiatura, o come ogni altro segno diacritico, essere considerato come una descrizione metalinguistica della struttura del testo, o come un'estensione del sistema stesso di scrittura, che permette di esplicitare caratteri altrimenti impliciti del testo. Il *markup* può quindi essere considerato, rispettivamente, come una forma di notazione metalinguistica o come un'ampliamento delle risorse espressive del linguaggio-oggetto, del linguaggio stesso costitutivo del testo. Assumendo l'equivalenza effettiva delle due diverse forme di notazione, se ne trae che il *markup* svolge rispetto al testo una funzione **autoriflessiva** e che può essere considerato, rispettivamente, come un'estensione dell'espressione che ne esplicita la struttura e le regole d'uso, o come una forma di riferimento esterno ai suoi aspetti funzionali e strutturali. Nell'un caso la forma logica del linguaggio costitutivo del testo deve comprendere forme di predicazione del secondo ordine, ossia forme espressive in grado di rappresentare aspetti strutturali del suo stesso funzionamento; nell'altro si tiene separata dal testo la rappresentazione dei suoi aspetti strutturali e si ricorre ad un linguaggio diverso per poterli descrivere.

Il carattere essenzialmente **notazionale** del *markup* non impedisce peraltro che esso possa rappresentare aspetti strutturali del contenuto, oltre che dell'espressione del testo. Occorre tuttavia tener conto del modo in cui una forma di rappresentazione connessa necessariamente con la struttura dell'espressione del testo riesce a descrivere aspetti strutturali del testo che sono propri del suo contenuto e che sono svincolati, in generale, dalla struttura della sua espressione. Infatti, per la natura essenzialmente notazionale che gli è propria, il *markup* assegna necessariamente una struttura all'espressione del testo anche quando serve ad esprimere aspetti strutturali del suo contenuto.

### 5. [La struttura del contenuto e il database](#)

Una rappresentazione strutturale del contenuto del testo, indipendente dalla struttura della sua espressione, può invece essere ottenuta con un *database*. Un *database* viene di solito utilizzato come forma di rappresentazione astratta delle nostre conoscenze sul mondo ( [GL CS](#), 248); esso può quindi essere assunto come descrizione strutturale ed astratta dei contenuti specifici di cui parla il testo, in breve, come **modello** strutturale del suo contenuto. Il *database* prescinde dall'espressione del testo e dalla sua struttura. Mentre il *markup* è relativo al formato del documento e ne esprime la forma strutturale astratta, il *database* prescinde dalla forma di rappresentazione dei dati e ne rappresenta la struttura semantica formale ( [RTW DRDM](#), 3-6). Il *markup* costituisce

l'espressione astratta del formato o della struttura (sintattica) dei dati, il *database* costituisce la rappresentazione astratta del modello (semantico) dei dati e ne implementa il formalismo operativo ( [J DR](#), 75-76). Se si fa riferimento, come nel nostro caso, a insiemi di dati costituiti da sequenze di caratteri codificati, ossia alla rappresentazione digitale del testo, si può così affermare che il *markup* rende evidente e presenta in forma esplicita la struttura astratta dell'espressione del testo e che il *database* costituisce il modello o la rappresentazione astratta della struttura del suo contenuto.

Risulta quindi chiaro che il modello del testo offerto da un *database* non è vincolato alla struttura lineare della sua rappresentazione. Il *database* può essere in qualche modo concepito come la struttura profonda, assolutamente **non lineare**, da cui può essere generata la struttura superficiale e lineare del testo ( [B RDMT](#), 147-48). Per assicurare l'adeguatezza funzionale della rappresentazione digitale del testo occorre assicurare che non esistano vincoli limitativi o condizionamenti reciproci tra la forma di rappresentazione strutturale del contenuto e la forma di rappresentazione strutturale dell'espressione del testo. Il sistema di *markup* dev'essere in grado di progettare la struttura non necessariamente lineare del contenuto sulla struttura lineare dell'espressione e il sistema di gestione del *database* dev'essere in grado di offrire un modello operativo in grado di associare all'espressione lineare del testo la rappresentazione formale delle diverse relazioni strutturali non necessariamente lineari del suo contenuto. In generale, la **congruenza** fra le proprietà strutturali lineari dell'espressione e le proprietà strutturali del contenuto non può essere assicurata da forme di *markup* vincolato (*strongly embedded*), cioè da sistemi di *markup* in cui la posizione del marcatore sia significativa, ossia portatrice di informazione ( [RTW MR](#), 3-4). I sistemi di *markup* a base SGML non soddisfano, in generale, questa condizione.

