

CAPITOLO I

L'Informatica Umanistica: le biblioteche digitali e la codifica dei testi

[...] da incontri tra scienziati e letterati, al fine di individuare gli automatismi con mezzi elettronici sono venute grandi spinte all'impiego competente di tali risorse, senza che venisse alterata in alcun modo la natura e la consolidata tradizione delle specifiche discipline che se ne avvalgono. E così partendo senza ambizioni da bisogni elementari, ma trovandosi al crocevia degli scambi tematici più diversi, l'Informatica Umanistica si trova oggi quasi evocata a ricomporre una relazione perduta tra i due mondi del sapere.¹

¹ A. Celentano, A. Cortesi, P. Mastrandrea, *Informatica Umanistica*, pag. 46, consultabile all'indirizzo internet: http://www.mondodigitale.net/Rivista/05_numero_uno/Celentano_def.p.44-55.pdf

Introduzione

L'avvento di Internet ha contribuito a ridefinire le pratiche della diffusione delle informazioni, assegnandone il compito soprattutto alla tecnologia. Il testo, tradizionalmente cartaceo, è stato aperto all'intera comunità degli utenti grazie alla Rete e, in particolare, al *World Wide Web*, strumenti fondamentali per offrire l'accesso ai dati. È sempre più frequente, infatti, il ricorso a metodologie e ad applicazioni computazionali nell'ambito degli studi umanistici: le prospettive di utilizzo dei dispositivi elettronici si stanno progressivamente evolvendo, permettendo la conservazione, il trattamento e la diffusione del patrimonio letterario.

L'Informatica Umanistica - uno dei settori che utilizzano la codifica e il trattamento delle informazioni -, riflette sui metodi conoscitivi e comunicativi legati al concetto di informazione ed individua, in questo modo, un legame tra le discipline umanistiche e le applicazioni informatiche. C'è la necessità, tuttavia, di individuare delle regole specifiche al fine di esaltare le caratteristiche di entrambe e renderle funzionali ai diversi ambiti di studio. Per questo motivo, scelte e applicazioni adeguate orientano una corretta codifica.

1.1 L'Informatica Umanistica

La rivoluzione riguardante l'uso delle nuove tecnologie ha coinvolto in parte anche le discipline umanistiche: si è parlato di una svolta epistemologica che ha accompagnato la nascita di nuove figure professionali e di nuovi modi di fare ricerca e di insegnare. Per questo motivo i termini “informatica” e “umanistica” compaiono frequentemente e in contesti diversi.

L'Informatica Umanistica è il punto di contatto tra “scienze umane” e “scienze esatte” e propone l'integrazione di questi due settori, apparentemente troppo lontani. Essa non è nuova in Italia: già nel 1949 Padre Roberto Busa² promosse l'uso del computer finalizzato allo studio dei testi.

All'inizio l'Informatica Umanistica è stata considerata come una branca dell'informatica applicata alle discipline umanistiche: si presupponeva che gli strumenti informatici fossero al servizio delle scienze umane. Questa considerazione pose due limiti alla disciplina: il primo riguarda la natura tecnologica - se fosse così valutata, sarebbe limitata al solo utilizzo di strumenti informatici, perdendo di vista le possibili applicazioni - ; il secondo riguarda le possibilità di integrazione - si perderebbero, infatti, le potenzialità delle discipline umanistiche nei confronti dell'informatica e si dimenticherebbero i contributi importanti del linguista Noam Chomski rispetto ai linguaggi formali e di Alonzo Church, il fautore della “teoria dei tipi”.

Riguardo l'Informatica Umanistica deve prevalere, dunque, una valutazione completa, poiché in essa ogni campo di indagine e oggetto di studio utilizza strumenti elettronici e metodologie informatiche. In questo modo, è la conoscenza approfondita dell'oggetto di studio (opere, dati storici, siti archeologici, etc.) a proporre nuovi paradigmi di elaborazione adeguati a rappresentare le informazioni.

Ciò che distingue una scienza umana (discipline umanistiche) da una esatta (informatica) è il dualismo che vi è nell'oggetto di studio: per la linguistica avremo i significati e i significanti; per i testi l'interpretazione e la critica letteraria; per la storiografia i fatti storici e

² Padre Roberto Busa (Vicenza, 28 novembre 1913), gesuita e linguista italiano, è uno dei pionieri dell'uso dell'informatica nella linguistica (oggi Linguistica Computazionale) ed è stato il realizzatore dell'*Index Thomisticus*, collezione lemmatizzata dell'opera omnia di Tommaso D'Aquino e dei testi a lui associati. Questa iniziativa nacque durante la stesura della sua Tesi di Laurea (1946), momento in cui ebbe l'idea di fare una verifica integrale del lessico del filosofo e teologo medievale servendosi di macchine adeguate. Nel 1949, a New York, entrò in contatto con Thomas Watson Sr., amministratore delegato della IBM e lo convinse a sostenere il progetto. Si dedicò, pertanto, allo sviluppo dell'*Index Thomisticus*: prima si servì di schede perforate e, in seguito, di nastri magnetici. Dopo anni di lavoro, completò l'edizione a stampa in 56 volumi (1980), riuscendo a fare assumere al suo *Index* la forma di ipertesto (1989), grazie al sostegno di Piero Slocovich. Nel 2005 il progetto, promosso dalla *Fundación Tomás de Aquino* e dalla CAEL (Associazione per la Computerizzazione delle Analisi Ermeneutiche Lessicologiche), fu pubblicato sul Web.

i documenti; per la storia dell'arte le immagini e le rappresentazioni. Questo confronto distingue tra elementi certi e quantificabili e interpretazioni che sfuggono alla logica e all'oggettività. È su questo aspetto che l'Informatica Umanistica propone l'integrazione di due mondi e non soltanto l'utilizzo di applicazioni tecnologiche avanzate per i settori delle scienze umane.

1.1.1 La Linguistica Computazionale

La Linguistica è stata forse una delle prime discipline che meglio ha compreso l'importanza dell'integrazione delle scienze umane con l'informatica: le ricadute si sono avute con la nascita della Linguistica Computazionale. Essa è riuscita a far crescere i propri progetti di ricerca non solo in ambito accademico: sono stati sviluppati, infatti, correttori ortografici e sintattici, sistemi di riconoscimento vocale, programmi di redazione di riassunti automatici.

La Linguistica ha utilizzato approcci tradizionali dell'informatica per analizzare il testo, come il trattamento delle stringhe di caratteri, all'interno delle quali è possibile individuare particolari *pattern* che permettono, ad esempio, la suddivisione in parole, il raggruppamento delle forme, l'ordinamento alfabetico.

In questo contesto, il compito della Linguistica Computazionale è stato quello di elaborare schemi capaci di individuare, in modo semi-automatico, le relazioni linguistiche presenti tra i vari elementi; di procedere alla lemmatizzazione delle forme; di individuare le strutture sintattiche e di associare strutture semantiche adeguate.

L'Informatica Umanistica deve sfruttare questo valore, ponendo le basi di nuovi standard, di nuove applicazioni e di nuove regole, anche nel settore dei testi letterari.

1.1.2 Gli studi letterari

Nell'ambito dell'Informatica Umanistica sono mutati in parte anche gli studi letterari, i quali hanno tratto un forte ausilio dalla costituzione dei *corpora* di testi antichi e moderni. Queste sono state le prime forme di ricerca in materia di digitalizzazione di documenti che, assieme all'utilizzo dell'*Information Retrieval*³, permettono ricerche e analisi testuali.

³ L'*Information Retrieval* (recupero d'informazioni) è l'insieme delle tecniche utilizzate per il recupero mirato delle informazioni (documenti, metadati, file, banche dati) in formato elettronico. L'IR, il cui termine è stato coniato da Calvin Mooers alla fine degli anni '40 del Novecento, è un settore interdisciplinare che coinvolge la psicologia

Questo settore ha consentito di riunire parte dei materiali delle biblioteche, di garantire l'accesso ai dati e di ricercare informazioni nei metadati e nei dati primari (testuali e non) su qualunque campo letterario.

Le iniziative in materia di digitalizzazione di testi hanno permesso anche di sviluppare software adeguati capaci di soddisfare le esigenze più diverse: sia i progetti no profit sia quelli commerciali (ad esempio, il *Thesaurus Linguae*) hanno compreso che per separare le banche dati dai sistemi di interrogazione era necessario pubblicare le specifiche della codifica dei testi. E' su questo ambito che si è concentrata la *Text Encoding Initiative* (TEI), iniziativa che riflette sulle diverse tipologie testuali (poesia, prosa, dramma, etc.) al fine di proporre schemi di marcatura XML del testo e delle informazioni ad esso correlate.

1.2 Le Biblioteche Digitali

Seppure per alcuni l'evoluzione della tecnologia digitale sia stata un fattore di discontinuità con il passato rispetto a determinate discipline, essa ha permesso la nascita e lo sviluppo delle biblioteche digitali.

Paul Levy⁴, osservando gli sviluppi digitali, intravedeva la possibilità che tutto il sapere umano potesse essere accessibile perché non conservato in un singolo luogo, ma in più parti del mondo. Ma anche Vannevar Bush con il suo Memex evidenziò la necessità di dare un nuovo ordine ad una biblioteca, in quanto la sua struttura si rilevava incapace a gestire ricerche adeguate come sostegno della conoscenza. Negli anni '60 Joseph Carl Licklider⁵, nel suo *Libraries of the Future* anticipò il cambiamento delle biblioteche attraverso l'uso del computer, parlando di *procognitive utility net* come servizio di supporto all'apprendimento.

La biblioteca è una realtà dinamica sia per la continua acquisizione di pubblicazioni e di materiali di consultazione di varia natura, sia per i servizi che essa offre all'utenza. Con le biblioteche digitali questo valore non si perde, in quanto la catalogazione dei documenti presenti in un biblioteca tradizionale corrisponde all'assegnazione, alla gestione e alla ricerca dei metadati, intesi come insieme dei dati e delle informazioni capaci di descrivere un testo in formato digitale. L'attribuzione dei metadati è una delle caratteristiche più importanti per costituire una collezione di documenti, mentre un corretto sistema di gestione rappresenta il supporto per fornire all'utente le informazioni riguardanti il materiale a cui è interessato. Pertanto, la corretta digitalizzazione di un testo, da inserire all'interno di una biblioteca digitale, deve prevedere l'inserimento delle caratteristiche bibliografiche (titolo, autore, luogo e data di edizione) dell'originale cartaceo; delle metodologie e degli strumenti utilizzati per rappresentare il documento in formato digitale; della certificazione dell'autenticità e della qualità del documento digitalizzato (il livello di correttezza della trascrizione e gli interventi di codifica strutturale e semantica adottati). Un

⁴ Paul Pierre Lévy (Parigi, 1886–Parigi, 1971), matematico e statistico francese, si concentrò sulla “teoria della probabilità”, diventando uno dei massimi esponenti del XX secolo di questo settore.