Ma è possibile cercare di descrivere in qualche modo la relazione strutturale tra la forma dell'espressione e la forma del contenuto del testo?

## 6. **Instabilità e compensazione strutturale**

Tra la struttura dell'espressione e la struttura del contenuto del testo esiste una forma di **compensazione**. ( [B RDMT](#), 151 ss.) Ad un'espressione data è possibile assegnare diversi contenuti interpretativi e ad un contenuto interpretativo dato possono corrispondere diverse espressioni del testo. La fissità e l'invarianza dell'espressione, o del contenuto, comportano reciprocamente la varianza e l'indeterminazione del contenuto, o dell'espressione. In breve, ci possono essere diversi modi di intendere quello che si dice e diversi modi di dire quello che si intende. L'equivalenza di diverse **espressioni sinonime** presuppone l'identità di un medesimo contenuto interpretativo e, inversamente, la compatibilità di diversi contenuti interpretativi rimanda all'identità di una medesima **espressione polisemica**. Sicché il campo di variabilità del contenuto è vincolato dall'identità dell'espressione e il campo di variabilità dell'espressione è vincolato dall'identità del contenuto. L'indeterminazione di questo rapporto conferisce mobilità e dinamicità alle strutture testuali, che possono fissarsi come forme strutturali definite dell'espressione o del contenuto e determinare, reciprocamente, l'instabilità del contenuto o dell'espressione corrispondente. Di conseguenza, il complesso delle relazioni interne tra le parti costitutive del testo resta per la maggior parte implicito e la struttura del testo può essere definita come l'insieme delle relazioni latenti tra i suoi elementi strutturali. ( [S ATL](#), 44)

I processi dinamici del testo sono di natura olistica. Nella relazione di compensazione tra determinazione e indeterminazione, la diversità è riferita alle parti, l'identità al tutto. È l'intera espressione del testo che si riferisce, polisemicamente, a interpretazioni diverse ed è alla totalità del contenuto che si riferiscono, sinonimicamente, le espressioni diverse che lo descrivono. Sicché l'espressione, nella sua interezza, è compatibile con analisi diverse e con diverse configurazioni delle parti del modello, così come il modello, nella sua totalità, è compatibile con articolazioni diverse della struttura dell'espressione e diverse disposizioni delle sue parti. Ciò che entra invariabilmente in gioco è il **rapporto tra le parti e il tutto**, benché di volta in volta si assuma come totalità definita e identica a se stessa il contenuto o l'espressione del testo e come rispettiva articolazione strutturale indeterminata l'insieme delle varianti testuali o l'insieme delle varianti interpretative. Sicché le parti del testo in cui si localizzano i fenomeni di **instabilità** dinamica, le sue **zone critiche**, vanno rispettivamente ricercate in quelle forme strutturali dell'espressione e del contenuto in cui si determina la relazione tra il tutto e le sue parti. I luoghi in cui agiscono e trovano diretta applicazione le **regole** dell'articolazione strutturale del tutto possono essere considerati zone critiche del testo. Nel caso dell'espressione, la zona critica, sede dell'instabilità e dell'indeterminazione, va ricercata in quella porzione del linguaggio-oggetto che reca informazione autoriflessiva sulla sua stessa struttura. Nel caso del contenuto, la zona critica va ricercata nelle relazioni strutturali complesse che le regole operative istituiscono tra gli elementi costitutivi del modello.