⁵ Joseph Carl Robnett Licklider (Stati Uniti, 1915–1990), informatico, è oggi considerato una delle figure più importanti nella storia dell'informatica e del calcolo generale. In particolare, egli si concentrò sulla tecnologia dell'informazione, anticipando lo sviluppo di Internet. Previde, infatti, la nascita di calcolatori in Rete dotati di interfacce-utente. Il suo articolo sulla *Simbiosi Uomo - Computer* preannunciava la computazione interattiva e fu alla base delle prime ricerche riguardo lo sviluppo di applicazioni. Inoltre, nel 1968, pubblicò un articolo dal titolo *Il calcolatore come dispositivo di comunicazione*, prevedendo l'utilizzo delle reti di calcolatori a sostegno delle comunità degli utenti e indipendenti dalla localizzazione geografica.

requisito importante è la capacità dei sistemi di gestione di interagire con i tradizionali sistemi catalografici e di essere inseribili all'interno di OPAC⁶.

Il ruolo dei metadati ha avuto un notevole sviluppo quando sono state elaborate una serie di proposte, delle quali una delle più importanti è il *Resource Description Framework* (RDF), un progetto tuttora in corso di elaborazione e promosso dal *World Wide Web Consortium*. Si tratta di un meta-linguaggio astratto che definisce marcatori capaci di descrivere ogni risorsa elettronica. Le specifiche di RDF si basano sulla sintassi XML ed è funzionale alla formalizzazione dei metadati in modo tale che siano direttamente trattabili da sistemi software senza alcun intervento umano.

Le biblioteche digitali, seppure con queste anticipazioni, hanno una storia recente e molto discussa, in quanto ancora oggi ci si chiede quale sia il loro ruolo e quali le loro caratteristiche fondamentali.

1.2.1 La definizione di biblioteca digitale

Un settore nascente e in forte sviluppo implica l'introduzione di concetti e strumenti nuovi: la terminologia consente, infatti, di evitare quelle ambiguità e imprecisioni che potrebbero derivare dall'utilizzo di termini preesistenti. Con questa logica si iniziò a definire "biblioteca digitale" quella che un tempo era conosciuta come "biblioteca elettronica".

La prima vera definizione si ebbe nel 1993, quando l'autrice Christine Borgman⁷ utilizzò il termine "biblioteca digitale" per definire la combinazione di servizi, di architettura di rete e di risorse informative (testi, dati numerici, immagini, documenti sonori e video). Una seconda definizione fu poi redatta da William Arms⁸, il quale si soffermò principalmente sulla necessità che essa assumeva nell'ottica di saper organizzare e gestire collezioni digitali.

⁶ OPAC, acronimo di *Online Public Access Catalogue*, è un sistema di accesso ai cataloghi delle biblioteche. Esso ha segnato il passaggio dalla catalogazione cartacea a quella elettronica: solo con l'informatizzazione delle biblioteche negli anni '80, infatti, è stato possibile creare dei sistemi di consultazione di cataloghi elettronici, accessibili agli utenti in modo autonomi. L'obiettivo di un OPAC è, dunque, quello di rendere accessibili i cataloghi delle biblioteche anche a coloro che non hanno le competenze necessarie. Tuttavia, negli ultimi anni, la maggior parte delle biblioteche hanno scelto di dotarsi di un OPAC: questo aspetto ha portato un notevole incremento dei sistemi di accesso ai cataloghi online, quantità che dal 2004 al 2009 è raddoppiata. Proprio del 2004 è lo studio di Aib e delle Università di Pavia e Ca' Foscari di Venezia sugli OPAC italiani.

⁷ Christine L. Borgman (Stati Uniti, 1951), docente presso l'Università della California, già professoressa presso l'Oxford Internet Institute dell'Università di Oxford, si occupa dello studio delle biblioteche digitali, della ricerca di informazioni, dell'interazione uomo-macchina.

⁸ William Arms è docente presso la Cornell University e si occupa di Informatica e Scienze dell'Informazione, dedicandosi soprattutto all'utilizzo dei computer nelle attività didattiche e alla nascita delle biblioteche digitali. L'interesse per questo materiale risale agli '70, quando l'Università di Dartmouth è stata la prima a creare un catalogo online. Dopo aver guidato lo sviluppo della *Electronic Library Mercurio*, si è dedicato allo sviluppo di programmi per gestire biblioteche digitali: è stato il principale ricercatore per la promozione del programma di CSTR DARPA (1992); il consulente per la NSF/DARPA/ NASA (1994); creatore, per conto della NSF, della *National Science Digital Library* (2000).

Un'altra definizione venne poi introdotta da Charles Oppenheim⁹ nel 1999, ripresa successivamente da Robert Smithson, il quale pone maggiore attenzione alle tecnologie utilizzate per realizzare queste collezioni: per entrambi, il ruolo della biblioteca digitale era quello di fornire un servizio informativo, nel quale tutte le risorse fossero disponibili in formato digitale, garantendo, per questo, possibilità di acquisizione, archiviazione, preservazione, recupero e accesso. A questa definizione si aggiunse anche quella di Chris Rusbridge¹⁰, il quale definì, nell'articolo *Toward the Hybrid Library*, la biblioteca digitale come "biblioteca ibrida", in quanto in essa vede un strumento di struttura fisica integrata all'interno di ambienti virtuali e di configurazioni fisiche tradizionali a cui si aggiungono archivi di collezioni numeriche. Gary Marchionini¹¹ ed Edward Alan Fox¹², in un saggio del 1999, focalizzarono l'attenzione sulla nascita della biblioteca digitale all'interno di uno spazio con quattro dimensioni: la comunità degli utenti (aspetti sociali, politici e legislativi del contesto socio-culturale); la tecnologia (progressi tecnologici; recupero delle informazione; multimedialità; interfacce, etc.); i servizi (*reference* con domande-risposte in tempo reale; aiuto online; corsi di educazione dell'utenza; personalizzazione dei servizi); i contenuti (tipologie di documenti).

Il termine "biblioteca digitale" fu utilizzato, in un primo tempo, soltanto da quei settori interessati a lavorare sull'accessibilità dei contenuti e sulla possibilità per l'utente di giudicare il valore e l'utilità del materiale raccolto. In questo contesto, il concetto di biblioteca digitale muta, perché *non è quello di una collezione digitale dotata di strumenti di gestione dell'informazione. È piuttosto uno spazio in cui mettere insieme collezione, servizi e persone a supporto dell'intero ciclo di vita della creazione, uso, preservazione di dati, informazione e conoscenza.*¹³ Seppure, in questa definizione emerga l'importanza della collezione, dell'accesso e dell'utente, la definizione di "biblioteca digitale" con i servizi ad essa correlati resta comunque suscettibile a modifiche. La *Digital Libraries Federation (DLF)* vede in esse il ruolo di *organizzazioni che forniscono le risorse, compreso il personale specializzato, per selezionare, organizzare, dare l'accesso intellettuale, interpretare, distribuire, preservare l'integrità e assicurare la persistenza nel tempo delle*

⁹Charles Oppenheim, professore emerito di Scienze dell'Informazione presso la Loughborough University di Leicestershire (Gran Bretagna), è noto per le sue ricerche in merito alle problematiche tipiche della comunicazione, alla gestione della conoscenza, all'informazione elettronica e all'editoria. In particolare, le sue aree di insegnamento e di ricerca si rivolgono alla scienza dell'informazione e alla gestione delle informazioni elettroniche.

¹⁰ Chris Rusbridge, direttore del Programma e-Lib (*Electronic Libraries Program*), svolge ricerche negli ambiti di supporto e di sviluppo in materia di conservazione digitale presso l'Università di Glasgow, progetti finanziati dalla JISC.

¹¹ Gary Marchionini, preside e docente della Scuola di Scienze dell'Informazione dell'Università della North Carolina, si è occupato delle interazioni uomo-informazione, di progettazione di interfacce e di biblioteche digitali. In passato è stato, invece, professore presso il Collegio dei servizi bibliotecari e dell'informazione dell'Università del Maryland e membro della *Human-Computer Interaction Laboratory*.

¹² Edward Alan Fox, professore presso il Dipartimento di Scienza dell'Informazione dell'Università della Virginia.

¹³ Luca Bardi, *Prende forma la digital library*, Biblioteche oggi, 1998, n. 10, pagg. 6-12.

*collezioni digitali così che queste possano essere accessibili prontamente ed economicamente per una comunità definita o per un insieme di comunità.*¹⁴

Come sottolinea Christine Borgman, il problema risiede nella mancata consapevolezza dell'importanza del ruolo delle biblioteche digitali all'interno di una società dell'informazione: l'approccio tecnologico e quello umanistico devono essere complementari. Il primo riguarda l'estensione e il miglioramento dell'uso tradizionale delle tecnologie e del recupero dell'informazione, includendo in esso anche l'organizzazione degli oggetti digitali e dei metadati; il secondo si riferisce alla strategia di consultazione e di ricerca di materiale di diverso tipo, ricerca che, precedentemente, trovava ostacoli nella possibilità di reperire, raccogliere, sintetizzare informazioni conservate in strutture diverse, spesso lontane tra loro.

Il termine "biblioteca digitale" coniato da Christine Borgman si affermò poi in Italia nel 1998, quando Michael Malinconico¹⁵ dimostrò come *le tecnologie digitali facilitano l'accesso alle raccolte bibliotecarie, trasferendo i contenuti delle fonti d'informazione o loro fedeli rappresentazioni attraverso lo spazio, dal luogo in cui sono conservate a quello in cui sono richieste. Le stesse tecnologie potrebbero essere utilizzate per trasportare la sostanza dei materiali nel tempo, contribuendo in tal modo alla loro conservazione.*

In seguito, i contributi pratici furono quelli di Luca Bardi¹⁶, di Claudio Leombroni¹⁷ e di Antonio Scolari¹⁸, il quale affrontò il tema di una federazione di biblioteche, aprendo, in questo modo, la strada a progetti come quello di AIBWEB¹⁹. Alberto Salarelli²⁰ e Anna

¹⁴ Carla Basili, Corrado Pettenati, *La biblioteca virtuale*, Milano: Bibliografica, 1994.

¹⁵ Sir Michael Malinconico, docente presso la *School of Library and Information Studies* dell'Università dell'Alabama, è oggi uno dei maggiori esperti mondiali nel campo delle tecnologie informatiche applicate alle biblioteche. Prima analista alla NASA, poi Direttore della Sezione Servizi Tecnologici della *New York Public Library* e Preside della *School of Computer, Information and Library Sciences del Pratt Institute*, è attualmente professore di *Information Studies* presso l'Università dell'Alabama.

¹⁶ Luca Bardi è Dirigente Area Sistema Bibliotecario di Ateneo presso il Politecnico di Milano.

¹⁷ Claudio Leombroni, membro della Commissione Nazionale Biblioteche e Servizi Nazionali, fa parte anche Gruppo di Lavoro sulle biblioteche digitali, di cui è stato Presidente dal 2005 al 2008. Attualmente si occupa del Servizio Reti Risorse Sistemi della Provincia di Ravenna.

¹⁸ Antonio Scolari si occupa del sistema bibliotecario d'Ateneo dell'Università di Pavia, costituito dalle biblioteche che gestiscono il patrimonio librario dell'Università di Pavia e dalle strutture centralizzate (Divisione Gestione Risorse documentarie; Divisione Biblioteca Digitale e Divisione Servizi Finanziari e Amministrativi) che ne regolamentano le attività e che gestiscono i servizi comuni.

¹⁹ Si tratta di un progetto promosso dall'AIB, l'Associazione Italiana delle Biblioteche nata nel 1930. AIBWEB è, pertanto, come trasposizione delle attività tradizionali sulla Rete. Il suo scopo è quello di fornire notizie ed informazioni in merito alle biblioteche presenti su Internet: all'interno del sito, visualizzabile al seguente indirizzo <http://www.aib.it/> è possibile trovare il MetaOpacAzalai (MAI), il Segnaweb e altri importanti materiali, come quelli della Documentazione di Fonte Pubblica in Rete (DFP), destinati sia ai bibliotecari che agli utenti meno esperti.

²⁰ Alberto Salarelli, docente di "Fondamenti di scienza dell'informazione" presso il Corso di laurea in Beni Artistici, Teatrali, Cinematografici e dei Nuovi Media dell'Università di Parma, svolge attività di ricerca negli ambiti della documentazione in formato digitale e, in particolare, nel rapporto che esiste tra la biblioteconomia e le nuove tecnologie.

Maria Tammaro²¹, nel volume *Biblioteca Digitale*, ampliano ulteriormente il significato del termine, che indica *uno spazio informativo in cui le collezioni digitali, i servizi di accesso e le persone interagiscono a supporto del ciclo di creazione, preservazione, uso del documento digitale*. E ancora, Fabio Ciotti²² e Gino Roncaglia²³ analizzano l'organizzazione dei documenti e dei metadati, arrivando a definire la biblioteca digitale come *collezione di documenti digitali strutturati (sia prodotti mediante digitalizzazione di originali materiali, sia realizzati ex-novo), dotata di un'organizzazione complessiva coerente di natura semantica e tematica, che si manifesta mediante un insieme di relazioni interdocumentali e intradocumentali e mediante un adeguato apparato metainformativo*. In questo senso possiamo distinguere una biblioteca digitale da un insieme non organizzato di informazioni assolutamente eterogenee come *World Wide Web*, ma anche da molti archivi testuali che attualmente sono disponibili su Internet e che si presentano come "depositi testuali" piuttosto che come vere e proprie biblioteche. La definizione più recente è quella di Gabriele Mazzitelli²⁴, il quale afferma che la biblioteca digitale è *l'insieme di una o più collezioni di oggetti digitali, della descrizione di questi oggetti (che si effettua utilizzando i cosiddetti metadati), messi a disposizione di tutti gli utenti interessati grazie a un'interazione di tipo elettronico che può comprendere diversi servizi quali la catalogazione, l'indicizzazione, il servizio di recupero dei documenti e di fornitura di informazioni a distanza (in cui tutte le richieste degli utenti e le relative risposte si effettuano, pertanto, usando la rete)*. La biblioteca digitale si presenta come un sistema complesso organizzato in cui si mettono a disposizione dell'utenza, in maniera strutturata, dei contenuti che, oltre ad essere derivati da una raccolta cartacea, possono già essere disponibili in rete o essere il risultato di un'attività intellettuale originale posta in essere dalla biblioteca o dai membri della comunità che fa parte della sua utenza istituzionale (ad esempio in ambito universitario, dispense di corsi o relazioni tenute a convegni).

È ovvio che ogni biblioteca dovrà avere criteri precisi in merito alla selezione della collezione; alla strategia con cui essa viene pubblicata; alla ricerca. Le biblioteche digitali dovrebbero essere, dunque, combinazione delle caratteristiche tipiche (selezione e crescita del materiale) di una biblioteca tradizionale e dei servizi derivanti da una profonda comprensione dell'utenza, che può essere definita sulla base della sua tipologia (generica o

²¹ Anna Maria Tammaro, ricercatrice presso la sezione Beni Librari del Dipartimento dei Beni Culturali e dello Spettacolo dell'Università degli Studi di Parma, si occupa di biblioteche digitali e di editoria elettronica.

²² Fabio Ciotti, docente di "Editoria multimediale e Informatica per le scienze umane" presso le Università di Viterbo, IULM di Milano e di Parma, è stato tra i fondatori del Centro di Ricerche Informatica e Letteratura (CRILet), sezione del Dipartimento di Studi Filologici Linguistici e Letterari dell'Università di Roma "La Sapienza", dove tuttora conduce attività di ricerca teorica e pratica nell'ambito dell'Informatica Umanistica.

²³ Gino Roncaglia, ricercatore di filosofia presso l'Università della Tuscia di Viterbo, coordina, a partire dal 1996, un seminario di teoria e pratica degli ipertesti e si occupa, inoltre, di storia e di logica.

²⁴ Gabriele Mazzitelli, diplomato in Archivistica minore presso la Scuola di Paleografia e Diplomatica dell'Archivio Segreto Vaticano e in Biblioteconomia presso la Scuola di Biblioteconomia della Biblioteca Apostolica Vaticana, nel 2005 ha conseguito il master universitario di I livello in "esperto in indicizzazione di documenti cartacei, multimediali e elettronici in ambiente digitale" presso l'Università degli Studi di Roma "Tor Vergata". Dal 1988 è responsabile della Biblioteca dell'Area Biomedica dell'Università di "Tor Vergata".

specialistica) o dei suoi materiali. Inoltre, mentre in passato la presenza fisica era indispensabile per l'utente, adesso l'accesso dipende dal contesto sociale in cui essa è inserita.

1.2.2 La Piccola Biblioteca Digitale Romanza

La Piccola Biblioteca Digitale Romanza (<http://piccolabdr.humnet.unipi.it/>) nasce con lo scopo di raccogliere materiale inerente alla Lingua e alla Letteratura Romanza, consultabile in formato elettronico e ricercabile per mezzo di software adeguati agli studi umanistici. Con essa si è voluto tentare di superare i limiti del libro elettronico e di fornire all'utente uno strumento di didattica e di ricerca. Dal momento che si tratta di una biblioteca digitale, essa rispetta gli standard di qualità certificati per i documenti archiviati, le strategie di preservazione delle risorse digitali, le modalità di distribuzione e gli strumenti di accesso per gli utenti.



(Home-Page della Piccola Biblioteca Digitale Romanza)

La Piccola Biblioteca Digitale Romanza è strutturata come un ipertesto diviso in quattro macro-sezioni: per facilitare la consultazione, esse sono suddivise in ulteriori ripartizioni, “strumenti”, “poesia”, “testo” e “narrativa”.



Sezione “Strumenti” della Piccola Biblioteca Digitale Romanza



Lista “Strumenti” della Piccola Biblioteca Digitale Romanza



Sezione “Poesia” della Piccola Biblioteca Digitale Romanza

Delle varie sezioni previste dalla Piccola Biblioteca Digitale Romanza la più completa è quella della poesia, dove è possibile consultare la raccolta scritta nel 1556 *Les Petites*

Inventions (*Le Piccole invenzioni*) di Remy Belleau²⁵. La prima riguarda la traduzione delle *Odes d'Anacreon Teien poète grec* dello pseudo-Anacreonte e le seconde sono inni originali dedicati agli elementi della natura e della società.

The screenshot shows the 'Poesia' section of the Piccola Biblioteca Digitale Romana. The header includes the library's name and a navigation menu with options like 'il progetto', 'come funziona', and 'contatti'. The main content area is titled 'Remy Belleau, Les Petites Inventions' and contains a detailed description of the text's history and a list of editions. The text is in French and discusses the history of the work, its publication, and the editor's efforts to reconstruct the original text. The list of editions includes: 1556 - *Les Odes d'Anacréon... ensemble quelques petites hymnes de son invention*, Paris, Wechel, 1556; 1573a - *Odes d'Anacréon... Ensemble quelques petites hymnes de son invention...*, Paris, Granjon, 1573; 1573b - *Odes d'Anacréon... Ensemble quelques petites hymnes de son invention... troisième édition...*, Paris, Granjon, 1573.

I contenuti della sezione “Poesia” della Piccola Biblioteca Digitale Romana

L'utilizzo della codifica XML permette di revisionare un testo anche dopo la prima codifica, in modo tale da arricchire e correggere l'interpretazione dei testi e di realizzare un'edizione critica in continua evoluzione e in condivisione tra gli utenti.

Inoltre, grazie alla *Tei-Lite*, al *software* di ricerca *full-text Philologic*, strumento realizzato dal progetto ARTLF dell'Università di Chicago, oltre a consultare documenti in essa contenuti, si possono salvare in formato “pdf” o condurre su questi ricerche complesse (parole, frasi, contenuti, figure retoriche).