Rispetto agli elementi della struttura che esse definiscono, le regole hanno una posizione ambivalente. Esse sono ciò che pone in relazione i singoli elementi con la totalità del sistema, ma possono essere espresse in modo diverso e il loro status è essenzialmente ambiguo. Nel caso dell'espressione, esse possono essere espresse o metalinguisticamente, o con asserzioni del secondo ordine del linguaggio-oggetto; nel caso del contenuto, o come specificazione delle componenti strutturali di ordine superiore e delle relazioni strutturali complesse che esse determinano tra gli elementi del modello, o come vincoli e restrizioni applicate alle operazioni che esse definiscono sui medesimi elementi. I diversi modi di espressione delle regole tra loro equivalenti ed alternativi determinano il luogo in cui assume forma esplicita l'**ambiguità** della struttura. Le relazioni strutturali compensative tra le parti e il tutto dei due sottosistemi, espressione e contenuto, che compongono la totalità del testo dipendono dalle regole che ne definiscono la struttura e dai loro vincoli operativi. È quindi l'espressione delle regole e dei loro vincoli operativi ciò che produce i fenomeni di instabilità strutturale e di ambivalenza dinamica del testo. I rapporti di compensazione tra la determinazione e l'indeterminazione della struttura dell'espressione e del contenuto del testo e le relazioni formali tra le loro rappresentazioni astratte, costituite rispettivamente dal sistema di *markup* applicato all'espressione e dal sistema di gestione della base di dati che ne descrive il contenuto, sono perciò definiti dalle formulazioni alternative delle rispettive regole di organizzazione strutturale.

La distinzione primaria tra l'espressione e il contenuto è già essa stessa il risultato di un'analisi dell'unità costitutiva della totalità del testo. L'instabilità strutturale della rappresentazione del testo è il risultato immediato di questa distinzione. Espressione e contenuto costituiscono due **sottosistemi** dell'intero sistema testuale e considerati come unità a sé stanti costituiscono due nuove totalità distinte e correlate. Analizzando l'espressione se ne fissa la struttura in relazione alla sua totalità integrale. Ma ad una struttura determinata dell'espressione e alla sua identità corrispondono diverse possibilità di articolare il contenuto. Fissando l'identità del testo nell'unità parziale costituita dalla sola espressione, si lascia indeterminata l'altra unità parziale costituita dal contenuto. Un fenomeno simmetrico e del tutto analogo si produce fissando l'identità del testo nell'altra unità parziale costituita dal contenuto. I fenomeni di instabilità e di indeterminazione della struttura testuale si producono assumendo come identità integrale del testo l'identità con se stessa di una delle sue sottounità parziali.

## 7. **Rappresentazione formale della struttura del testo**

Le forme di **rappresentazione formale** della struttura delle due sottounità parziali del testo possono essere costituite rispettivamente dal *markup* dell'espressione e dal modello del contenuto. La relazione strutturale tra espressione e contenuto può quindi essere rappresentata dalla relazione tra il sistema di *markup* applicato all'espressione e il sistema di gestione del *database* che ne modella il contenuto. L'indeterminazione strutturale dell'espressione si esprime nell'insieme delle regole del sistema di *markup*

e nei suoi gradi di libertà. L'indeterminazione strutturale del contenuto si esprime nell'insieme delle regole operative del sistema di gestione del *database* e nei suoi gradi di libertà, ossia nella potenza virtuale del formalismo applicato agli elementi strutturali del modello.

La rappresentazione delle relazioni strutturali tra gli elementi costitutivi del contenuto, ovvero l'espressione dei risultati delle operazioni applicate a tali elementi, può essere ottenuta applicando diverse **regole di linearizzazione**. Si ottengono così descrizioni, o rappresentazioni diverse delle relazioni interne alla totalità del contenuto. Allo stesso modo, ad una medesima rappresentazione lineare del contenuto del testo possono essere assegnate, attraverso il *markup*, **forme logiche** diverse. Si ottengono così modelli diversi delle relazioni interne alla totalità dell'espressione.

Ciò che esprime le relazioni interne alla totalità dell'espressione può a sua volta essere considerato come parte dell'espressione o come qualcosa di separato ed esterno. Allo stesso modo, le regole delle operazioni applicabili agli elementi costitutivi del contenuto possono essere definite e descritte all'esterno del modello, oppure possono essere rappresentate direttamente al suo interno attraverso la specificazione di tutti gli elementi del dominio e del codominio che esse mettono rispettivamente in corrispondenza tra loro. Questa **ambivalenza** delle forme di rappresentazione delle regole di articolazione strutturale delle sottounità costitutive del testo permette di convertire varianti interpretative in varianti testuali e viceversa.