²⁵ Remi Bellau è stato un poeta francese (Nogent-le-Rotrou 1528 - Parigi 1577). Dedicatosi agli studi umanistici, studiò al Collegio di Coqueret e riconosciuto da Ronsard fin dal 1554 come uno dei sette poeti della Pléiade, si arruolò in seguito nella cavalleria del marchese d'Elbeuf, fratello del duca di Guisa, scendendo con lui in Italia nel 1556. Al ritorno in Francia, nel 1557 continuò ad averlo come protettore, ma fu anche appoggiato dai sovrani Enrico II, Carlo IX e da Enrico III. Esordì nel 1557 con l'opera *Les Petites Inventions* (*Le piccole invenzioni*), ma si fece conoscere nel 1565 con *La Bergerie* (*La pastorale*), imitazione dell'*Arcadia* del Sannazaro, composta di prosa e di poesia. Poco prima di morire, nel 1576, pubblicò *Amours et nouveaux échanges des pierres précieuses* (*Amori e nuove metamorfosi delle pietre preziose*), considerato il suo capolavoro. Il tema, di antica tradizione e già trattato da autori medievali, è svolto con immaginazione epica e arte sottile, in quanto nell'opera è presente un evidente sentimento panteistico. Belleau è anche autore di un romanzo, *La reconneue* (1556) e delle traduzioni del *Cantico dei Cantici* e delle *Anacreontee*.



Sezione “Come Funziona” della Piccola Biblioteca Digitale Romanza



Sezione “Ricerca” della Piccola Biblioteca Digitale Romanza

La Piccola Biblioteca Digitale Romanza è attualmente priva di materiali nella sezione “teatro” e “prosa”. Il mio lavoro di Tesi è inserito all’interno di questo progetto. *Le Théâtre Italien ou Recueil de toutes le scenes françoises...* è una raccolta della *Comédie Italienne* in ambientazioni francesi del Seicento. L’autore è Evaristo Gherardi, figlio del comico Giovanni, conosciuto anche come interprete della maschera di Arlecchino. Gran parte delle opere della raccolta è in francese, ma vi sono passi, seppure poco frequenti, in italiano e in latino. I testi sono caratterizzati dal fatto di essere corredati da incisioni e da accompagnamenti vocali in entrambe le lingue.

1.3 La Codifica dei Testi

È a Fabio Ciotti che si deve forse una delle definizioni più complete di codifica informatica dei testi, intesa come *la rappresentazione formale di un testo ad un qualche livello descrittivo, su di un supporto digitale, in un formato utilizzabile da un elaboratore elettronico (Machine Readable Form) mediante un opportuno linguaggio informatico.*²⁶ Questa affermazione implica due livelli di analisi: da una parte vi è, infatti, il supporto materiale del testo; dall'altra, si trova l'aspetto astratto, basato essenzialmente sulla sua comprensione da parte dell'uomo e del computer.

Dal momento che la codifica elettronica garantisce il passaggio dei testi dal mondo materiale (atomi) a quello digitale (bit), con questo termine si deve far riferimento anche al concetto di codice, lo strumento per mezzo del quale è possibile rappresentare l'informazione trasmessa da un emittente a un ricevente e veicolata da un canale di comunicazione.²⁷ Affinché la trasmissione dell'informazione sia compresa, è indispensabile che l'emittente e il ricevente condividano lo stesso codice. Dunque questo comporta che durante la comunicazione si abbia *un processo di rappresentazione che implica una serie di operazioni di selezione e di classificazione degli elementi rilevanti in funzione di un determinato punto di vista.*

Anche Giuseppe Gigliozzi sottolinea che *codificare un testo significa ... esplorare il codice di partenza. Studiarne regolarità e funzioni. Confrontarle con le costrizioni proprie del codice prescelto e riproporre le relazioni (che sono già in un codice) del testo originale nel nuovo codice che, nel nostro caso, dovrà essere utilizzabile per un'elaborazione elettronica,...* Lavoreremo, quindi con due codici: il codice del documento da cui muoviamo e il codice che renderà questo documento disponibile per un'elaborazione elettronica.²⁸ Gigliozzi afferma inoltre che *col processo di codifica si realizza anche un strumento dell'oggetto testo, strumento che può essere qualcosa di più piccolo del testo*, in quanto vengono esplicitate informazioni maggiori rispetto all'originale. È opportuno che vengano rispettate

²⁶ Fabio Ciotti, *Cosa è la codifica informatica dei testi. Atti del Convegno Umanesimo & Informatica* (Trento 24-25 maggio 1996), a cura di Daniela Gruber e Patrick Pauletto, consultabile all'indirizzo Internet: [http://circe.lett.unitn.it/circe/\[...\]/ciotti.pdf](http://circe.lett.unitn.it/circe/[...]/ciotti.pdf)

²⁷ Si fa riferimento alla "teoria della comunicazione verbale" di Roman Jakobson. Egli schematizzò sei aspetti fondamentali della comunicazione, riconducibili anche ad altre forme di comunicazione, come quelle che utilizzano un linguaggio non verbale (suoni e gesti). In particolare, egli individuò un mittente (colui che invia il messaggio); un destinatario (il contesto, inteso come l'insieme della situazione generale e delle situazioni nelle quali è inserito il messaggio); un codice (lo strumento per mezzo del quale il messaggio risulta comprensibile sia al mittente che al destinatario); un contatto (il canale fisico e la connessione psicologica fra il mittente e il destinatario). Secondo Jakobson a questi sei fattori corrispondevano altrettanti funzioni: la funzione referenziale (contesto); la funzione emotiva (mittente); la funzione conativa (destinatario); la funzione fática (contatto); la funzione poetica (messaggio); la funzione meta-linguistica (codice).

²⁸ G. Gigliozzi, *Il testo e il computer. Manuale di informatica per studi letterari*, Milano, Bruno Mondadori, 1997, p. 56

*le leggi dell'isomorfismo. Dove per isomorfismo potremmo intendere una trasformazione che conservi l'informazione.*²⁹

La codifica, tuttavia, non può prescindere dall'attività interpretativa, la quale implica scelte e valutazioni da parte del codificatore, che è sia un trascrittore del testo, sia un "traduttore", poiché traspone il testo in una forma comprensibile al computer.

1.3.1 Il testo e il documento

*Quando parliamo di elementi del testo, possiamo alludere tanto a elementi del significante, quanto ad elementi del significato. Va ribadito che i secondi sono impliciti, e come sviluppati dai primi. Qualunque analisi del testo deve dunque avere il punto di partenza nel fatto che il testo è costituito da una successione di significanti grafici, o monemi, formati da gruppi di grafemi. È questa successione, immutabile se si prescinde dai guasti della tradizione, che sviluppa poi, nell'atto della lettura, i significati.*³⁰

Il significato del termine "testo", dal latino "*textus*" (tessuto o trama), è ambiguo, in quanto può riferirsi sia alla sequenza dei caratteri contenuti in esso sia al loro significato. I testi sono successioni di monemi rappresentabili graficamente tramite grafemi: il testo nasce proprio dalla lettura e dalla percezione dei monemi.

Questa premessa è indispensabile per analizzare i segni verbali che costituiscono il testo e che sono legati ad un singolo oggetto (carta e inchiostro). Quando si traspone un testo in MRF (*Machine Readable Form*), come vedremo successivamente, si attua un'operazione di memorizzazione del testo per come esso appare in quel contesto.

Un documento, invece, è il supporto materiale sul quale è conservato un testo (ad esempio, un libro), il quale contribuisce ad orientarne la lettura, focalizzando l'attenzione del lettore e scandendo le divisioni semantiche del testo. Il semiologo Gérard Genette ha dedicato un saggio a tutte quelle pratiche discorsive, iconiche e materiali che accompagnano il testo sia spazialmente che cronologicamente nel contatto con il fruitore, definendole con il termine di *paratesto*.³¹ Gli elementi a cui allude Genette possono essere sia di natura testuale (epigrafi, commenti, etc.) sia di natura visuale (composizione tipografica, impaginazione, etc.), la cui funzione è documentata dall'evoluzione della stampa. Esiste, tuttavia, una tradizione secolare, nata già prima della nascita del libro stampato: i *carmina intexta* medievali e i *carmina figurata* barocchi sono stati esempi delle possibilità estetiche che derivano da una semantizzazione degli elementi non - verbali del testo scritto.

²⁹ G. Gigliozzi, p.59.

³⁰ C. Segre, *Avviamento all'analisi del testo letterario*, p. 45.

³¹ G. Genette, *Soglie. I dintorni del testo*, trad. it., Torino, Einaudi, 1989, p. 6.

Il rapporto tra testo e paratesto è, pertanto, più complesso di quanto possa sembrare, in quanto ogni scelta fatta nell'edizione di riferimento risponde a specifiche pratiche, che possono essere veicolo semantico, sociale e storico. Rientrano in questa categoria elementi come i capitoli, i paragrafi, i capoversi, informazioni che, seppure percepite come caratteristiche del testo, sono trasmesse da elementi peritestuali (impaginazione, caratteri, spaziature, etc.). Elementi, questi, alla base delle relazioni strutturali e semantiche di un testo e da considerare ogni qual volta si trasferisce in MRF.

1.3.2 La *Machine Readable Form* (MRF)

Dal momento che un documento digitale è un testo memorizzato su un supporto elettronico, esso, per poter essere significativo da un punto di vista semantico, deve essere codificato dall'emittente in modo tale che il computer lo comprenda senza ambiguità. Con ciò si evidenziano gli elementi comuni ai processi informatici e a quelli comunicativi: per questo motivo, la memorizzazione di un testo non è solo trasposizione meccanica, ma un processo di trasmissione di informazioni ad un calcolatore. *Se si vuole trasferire un testo (a stampa o manoscritto) sull'elaboratore per sottoporlo a trattamento automatico (o trasmissione) si deve cercare di ridurre al minimo la perdita di informazione. La trasmissione deve rappresentare, infatti, tutta l'informazione contenuta nel testo. E' da qui che nasce il problema della codifica*³², alla quale il computer pone dei vincoli: ad esempio, la rappresentazione binaria delle informazioni.

La codifica risulta, quindi, necessaria per individuare gli elementi significativi del testo originale, sempre, però, tenendo presente che essa deve adeguarsi alla natura fisica del canale e del destinatario per rappresentare al meglio tutte le informazioni rilevanti, anche quelle che ad una prima lettura non risultano importanti. Pertanto, un testo trasposto in MRF deve tenere conto degli standard: se per un documento cartaceo il suo utilizzo è possibile in ogni contesto, questo non è valido per i documenti digitali, i quali devono, invece, avvalersi di strumenti hardware standardizzati e portabili³³ che ne consentano l'utilizzo.

³² D. Buzzetti, A. Tabarroni, *Informatica e critica del testo, Schede umanistiche*, Nuova Serie, p. 186.

³³ L'*International Organization for Standardization* è la più importante organizzazione a livello mondiale per la definizione di norme tecniche. Fondata il 23 febbraio 1947, ha sede a Ginevra. Ne sono membri gli organismi nazionali di standardizzazione di 157 Paesi del mondo. In Italia, le norme ISO vengono recepite e diffuse dall'UNI, rappresentante dell'Italia all'attività normativa dell'ISO. L'ISO coopera strettamente con l'IEC, responsabile per la standardizzazione degli equipaggiamenti elettrici.

1.3.2 La codifica dei caratteri

Il livello più basso della codifica informatica dei testi è la codifica dei caratteri, in quanto i computer gestiscono sequenze di numeri in formato binario (0 e 1). Affinché l'elaboratore decodifichi lettere o altri simboli sono necessarie delle tabelle per la conversione dei numeri in caratteri (lettere, numeri, simboli), tali da permettere un'assegnazione univoca.

Fino alla fine degli anni '50 la comunicazione tra operatore umano ed elaboratore avveniva soltanto con l'immissione di dati e con istruzioni in codice binario, ma, dal momento che i computer non sono solo calcolatori numerici, ma macchine universali in grado di manipolare simboli secondo specifiche regole, è sorta l'esigenza di rappresentare in maniera più efficiente nella memoria degli elaboratori segni come quelli delle notazioni alfabetiche, rappresentabili sempre per mezzo di 0 e di 1.

È essenziale che la codifica non trascuri quella serie di concetti che determinano una corrispondenza biunivoca nella rappresentazione dei caratteri: il *character repertoire*, insieme dei simboli che si vogliono rappresentare, e il *code set*, codici numerici che fanno riferimento ai simboli del *character repertoire*, sono, a tale proposito, indispensabili. Il risultato che si ottiene dall'unione di questi due insiemi viene detto *coded character set* o *charset*: esse sono tabelle contenenti i simboli con i rispettivi codici numerici, definite anche *code table*. Le *charset* contengono sia caratteri grafici sia istruzioni o codici di controllo che permettono la visualizzazione dei dati o altre operazioni su di essi.

In passato poteva accadere che questa corrispondenza biunivoca venisse meno nel momento in cui sistemi diversi utilizzavano formati numerici differenti per memorizzare il testo. La soluzione si ebbe nel 1963 quando un gruppo di studiosi sviluppò un sistema di codifica standard per l'interoperabilità e l'interscambio di informazioni tra computer. Il progetto fu approvato nel 1968 dall'Ente Nazionale Americano per la Standardizzazione (ANSI), il quale registrò la tavola ASCII (*American standard code for information interchange*) come standard nazionale. Sebbene questo standard sia ancora l'unico set supportato da ogni piattaforma hardware e software, presenta, tuttavia, il forte limite di non essere in grado di rappresentare tutte le lingue con le rispettive caratteristiche: comprende, infatti, 128 simboli, che vanno da 0 a 127, rappresentabili per mezzo di 7 *bit*.

Dec	Hx	Oct	Char	Dec	Hx	Oct	Htrl	Chr	Dec	Hx	Oct	Htrl	Chr	Dec	Hx	Oct	Htrl	Chr
0	0	000	NUL	(null)	32	20	040	␣	64	40	100	␣	␣	96	60	140	␣	␣
1	1	001	SOH	(start of heading)	33	21	041	!	65	41	101	␣	␣	97	61	141	␣	␣
2	2	002	STX	(start of text)	34	22	042	"	66	42	102	␣	␣	98	62	142	␣	␣
3	3	003	ETX	(end of text)	35	23	043	#	67	43	103	␣	␣	99	63	143	␣	␣
4	4	004	EOT	(end of transmission)	36	24	044	\$	68	44	104	␣	␣	100	64	144	␣	␣
5	5	005	ENQ	(enquiry)	37	25	045	%	69	45	105	␣	␣	101	65	145	␣	␣
6	6	006	ACK	(acknowledge)	38	26	046	&	70	46	106	␣	␣	102	66	146	␣	␣
7	7	007	BEL	(bell)	39	27	047	'	71	47	107	␣	␣	103	67	147	␣	␣
8	8	010	BS	(backspace)	40	28	050	(72	48	110	␣	␣	104	68	150	␣	␣
9	9	011	TAB	(horizontal tab)	41	29	051)	73	49	111	␣	␣	105	69	151	␣	␣
10	A	012	LF	(NL line feed, new line)	42	2A	052	*	74	4A	112	␣	␣	106	6A	152	␣	␣
11	B	013	VT	(vertical tab)	43	2B	053	+	75	4B	113	␣	␣	107	6B	153	␣	␣
12	C	014	FF	(NP form feed, new page)	44	2C	054	,	76	4C	114	␣	␣	108	6C	154	␣	␣
13	D	015	CR	(carriage return)	45	2D	055	-	77	4D	115	␣	␣	109	6D	155	␣	␣
14	E	016	SO	(shift out)	46	2E	056	.	78	4E	116	␣	␣	110	6E	156	␣	␣
15	F	017	SI	(shift in)	47	2F	057	/	79	4F	117	␣	␣	111	6F	157	␣	␣
16	10	020	DLE	(data link escape)	48	30	060	␣	80	50	120	␣	␣	112	70	160	␣	␣
17	11	021	DC1	(device control 1)	49	31	061	1	81	51	121	␣	␣	113	71	161	␣	␣
18	12	022	DC2	(device control 2)	50	32	062	2	82	52	122	␣	␣	114	72	162	␣	␣
19	13	023	DC3	(device control 3)	51	33	063	3	83	53	123	␣	␣	115	73	163	␣	␣
20	14	024	DC4	(device control 4)	52	34	064	4	84	54	124	␣	␣	116	74	164	␣	␣
21	15	025	NAK	(negative acknowledge)	53	35	065	5	85	55	125	␣	␣	117	75	165	␣	␣
22	16	026	STN	(synchronous idle)	54	36	066	6	86	56	126	␣	␣	118	76	166	␣	␣
23	17	027	ETE	(end of trans. block)	55	37	067	7	87	57	127	␣	␣	119	77	167	␣	␣
24	18	030	CAN	(cancel)	56	38	070	8	88	58	130	␣	␣	120	78	170	␣	␣
25	19	031	EM	(end of medium)	57	39	071	9	89	59	131	␣	␣	121	79	171	␣	␣
26	1A	032	SUB	(substitute)	58	3A	072	:	90	5A	132	␣	␣	122	7A	172	␣	␣
27	1B	033	ESC	(escape)	59	3B	073	;	91	5B	133	␣	␣	123	7B	173	␣	␣
28	1C	034	FS	(file separator)	60	3C	074	<	92	5C	134	␣	␣	124	7C	174	␣	␣
29	1D	035	GS	(group separator)	61	3D	075	=	93	5D	135	␣	␣	125	7D	175	␣	␣
30	1E	036	RS	(record separator)	62	3E	076	>	94	5E	136	␣	␣	126	7E	176	␣	␣
31	1F	037	US	(unit separator)	63	3F	077	?	95	5F	137	␣	␣	127	7F	177	␣	␣

Tavola dei caratteri ASCII

Analizzando la tavola, si può notare, infatti, come siano assenti gli accenti e le altre lettere fondamentali per la corretta riproduzione di una lingua che non sia l'inglese. Ci fu, pertanto, la necessità, soprattutto per i paesi non anglofoni, di porre rimedio all'inadeguatezza del *charset* ASCII. La prima soluzione fu quella di implementare nuovi *set* di caratteri a 8 bit e a 16 bit. Nacque allora l'ASCII esteso, conosciuto anche con il nome di ISO 8859-1 o ISO Latin – 1, il quale contiene i caratteri principali delle lingue occidentali e anglosassoni derivanti dall'alfabeto latino. La dicitura “-1” indica la versione (in questo caso, la prima tavola di una serie): dopo questa, infatti, vennero sviluppate anche la ISO 8859 – 2 (usata per le scritture latine dell'Europa centrale e dell'est) e la ISO 8859-15 (versione più recente), alla quale è stato aggiunto il simbolo dell'euro.

1.4 I linguaggi di codifica

La codifica dei caratteri, come si è spiegato precedentemente, è il livello più basso della codifica informatica di un testo.

Per codificare un testo, tuttavia, devono essere esplicitati tutti quegli elementi che permettono una corretta conservazione e riproduzione dell'originale cartaceo su un supporto digitale. Digitalizzare un testo, pertanto, implica un processo di analisi che parte dal singolo carattere alfabetico e che può arrivare a rappresentare le forme e le strutture più interne e complesse, come le figure retoriche. Si tratta di un processo di traduzione realizzato con linguaggi diversi. Definiti *markup language* (linguaggi di marcatura), essi riproducono gli elementi astratti (segmentazione logica e partizioni interne) e le informazioni strutturali, semantiche e tipografiche del testo.

I linguaggi di marcatura si dividono in due categorie: gli *specific markup language* (linguaggi procedurali) e i *generic markup language* (linguaggi dichiarativi).

1.4.1 I Linguaggi Procedurali

I linguaggi procedurali forniscono istruzioni operative allo stesso modo dei programmi informatici sulle porzioni di testo a cui fanno riferimento e descrivono il documento, permettendo al computer di rappresentarlo anche nel suo aspetto grafico.

Appartiene a questa categoria il *Rich Text Format* (RTF)³⁴. Realizzato dalla *Microsoft*, ebbe lo scopo principale di facilitare lo scambio di documenti in diverse applicazioni mediante codifica di tipo tagged. I linguaggi *Script*, *TROFF* e *TEX* sono efficaci nel definire la grafica finale di un testo, ma incapaci nell'analizzare il messaggio contenuto in esso.