La rappresentazione formale della struttura interna alle sottounità testuali permette anche di rendere esplicite le forme di compensazione tra determinazione e indeterminazione strutturale del testo e dà la possibilità di analizzare i processi di instabilità dinamica della struttura testuale. Tale rappresentazione esplicita si può ottenere proprio grazie alla natura specifica della **rappresentazione digitale**. Le proprietà semiotiche degli elementi significativi primari dell'espressione e del contenuto non costituiscono proprietà elementari non ulteriormente analizzabili, ma possono essere determinate in funzione dalle proprietà semiotiche di unità ancora più elementari delle strutture di dati che le rappresentano, quali sono in effetti i singoli caratteri codificati. E la rappresentazione digitale permette di applicare procedure rigorose e automatiche di elaborazione e di analisi alle unità minime costitutive delle strutture di dati più complesse che rappresentano le proprietà semiotiche delle unità significative elementari del testo. ( *SMcG DI*, 35) La rappresentazione digitale permette di considerare gli elementi dell'instabilità dinamica del testo non come un dato fisso e imm modificabile, ma come il risultato controllabile e analizzabile di processi più elementari di costituzione della struttura del testo e di produzione di fenomeni testuali complessi.

## 8. Formulazione della legge di compensazione

La rappresentazione formale della struttura dell'espressione e della struttura del contenuto del testo può essere costituita rispettivamente dal **markup** dell'espressione e dal **database** del contenuto. Il *markup* può essere considerato come la forma di esplicitazione lineare, in quanto legato all'espressione del testo, di complesse relazioni interne, rappresentate esplicitamente in forma non lineare da un *database* che ne descrive il contenuto. Come ci si può valere di queste rappresentazioni formali delle sottounità strutturali del testo per esprimerne la **legge di compensazione**?

La legge di compensazione è stata proposta da Jerome **McGann** nella forma seguente:

$$A = A \iff A \neq A \quad . \quad ( \text{McG WIT}, 81; \text{McG VIB} )$$

La legge esprime il principio paradossale che ogni testo, *A*, non è identico a se stesso. Mc Gann ricava questa legge dalla relazione tra identità del tutto e distinzione prodotta dalla sua partizione primaria, che George **Spencer Brown** ha espresso formalmente attraverso l'introduzione del concetto di *forma della distinzione* ( *SB LF*, 1 a ) e delle sue leggi. Applicata al testo, la specificazione della partizione primaria di Spencer Brown può essere costituita dalla distinzione delle sue sottounità primarie, l'espressione e il contenuto. L'*indicazione* dell'espressione o l'*indicazione* del contenuto ( *SB LF*, 1 b ) presuppongono infatti la loro distinzione, prodotta dalla partizione primaria operata sulla totalità del testo. L'*indicazione* dell'espressione pone l'espressione come sottounità del testo identica a se stessa e ne determina la struttura. La determinazione e l'identità dell'espressione con se stessa è espressa formalmente dalla legge di idempotenza dell'espressione rispetto alla sua rappresentazione - primo assioma di Spencer Brown, o *law of calling* ( *SB LF*, 1 c ), e *forma della condensazione*. ( *SB LF*, 5 a ) L'*indicazione* del contenuto pone il contenuto come sottounità del testo identica a se stessa e ne determina la struttura. Allo stesso modo, la determinazione e l'identità del contenuto con se stesso è espressa formalmente dalla legge di idempotenza del contenuto rispetto alla sua rappresentazione. Ma come può l'**identità** del testo con se stesso, formulata come idempotenza delle sue sottounità parziali rispetto alla loro rappresentazione, dipendere ed essere tolta ad un tempo dalla **distinzione** primaria che le definisce?

Si può mostrare che la legge di compensazione presuppone ed implica un **endomorfismo** ( *f* ) tra gli elementi strutturali costitutivi del testo

$$( A = A \iff A \neq A ) \iff A \rightarrow A \quad . \quad ( \text{B RDMT}, 156 )$$

Tale endomorfismo può mettere in corrispondenza elementi del contenuto con elementi dell'espressione o, inversamente, elementi dell'espressione con elementi del contenuto. La compensazione tra l'espressione e il contenuto è rappresentata dall'**inversione** tra il dominio e il codominio dell'endomorfismo. L'endomorfismo tra gli elementi strutturali del testo può essere espresso dal *markup*. Lo status logico del *markup* è ambivalente e l'**ambiguità del markup** costituisce l'espressione esplicita dell'inversione dell'endomorfismo. Ciascuna delle due sottounità parziali, espressione e contenuto, prodotte dalla partizione primaria del testo può essere di volta in volta assegnata come dominio o come codominio dell'endomorfismo. L'inversione si produce sostituendo forme di *markup* tra loro equivalenti e la compensazione prodotta dalla dinamica olistica interna al testo si manifesta esplicitamente nell'oscillazione tra l'equivalenza e la distinzione funzionale delle diverse forme di *markup*. Cerchiamo di descrivere in modo più preciso questo fenomeno di compensazione.