1.4.2 I Linguaggi Dichiarativi

A differenza dei linguaggi procedurali, i linguaggi dichiarativi trasmettono all'elaboratore il contenuto del testo da codificare, identificando in esso la funzione strutturale e semantica di ogni blocco testuale.

³⁴ Il *Rich Text Format* è un formato proprietario per documenti multiplatforma. Sviluppato nel 1981 da Charles Simoni, poi membro della Microsoft, nel 1987 creò la versione definitiva con David Luebbert. Nonostante la maggior parte degli editor di testo e dei word processor per Microsoft, Windows, Mac OS e Linux siano in grado di leggere e scrivere documenti RTF, sui sistemi operativi Microsoft non è garantito il mantenimento della formattazione originaria, in quanto il formato non è mai stato standardizzato.

Un documento RTF è un file ASCII con stringhe di comandi speciali capaci di controllare le informazioni riguardanti la formattazione del testo, quali il tipo di carattere e il colore, i margini, i bordi del documento, etc.

Anche in questo tipo di codifica si utilizzano i tag. Le etichette o marcatori vengono inseriti nel corpo del testo e sono costituiti da sequenze di caratteri delimitati dai segni speciali delle parentesi angolate (<...>) che permettono all'elaboratore di distinguerli dal testo. Un esempio di testo codificato con un linguaggio di codifica dichiarativo è il seguente:

```
<chapter n="" id="">
<title type=""></title>
<p>
<name type=""></name>
</p>
[...]
```

Un linguaggio dichiarativo permette di avere dei vantaggi nel trattamento informatico delle informazioni: ad esempio, all'interno di un'applicazione per l'estrazione delle informazioni (*Information Retrieval*), il computer è in grado di trasmettere dati riguardanti la frequenza di una parola all'interno di un determinato blocco.

Inoltre, dal momento che non forniscono istruzioni capaci di specificare l'aspetto formale del testo, i linguaggi dichiarativi possono fornire un *output* formattato mediante il ricorso ad informazioni esterne, come i fogli di stile. Questi consentono all'utente di descrivere l'aspetto finale del testo una volta terminata la fase di codifica. Oltre a facilitare la separazione del contenuto dalla forma, i fogli di stile sono in grado di realizzare numerosi *output* con caratteristiche tipografiche e di impaginazione, partendo da uno stesso documento sorgente.

Le possibilità che i linguaggi di marcatura offrono agli umanisti sono molte, in quanto *se il linguaggio scelto è abbastanza potente, è possibile rappresentare il testo (i suoi elementi e i rapporti che intercorrono tra loro) a vari livelli. Si possono costruire vari modelli del testo: accanto alla struttura editoriale possiamo immaginare di codificare la sua struttura grammaticale (per esempio i tempi verbali), la struttura retorica (le figure, ma anche la distribuzione del discorso diretto e indiretto), i luoghi; e poi è solo questione di voglia e fantasia.*³⁵

Vedremo ora nel dettaglio alcuni linguaggi di tipo dichiarativo, quali l'SGML e l'XML e la sua applicazione più conosciuta, l'HTML e l'XML.

³⁵ G. Gigliozzi, *Il testo e il computer. Manuale di informatica per studi letterari*, Milano, Mondadori, 1998, p. 109.

1.5 L'SGML

L'SGML, acronimo di *Standard Generalized Markup Language*, è un metalinguaggio standard ISO (ISO 8879:1986 SGML) utilizzato per la definizione formale di documenti in modo tale da renderli indipendenti dal dispositivo, dal sistema e dall'applicazione con cui sono realizzati.

La sua ideazione fu necessaria per rimediare ai problemi legati all'interscambio dei documenti tra computer ed applicazioni diverse dovuti essenzialmente ad un *markup* di tipo procedurale. Per tale motivo, si pensò ad un nuovo linguaggio, nel quale le istruzioni per rappresentare i dati testuali fossero separate dal contenuto del documento, da evidenziare mediante opportuni marcatori (tag). La prima elaborazione avvenne presso i laboratori della GCA (*Graphic Communications Association*), dove venne definito il "*GenCode Concept*": tale codice stabiliva che erano necessari marcatori diversi a seconda della tipologia del testo. Inoltre, si basava sul fatto che documenti più piccoli potessero essere incorporati come elementi di documenti più grandi. Nel momento in cui il progetto si trasformò nel Comitato *GenCode*, si elaborò la prima versione del *Generalized Markup Language* (GML), linguaggio realizzato nell'ambito di un progetto di ricerca promosso dalla IBM sui sistemi integrati per la gestione dei documenti legali. Con esso si introdusse il concetto di documento formalmente definito con una struttura annidata, da implementare su *mainframe* dell'IBM, la quale lo adottò poi come standard di codifica per i propri file. Al completamento di GML, la ricerca non si concluse e i nuovi concetti che vennero elaborati finirono nell'SGML. Nel 1978, l'*American National Standards Institute* (ANSI) istituì un comitato per la definizione di uno standard capace di trasmettere e di archiviare i documenti. Supportato dal Comitato *GenCode* e, partendo da GML come base, fu avviata la standardizzazione di un linguaggio descrittivo, l'SGML. La prima versione delle specifiche dello standard fu pubblicata nel 1980, ma negli anni seguenti lo sviluppo fu condotto insieme all'ISO che lo adottò come standard col nome di ISO 8879 (1986).

La codifica prevista in SGML consente di rendere chiara ed esplicita l'interpretazione di ogni parte di un documento, descritta usando parole, lettere e altri caratteri stampabili. Essi descrivono il documento e indicano l'inizio e la fine di una parola o dei blocchi testuali con cui è composto il documento (paragrafi, sezioni, capitoli, etc.). Ad esempio, una delle applicazioni di SGML più conosciute è l'HTML (*HyperText Markup Language*), il quale, come vedremo successivamente, utilizza lo *Standard Generalized Markup Language* per definire un linguaggio che descriva la struttura dei documenti ipertestuali *World Wide Web*.

Esso, tuttavia, definendo un meccanismo formale per garantire al computer di interpretare correttamente un documento indipendentemente dal sistema su cui esso è stato realizzato, deve specificare le seguenti informazioni: gli elementi previsti dal linguaggio di *markup* e quelli necessari per la creazione di un documento; i loro delimitatori e la semantica che descrive il significato di ciò che si intende codificare; il contesto nel quale essi possono ricorrere.

Tuttavia, le caratteristiche fondamentali di SGML sono definite nella sua definizione: *Markup Language*³⁶*Generalized*³⁷ e *Standard*³⁸: l'indipendenza dall'hardware³⁹; l'indipendenza dal software⁴⁰, l'indipendenza dai sistemi di codifica dei caratteri⁴¹; l'indipendenza logica dalle tipologie di elaborazione⁴².

Affinché il testo codificato sia corretto, l'SGML ricorre alla *Document Type Definition* (DTD)⁴³, strumento per mezzo del quale si definisce il tipo di documento e si crea un'applicazione SGML, in quanto ... *mette a disposizione un strumento per la marcatura di documenti che indica la presenza, l'ordine e la posizione degli elementi e dei loro attributi in un documento.*⁴⁴ L'importanza della DTD risiede nel fatto che è in essa che sono elencati e definiti tutti gli elementi necessari all'elaborazione di un documento. Al suo interno si trovano, infatti gli elementi⁴⁵; il *Content Data Model*⁴⁶ e le entità⁴⁷.

³⁶ SGML è un meta-linguaggio per la codifica testuale, basato su un sistema di tag che vengono associati al contenuto di un testo per esplicitarne le struttura, gli elementi che la compongono, le relazioni di dipendenza e di ricorrenza fra questi.

³⁷ SGML è un tipo di codifica dichiarativo, astratto e generalizzato.

³⁸ SGML è uno standard, in quanto, come ho sopra accennato, è stato sviluppato dalla *International Standardization Organization* sotto la dicitura di ISO 8879:1986. 000

³⁹ SGML è un linguaggio autonomo rispetto alla piattaforma informatica utilizzata ed è un sistema di codifica astratto che può essere usato per archiviare dati su formati diversi.

⁴⁰ SGML è utilizzabile su qualsiasi sistema operativo e su qualsiasi editor di testo.

⁴¹ SGML è indipendente dalle lingue nazionali e dai sistemi di scrittura.

⁴² SGML rappresenta qualsiasi tipo di testo, indipendentemente dagli obiettivi e dai processi computazionali per il quale il testo è memorizzato e codificato.

⁴³ La DTD è una descrizione formale della struttura di una particolare classe di documenti. In essa sono specificati gli elementi che compongono un documento, le relazioni che intercorrono tra questi, il loro nome ed i dati che possono contenere. Ogni elemento può contenere altri elementi, dati o entrambi; nel *content data model* di ogni elemento è specificato che cosa questo può contenere e vengono fornite particolari notazioni, nelle quali ogni elemento è accompagnato da un simbolo: "+" (una o più ricorrenze dell'elemento); "*" (zero o più ricorrenze); "?" (zero o una ricorrenza).

Gli elementi possono essere raggruppati tramite parentesi tonde e separati dai seguenti simboli: "|" (almeno uno degli elementi della lista deve comparire); "&" (entrambi gli elementi separati da questo simbolo devono comparire, in qualsiasi ordine); "," (entrambi gli elementi separati da questo simbolo devono comparire, nell'ordine specificato).

⁴⁴ D. Shepherd, *XML. Guida completa*, trad. it., Milano, Apogeo, 2002, p. 54 (ed. orig. *Teach yourself XML in 21 Days second edition*, Sams, 2001).

⁴⁵ Sono i nomi dei marcatori adottati per identificare i blocchi strutturali e semantici che costituiscono l'albero dati e presenti nella DTD. Essi sono realizzati sulla base della loro funzione nel documento SGML e vengono identificati col proprio identificatore generico (GI, *Generic Identifier*). Inoltre, possono contenerne, a loro volta, altri elementi secondo le specifiche del *Content Data Model*.