Il *markup* può essere considerato come rappresentazione **metalinguistica** della struttura dell'espressione o come sua espressione diretta e autoriflessiva costituita da forme di predicazione del secondo ordine formulate nel **linguaggio-oggetto** ( *B RDMT*, 152-53). In quest'ultimo caso il *markup* assegna alla struttura dell'espressione una forma logica del secondo ordine, mentre nel primo caso le assegna una forma logica del primo ordine. Il riferimento alla totalità integrale del testo rende equivalenti le due diverse formulazioni del *markup* e ne genera l'ambiguità diacritica. A sua volta, l'**equivalenza** delle due forme di *markup* rispetto alla totalità integrale del testo rende possibile il trasferimento dell'indeterminazione dall'una all'altra delle sue sottounità parziali. Infatti, la struttura del contenuto può essere riferita alla forma dell'espressione attraverso una forma di *markup* metalinguistico e *out of line*, ossia non vincolato alla posizione dei marcatori nella successione dei caratteri codificati ( *RTW MR*, 4). Questo permette di assegnare all'espressione anche strutture non lineari o strutture gerarchiche *enjambées*. Diverse **varianti interpretative** possono essere riferite in questo modo alla struttura del testo. È però possibile trasformare le varianti interpretative in **varianti testuali** sostituendo il *markup* metalinguistico con forme di *markup* equivalenti costituite da espressioni autoriflessive del secondo ordine appartenenti al linguaggio-oggetto. Allo stesso modo, sostituendo le espressioni autoriflessive del secondo ordine che rendono esplicita la struttura dell'espressione all'interno del linguaggio-oggetto con forme equivalenti di *markup* metalinguistico, è possibile

trasformare varianti testuali in varianti interpretative.

Riferite alla totalità integrale del testo le due diverse forme di *markup* sono di fatto equivalenti e rendono sostanzialmente ambigua l'espressione dell'articolazione strutturale del testo e delle sue relazioni interne. Riferite invece a ciascuna delle due sottounità parziali del testo le due diverse forme di *markup* restano **funzionalmente distinte** e producono la determinazione e indeterminazione reciproca dell'espressione e del contenuto. Infatti, se la **forma logica** assegnata dal *markup* all'espressione del testo è rappresentabile con una logica del **primo ordine**, essa esibisce la sola struttura dell'espressione in modo affatto indipendente dalla struttura del contenuto. Il sistema formale per l'elaborazione di tale forma può avere la proprietà della **completezza**, ma non può avere la proprietà della categoricità. Se è completo, è possibile trovare un modello che soddisfa tutte le sue asserzioni dimostrabili, ma non tutti i modelli che soddisfano le asserzioni dimostrabili del sistema sono isomorfi. Ci troviamo quindi di fronte ad un caso di compensazione. Determinando la forma dell'espressione solo in relazione a se stessa si lascia indeterminata la struttura del contenuto. Viceversa, se la forma logica assegnata dal *markup* all'espressione del testo può essere rappresentata solo con una logica del **secondo ordine**, il sistema formale per l'elaborazione di tale forma può essere **categorico**, ma non può essere completo. Ciò significa che la forma logica dell'espressione esprime in questo caso anche la struttura del contenuto. Infatti, tutti i modelli che soddisfano le asserzioni dimostrabili del sistema sono isomorfi, benché non sia per principio possibile stabilire che tutte le asserzioni dimostrabili nel sistema sono verificabili. Di nuovo, alla determinazione strutturale della forma del contenuto viene a corrispondere l'indeterminazione della sua espressione. La legge di compensazione tra determinazione e indeterminazione reciproca dell'espressione e del contenuto può essere quindi rappresentata formalmente attraverso la relazione tra le rispettive rappresentazioni, ossia attraverso la relazione tra la forma logica assegnata al testo dal *markup* e il modello assegnatogli dal *database*. A sua volta, la relazione tra la **forma logica dell'espressione** (*markup*) e il **modello del contenuto** (*database*) può essere considerata un'esemplificazione del secondo assioma di Spencer Brown, o *law of crossing* ( [SBLF](#), 2), e della *forma della cancellazione* ( [SBLF](#), 5 b ). Infatti, il riferimento dell'articolazione strutturale di una delle due sottounità testuali alla totalità strutturale dell'altra ne toglie l'identità con se stessa e ne produce l'indeterminazione. In sostanza, dunque, l'identità del testo con se stesso è posta dalla partizione primaria tra espressione e contenuto ed è tolta dal riferimento incrociato (*crossing*) di ciascuna sottounità all'altra, che annulla l'identità separata delle unità parziali determinate e reintegra la totalità indeterminata del testo. Il testo può essere considerato e descritto, in ultima analisi, solo come unità olistica.

Alla luce di queste considerazioni si comprende perché le forme di *markup* vincolato dipendano dalla struttura della sola espressione (nel caso di testi rappresentati in formato SGML, dalle *context-free grammars* definite dalle rispettive DTD). Il tentativo di progettare questa struttura sulla forma del contenuto infrange la legge di compensazione. Una *context-free grammar* propria della sola espressione viene assunta indebitamente come *constraint language* per il sistema di gestione della base di dati ( [SHCDH](#), 30 b ).

## 9. [Per un'analisi delle rappresentazioni strutturali del testo](#)

Si è visto che per la comprensione dei fenomeni della rappresentazione digitale del testo è essenziale la considerazione di unità costitutive di natura ancor più elementare degli elementi ultimi della forma logica assegnata all'espressione o delle entità strutturali astratte costitutive del modello del contenuto. Le **categorie espressive** e le **categorie ontologiche** che costituiscono rispettivamente la base dell'articolazione strutturale dell'espressione e del contenuto, sono a loro volta il risultato dell'aggregazione di **costituenti più elementari**. I valori delle variabili dei linguaggi di programmazione applicabili rispettivamente alla struttura dell'espressione e alla struttura del contenuto sono a loro volta il risultato della composizione di dati codificati di natura più elementare. Il *markup* può essere riferito non solo alle unità logiche significative dell'espressione del testo, ma ai singoli caratteri codificati che le costituiscono e lo stesso vale per la descrizione delle strutture di dati a cui è applicabile il formalismo del sistema di gestione del *database* che ne rappresenta il contenuto.

È a questo livello estremamente elementare della rappresentazione che debbono essere applicate le procedure per l'elaborazione automatica dell'informazione testuale ed è di conseguenza a questo livello che deve sapersi esprimere la progettualità dei sistemi per la rappresentazione, l'edizione digitale e l'analisi automatica del testo. La filologia del testo digitale passa anche per queste strade.