⁴⁶ È il strumento di contenuto che indica, per ciascun elemento, il numero degli elementi che possono comparire dentro l'elemento di origine. Permette, inoltre, di identificare gli attributi da associare ad un elemento, che può, in questo

Lo *Standard Generalized Markup Language*⁴⁸ prevede, inoltre, la separazione del contenuto dalle istruzioni sulla rappresentazione, le quali devono essere poste all'esterno del documento SGML. L'utilizzo dei fogli di stile permette di rappresentare graficamente il testo e di aggiornarlo continuamente con informazioni specifiche a seconda delle situazioni.

modo, avere informazioni opzionali, utili sia per descrivere dettagliatamente la struttura dei dati, sia per essere veicolo di meta-informazione.

⁴⁷ Sono gruppi di dati immagazzinati in opportuni file. Sono utilizzate per separare files grandi in parti più piccole; per incorporare gli stessi dati in documenti diversi; per proteggere i dati non-SGML o caratteri di un documento che non fanno parte della tavola ISO 646 IRV (ASCII a 7 bit) dagli analizzatori sintattici, consentendo di includerli in un documento attraverso un riferimento; per creare riferimenti ai dati esterni, ad altri documenti o a file grafici. Le entità sono dichiarate nella DTD con un nome, delimitato da e commerciale (&) e dal punto e virgola (;).

⁴⁸ SGML è un tipo di codifica dichiarativo, astratto e generalizzato.

⁴⁹ SGML è un meta-linguaggio per la codifica testuale, basato su un sistema di tag che vengono associati al contenuto di un testo per esplicitarne le struttura, gli elementi che la compongono, le relazioni di dipendenza e di ricorrenza fra questi.

1.6 L'HTML

L'applicazione più diffusa del linguaggio SGML è l'*Hypertext Markup Language*, noto come HTML. Esso è un linguaggio di marcatura per ipertesti, i quali possono essere creati senza troppa difficoltà per la struttura semplice che esso presenta: si basa, infatti, su una DTD standard e non modificabile dall'utente.

Il suo ideatore è stato il ricercatore del CERN (*The European Organization for Nuclear Research*)⁵⁰ di Ginevra, Tim Berners-Lee che, nel marzo del 1989, presentò ai dirigenti dei laboratori una relazione dal titolo *Information Management: a Proposal*, nella quale illustrava la sua idea di *un sistema universale di informazioni fra loro collegate, in cui la generalità e la portabilità sono molto più importanti della bella grafica*.⁵¹ L'obiettivo di Berners-Lee fu quello di sviluppare un sistema di pubblicazione e di reperimento delle informazioni da distribuire geograficamente, in modo tale da tenere in contatto la comunità internazionale dei fisici.

Per il sociologo e filosofo Theodor Holm Nelson l'ipertesto è *scrittura non sequenziale, testo che si dirama e consente al lettore di scegliere: qualcosa che si fruisce al meglio davanti a uno schermo interattivo. Così come è comunemente inteso, un ipertesto è una serie di brani di testo tra cui sono definiti legami che consentono al lettore differenti cammini. [...] L'ipertesto include come caso particolare la scrittura sequenziale, ed è quindi la forma più generale di scrittura. Non più limitati alla sola sequenza, con un ipertesto possiamo creare nuove forme di scrittura che riflettano la struttura di ciò di cui scriviamo, e i lettori possono scegliere percorsi diversi a seconda delle loro attitudini, o del corso dei loro pensieri, in un modo finora ritenuto impossibile*.⁵²

Berners-Lee, insieme al Web, ideò un linguaggio di formattazione per creare ipertesti consultabili attraverso la Rete, l'HTML; il protocollo HTTP (*Hypertext Transfer Protocol*) e il metodo di collegamento dei documenti presenti su Internet, l'URL (*Universal Resource Locator*).

Sebbene per Berners-Lee l'obiettivo principale fosse il reperimento e la navigazione di informazioni e non il loro aspetto grafico, lo sviluppo del *World Wide Web* impose un cambiamento. Nel 1993 Marc Andressen ed Eric Bina, dottorandi del *National Center for Supercomputing Applications* (NCSA) dell'Università dell'Illinois, realizzarono la prima interfaccia grafica multiplatforma per l'accesso ai documenti presenti sul *World Wide Web*. Grazie al successo di *Mosaic*, Marc Andreesen e Jim Clark crearono la *Mosaic Communications* (in seguito, chiamata *Netscape Communications Corp.*), che, nel 1994, rilasciò la prima versione

⁵⁰ L'acronimo si riferisce alla dizione inglese: *European Organization for Nuclear Research*

⁵¹ T. Nelson, *Literary Machines 90.1*, Padova, Franco Muzzio Editore, 1992, cit. pp. 0/2.

⁵² T. Nelson, *Literary Machines 90.1*, Padova, Franco Muzzio Editore, 1992, cit. pp. 0/2.

del Browser Netscape. Il passaggio ai browser grafici segnò una svolta per Internet, in quanto permise di passare dall'ipertesto del CERN ad una impostazione di tipo multimediale.

La prima versione di HTML non ebbe mai un documento ufficiale di riferimento, in quanto fu soggetta a continue revisioni che, nel 1995, confluirono nella 2.0. Riguardo questa, nel novembre 1995, Tim Berners-Lee e Dan Connolly, pubblicarono le specifiche, per conto dell'*Internet Engineering Task Force* (IETF). Sin dalle prime revisioni, furono inseriti marcatori dal valore tipografico, come il corsivo `<i></i>`, il grassetto ``, le interruzioni di riga `
`, a cui si aggiunse poi la possibilità di inserire nel documento immagini e tabelle. La versione 3.2 di HTML incluse anche il linguaggio di *scripting* Javascript, favorendo, in questo modo, l'interattività con le pagine web. La versione 4.0 comprendeva, invece, le specifiche per il DHTML (*Dynamic HTML*). Nel 1996 fu rilasciata la prima versione delle specifiche per i CSS (*Cascading Style Sheets*), strumenti di formattazione per le pagine HTML attraverso i quali è possibile definire l'aspetto grafico dei documenti Web.

Grazie al successo del *World Wide Web*, Tim Berners-Lee fondò, nel 1994, il *World Wide Web Consortium* (W3C) presso il *Laboratory for Computer Science del Massachusetts Institute of Technology*, in collaborazione con il CERN e con il supporto del Dipartimento della Difesa Americano (*Defense Advanced Research Project Agency*, DARPA) e della Commissione Europea. Il W3C si costituì come consorzio di aziende del settore informatico interessato allo sviluppo di standard di riferimento per il *Web*, favorendo l'interoperabilità e la standardizzazione di tutte le tecnologie riguardanti Internet.

L'HTML ha rivoluzionato le modalità di invio e ricezione delle informazioni e ha permesso la visualizzazione dei dati. Questo linguaggio, tuttavia, ponendo attenzione al modo in cui esse sono presentate e non alla tipologia e alla loro struttura, si è rivelato inefficiente nella gestione degli aspetti morfo-sintattici e semantici di un documento. Per questo motivo, è stato sviluppato il linguaggio XML.

1.7 L'XML

L'XML, acronimo di *Extensible Markup Language*, è un linguaggio di *markup* aperto, che si basa su dati, che descrivono la struttura e la semantica dei blocchi testuali di un testo. Derivato da SGML come suo linguaggio semplificato, esso è ottimizzato per il Web, diventando un complemento dell'HTML.

Fu sviluppato da un gruppo di ricerca del W3C ed ebbe, fin da subito, l'obiettivo di utilizzare le funzionalità di SGML sul Web. Il risultato che ne derivò fu un linguaggio di marcatura che conteneva le caratteristiche principali del precedente, ma non quelle scarsamente utilizzate.

Lo sviluppo della prima stesura del progetto fu curata dall'*XML Working Group* (noto anche come *SGML Editorial Review Board*) e la prima pubblicazione si ebbe nel 1996 per merito del W3C. Nel febbraio del 1998 vennero poi rilasciate le specifiche del nuovo linguaggio di *markup* con il nome di *Extensible Markup Language 1.0*. Le basi del progetto risiedevano nel fatto che l'XML doveva essere utilizzabile in modo semplice su Internet, supportando un gran numero di applicazioni, ed essenziale nelle sue caratteristiche opzionali: i documenti codificati con questo linguaggio sono, pertanto, facilmente interpretabili dal computer. Tenuto poi conto dei limiti dell'HTML e dell'SGML, lo sviluppo dell'XML doveva essere rapido e la sua progettazione formale e concisa. Un documento steso dal W3C sottolineò come l'XML è un metodo per mettere dati strutturati in un file di testo. Per "dati strutturati" pensate a cose come fogli elettronici, agende, parametri di configurazione, transazioni finanziarie, disegni tecnici, ecc. I programmi che producono questi dati li salvano anche su disco, per cui possono usare sia il formato binario o sia quello testuale. Quest'ultimo permette, se necessario, di guardare i dati senza l'ausilio del programma che li ha creati. XML è un insieme di regole, di linee guida, convenzioni, in qualsiasi modo li vogliate chiamare, per progettare file di testo per questi dati, in un modo che produca file che siano facili da generare e leggere (da parte di un computer), che siano non ambigui, e che tenga lontano da pericoli comuni, come mancanza di estensibilità, mancanza del supporto per l'internazionalizzazione/localizzazione, e dipendenza dalla piattaforma [...] Come l'HTML, XML fa uso di tag (parole racchiuse tra '<' e '>') e attributi (della forma `name="value"`), ma mentre l'HTML specifica cosa ogni tag & attributo significhi (e spesso come un testo tra essi apparirà in un browser), XML usa i tag solo per delimitare pezzi di dati, e lascia l'interpretazione dei dati completamente all'applicazione

*che li legge. In altri termini, se vedi "<p>" in un file XML, non è detto che sia un paragrafo. A seconda del contesto, può essere un prezzo, un parametro, una persona,*⁵³

L'XML è, dunque, un metalinguaggio, in quanto consente di creare marcatori personalizzati spesso non associati ad una DTD.