---

### Bibliografia

- [ **B FA** ] D. Buzzetti, *La faculté des arts dans les universités de l'Europe méridionale: Quelques problèmes de recherche*, in *L'enseignement des disciplines à la Faculté des arts* (Paris et Oxford, XIIIe-XVe siècles), Actes du colloque international édités par O. Weijers et L. Holtz, Turnhout, Brepols, 1997 (Studia Artistarum, 4), pp. 457-466. [ [riferimenti: 465](#) ]
- [ **B IP** ] D. Buzzetti, *Image Processing and the Study of Manuscript Textual Traditions*, in «Historical Methods», 28 (1995), n. 3, pp. 145-154. [ [riferimenti: 145-48](#), [145-54](#) ]
- [ **B RDMT** ] D. Buzzetti, *Rappresentazione digitale e modello del testo*, in *Il ruolo del modello nella scienza e nel sapere* (Roma, 27-28 ottobre 1998), Contributi del Centro Linceo Interdisciplinare «Beniamino Segre», n. 100, Roma, Accademia Nazionale dei Lincei, 1999, pp. 127-161. [ [riferimenti: 130](#), [146](#), [147-48](#), [151ss](#), [152-53](#), [156](#) ]
- [ **B TF** ] D. Buzzetti, *Il testo 'fluidò': Sull'uso dell'informatica nella critica e nell'analisi testuale*, in *Filosofia & informatica, Atti del primo incontro italiano sulle applicazioni informatiche e multimediali nelle discipline filosofiche* (Convegno Nazionale della Società Filosofica Italiana: Roma, 23-24 novembre 1995), a cura di Luciano Floridi, Torino, Paravia, 1996, pp. 85-93. [ [riferimenti: 87-89](#) ]
- [ **B VRJ** ] D. Buzzetti, *Digital Editions: Variant Readings and Interpretations*, in *ALLC-ACH'96: Conference Abstracts*, The 1996 ALLC and ACH Joint International Conference (Bergen, 25-29 June 1996), University of Bergen, 1996, pp. 254-256. ([Adobe Acrobat PDF format version](#)) [ [riferimenti: 255](#) ]
- [ **BPT LM** ] D. Buzzetti, P. Pari, A. Tabarroni, *Libri e maestri a Bologna nel XIV secolo: Un'edizione come database*, in «Schede umanistiche», N.S., 2 (1992), n. 2, pp. 163-169. [ [riferimenti: LM](#) ]
- [ **BR TFDE** ] D. Buzzetti and M. Rehbein, *Textual Fluidity and Digital Editions*, in *Text Variety in the Witnesses of Medieval Texts*, Proceedings of the International Workshop, Sofia, 21-23 September 1997, edited by Milena Dobrova, Sofia, Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences, 1998, pp. 14-39. [ [riferimenti: 14-39](#), [36-37](#) ]
- [ **BT ICT** ] D. Buzzetti e A. Tabarroni, *Informatica e critica del testo: il caso di una tradizione 'fluida'*, in «Schede umanistiche», N. S., 1 (1991), n. 2, pp. 185-193. [ [riferimenti: ICT](#) ]
- [ **GL CS** ] L. Goldschlager and A. Lister, *Computer science: A Modern Introduction*, London, Prentice Hall, 19882. [ [riferimenti: 248](#) ]
- [ **J DR** ] V. Joloboff, *Document Representations: Concepts and Models*, in J. André, R. Furuta, V. Quint (eds.), *Structured Documents*, Cambridge, Cambridge University Press, 1989, pp. 75-105. [ [riferimenti: 75-76](#) ]

- [**McG WIB**] J. McGann, *Visible and Invisible Books: Hermetic Images in N-Dimensional Space*. [riferimenti: [VIB](#)]
- [**McG WIT**] J. McGann, *What Is Text?*, in *ACH-ALLC'99 Conference Proceedings*, Charlottesville VA, University of Virginia, 1999, Draft, pp. 80-81. [riferimenti: [81](#)]
- [**R DDT**] Malte Rehbein, *Die dynamische digitale Textedition: Ein Modell*, in AA. VV., *Vom digitalen Archiv zur digitalen Edition*, Göttingen, Max-Planck-Institut für Geschichte und Stadtarchiv Duderstadt, 1998, pp. 5-22. [riferimenti: [DDT](#)]
- [**RTW MR**] 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 [1-20]. [riferimenti: [1-4 a](#), [1-4 b](#), [3-4](#), [4](#)]
- [**RTW DRDM**] 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), pp. [1-22]. [riferimenti: [3-6](#)]
- [**S ATL**] C. Segre, *Avviamento all'analisi del testo letterario*, Torino, Einaudi, 1985. [riferimenti: [44](#)]
- [**SB LF**] G. Spencer Brown, *Laws of Form*, London, Allen and Unwin, 1969. [riferimenti: [1 a](#), [1 b](#), [1 c](#), [2](#), [5 a](#), [5 b](#)]
- [**SH CDH**] C. M. Sperberg-McQueen and C. Huitfeldt, *Concurrent Document Hierarchies in MECS and SGML*, in «Literary and Linguistic Computing», 14 (1999), pp. 29-42. [riferimenti: [30 a](#), [30 b](#), [41](#)]
- [**SMcG DJ**] L. Samuels and J. McGann, *Deformance and Interpretation*, in «New Literary History», 30 (1999), 25-56. [riferimenti: [35](#)]
- [**T K**] M. Thaller, *'kleio': A Database System*, St. Katharinen, Max-Planck-Institut für Geschichte i.K.b. Scripta Mercaturae Verlag, 1993. [riferimenti: [K](#)]

|                                  |   |
|----------------------------------|---|
| <a href="#">Torna all'inizio</a> | <a href="#">Soluzioni informatiche e telematiche per la filologia</a> |
|                                  | Pavia, 30-31 marzo 2000   |
| FACOLTÀ DI LETTERE E FILOSOFIA   | PUBBLICAZIONI TELEMATICHE   |