In seguito venne introdotto anche il concetto di documento "ben formato", dove la verifica della validità e della corretta formazione di un documento XML è affidata a software detti *parser*. Per poter essere corretto, deve gestire il fatto che gli elementi, gli attributi e le entità sono *case sensitive* e, quindi, sensibili alla differenza tra maiuscolo e minuscolo; verificare che tutti gli elementi presentino sia il tag di apertura sia quello di chiusura, in quanto non sono consentite semplificazioni; controllare che i nomi degli elementi e degli attributi contengano solo caratteri alfanumerici (lettere dalla "a" alla "z"; dalla "A" alla "Z" e numeri da 0 a 9), sebbene siano accettati anche caratteri di punteggiatura, come il carattere di sottolineatura o underscore ("_"), il trattino ("-") ed il punto ("."). Inoltre, deve controllare che la marcatura sia separata dal contenuto testuale attraverso i caratteri speciali delle parentesi angolari ("< >"), valide per i tag, e la e commerciale ("&") per le entità. Nel momento in cui questi segni devono essere inseriti come elementi del testo, essi sono da codificare per mezzo delle entità predefinite XML: "&" per la e commerciale "&"; "<" per il simbolo di minore "<"; ">" per il simbolo di maggiore ">"; "'" per le virgolette singole (') e """ per le virgolette doppie (").

Inoltre, un documento XML si basa su una struttura gerarchica: esso è, infatti, costituito da un insieme di elementi, al cui interno si possono trovare sia informazioni aggiuntive (attributi) sulle proprietà sia altri elementi (sotto-elementi). L'organizzazione gerarchica degli elementi prevede un elemento principale o elemento radice, il quale contiene l'insieme degli altri elementi presenti all'interno del documento XML.

I marcatori XML sono memorizzati in formato ASCII e utilizzano il sistema di codifica Unicode/ISO10646, in particolare l'UTF-8 o l'UTF-16.

Altre caratteristiche importanti di questo linguaggio risiedono nella sua indipendenza dal tipo di piattaforma hardware o software utilizzata e dai dispositivi di archiviazione e di visualizzazione e nella sua capacità di rappresentare qualsiasi tipo di documento con finalità applicative diverse. Non a caso, Tim Bray lo definì l'ASCII del futuro.

⁵³ B. Bos, *XML in 10 punti*, trad. it. Emiliano Tellina, in *Latoserver.it* <http://www.latoserver.it/XML/in10punti/> (ed. orig. *XML in 10 points*, <<http://www.w3.org/XML/1999/XML-in-10-points>>)

Negli ultimi anni il W3C ha concentrato le proprie ricerche della materia soprattutto nello sviluppo di linguaggi basati sull'XML: ne sono esempio *XLink*⁵⁴; *Xpointer*⁵⁵; *Xpath*⁵⁶; *Xsl* e *Xslt*⁵⁷; *Xhtml*⁵⁸.

⁵⁴ Un linguaggio che permette di collegare tra di loro diversi documenti XML.

⁵⁵ Un linguaggio che permette di specificare le aree, dette *location*, a cui ci si indirizza all'interno di un documento.

⁵⁶ Un linguaggio che permette di scrivere espressioni con lo scopo di fornire informazioni all'interno di un documento.

⁵⁷ Due linguaggi che permettono di costruire fogli di stile atti a regolamentare la visualizzazione sul *Web* dei documenti XML.

⁵⁸ Si tratta dell'implementazione di HTML basata sul linguaggio di *markup* XML.

CAPITOLO II

Il Théâtre Italien di Evaristo Gherardi

È poesia popolare se mai quella non dal popolo, ma per il popolo, [...] e pare ovvio riconoscere che niente è più facile ad una cultura ben definita nei suoi compiti e disciplinata nell'opera, che scegliersi un'antitesi ed esercitarvisi ironicamente, a rappresentare quella che non è.⁵⁹

⁵⁹ Mario Apollonio, *Teatro e culture della rappresentazione. Alle origini della Commedia*, Bologna, Il Mulino, 1988, pag. 255.

Introduzione

Il Teatro Francese nel Seicento fu particolarmente vitale e la sua produzione cospicua fin dall'inizio del secolo. A partire, infatti, dal XVI secolo, furono attive le prime compagnie di attori girovaghi e dilettanti che, diversamente da quelle italiane, non si svilupparono spontaneamente, ma nacquero piuttosto nell'ambito di associazioni ricreative. La loro origine "borghese" determinò, per gran parte del Cinquecento, scelte di repertorio orientate a *moralité*, *sotties*⁶⁰ e farse⁶¹. In seguito, grazie all'interesse del pubblico, all'appoggio del Re (si fa riferimento, in particolare, al periodo di Luigi XIV), il quale vedeva nel teatro uno strumento di potere a sostegno della propria politica culturale, e all'arrivo dei comici dell'arte italiani, tale repertorio fu ampliato anche al genere delle commedie e delle tragedie di derivazione italiana. Fu soprattutto nella capitale che si ebbe una crescita della produzione teatrale e lo sviluppo dei teatri, ciascuno con la propria compagnia stabile: l'*Hôtel de Bourgogne*, il *Marais*, il *Petit-Bourbon* e il *Palais Royal*, regno di *Molière*, ne sono l'esempio. Furono molte le compagnie di attori professionisti italiani che allestirono spettacoli in tutta la Francia per guadagnarsi da vivere e che, in seguito, gestirono teatri, sempre in concorrenza una con l'altra per assicurarsi le *pièces* di successo e senza risparmiarsi attacchi personali.⁶²

In questi anni, il teatro comico, nel quale era forte l'origine popolare, prese il sopravvento sul teatro tragico, orientato sempre di più ad un pubblico di *élite*. La presenza dei comici dell'arte italiani risultò essenziale. Dopo varie *tournées* in Francia, intorno al 1660, una compagnia di comici dell'arte italiani si stabilì a Parigi restandovi quasi fino alla fine del secolo ed esercitando sul teatro francese una profonda influenza. Protetti da Maria de' Medici e da Mazzarino essi contribuirono alla rinascita del Teatro Francese.

Il teatro italiano ebbe, per questo, un tale successo che *Molière* collaborò con i comici dell'arte al *Petit-Bourbon*. Esso non fu, quindi, rappresentato soltanto in luoghi chiusi e di prestigio, ma anche nelle piazze con farse improvvisate, dando vita alla *Comédie Italienne*, molto gradita anche alla corte francese.

⁶⁰ Le *sotties* erano quelle rappresentazioni caratterizzate da un forte contenuto morale, dalla presenza di personificazioni allegoriche e da quella del *sot*, il "matto", il quale, vestito come un giullare, criticava i vizi della società mediante lunghi monologhi.

⁶¹ La farsa era una rappresentazione drammatica tipica del teatro francese del XV secolo, anche se si protrasse fino agli inizi del XVII. Si basava sulla presenza di alcuni elementi formali esteriori, quali il metro del verso, il numero dei personaggi (solitamente non più di quattro) e la lunghezza del testo (non superava mai i cinquecento versi), ma era priva di uno schema fisso. Essa poteva, infatti, avvicinarsi al realismo della commedia borghese o toccare toni violenti e grotteschi. I temi che solitamente la farsa esprimeva riguardavano la politica e il sociale, tanto che uno di quelli più ricorrenti era la polemica contro il clero o la corruzione dell'amministrazione della giustizia.

⁶² Lo stesso *Molière*, in merito all'*Improvvisazione di Versailles*, mise in scena una parodia riguardante gli attori dell'*Hôtel de Bourgogne*

I comici italiani, emigrati in Francia non per volontà, ma perché spinti dalla fame e spesso indesiderati e rifiutati dal potere delle corti, si radicarono in terra straniera: riproposero qui il “teatro degli italiani”, fatto di improvvisazione e di creatività, che contagiò altri attori, altri autori e altre forme di arte.

2.1 La nascita della Commedia dell'Arte

La Commedia dell'Arte è stato un importante genere teatrale: basata sull'improvvisazione dell'attore insieme ad altri attori, essa pose l'attenzione sui temi della vita sociale e familiare, dell'amore, della fame e della povertà, della sopraffazione e del potere; mise in scena tutto quello che ruota intorno all'uomo e alla società.

Nata come forma di reazione al Teatro di Corte, portato avanti da nobili che si improvvisavano attori per mezzo di testi poetici, letterari ed elitari, la Commedia dell'Arte nacque piuttosto da attori provenienti dal popolo. Sviluppata, dunque, come risposta al potere dei regnanti, degli aristocratici, degli ecclesiastici, essa mise in scena le proprie rappresentazioni nelle piazze, nei mercati, nelle strade, ma anche ai confini della città, poiché spesso il suo ingresso era vietato dalle autorità locali. I comici dell'arte utilizzavano semplici palchi o uscivano da dietro delle tende e, fin dall'inizio, si distinsero per l'utilizzo di costumi variopinti, arricchiti da elementi vistosi, come cappelli piumati, o per il ricorso a strumenti musicali al fine di richiamare l'attenzione dei passanti e di dare una scansione ritmica alle scene improvvisate. Inoltre, per rendere le rappresentazioni ancora più "teatrali" ricche, complesse e – per quanto possibile – irregolari, ricorsero all'uso della maschera, creata sulla base di tipi fissi della loro epoca.



**Jacques Callot⁶³, Danza di due zanni, 1621,
incisione tratta da *I balli di Sfessania***

Tuttavia, l'evoluzione di queste forme improvvisate in Commedia dell'Arte è frutto di una serie di una serie di contingenze politiche e sociali, che precludevano la possibilità di realizzare una produzione teatrale incentrata sull'impegno politico e sulla polemica sociale. A contribuire al suo sviluppo fu, senza dubbio, la nascita dei teatri privati, soprattutto a Venezia, dove le famiglie nobili si impegnarono a creare e a diffondere spazi preposti alla recitazione di commedie e di melodrammi. Fu in questi anni che le famiglie dei *Tron* e dei

⁶³ Jacques Callot (Nancy, Francia, 1592 - Nancy, 1635) era un incisore francese.

Michil fecero costruire due sale nella zona di San Cassiano, che, secondo una testimonianza del 1581 di Francesco Sansovino, figlio dell'architetto Jacopo Sansovino, furono i primi due teatri aperti al popolo. Fino a quel momento esso era stato escluso da ogni tipo di rappresentazione teatrale, riservata esclusivamente alle corti. Dopo i *Tron* e i *Michil* altre famiglie si dedicarono alla costruzione di teatri come forma di investimento: ad esempio, i *Vendramin* e i *Grimani* fecero costruire il Teatro dei SS. Giovanni, il Teatro San Luca e quello di Sant'Angelo. Per questo motivo, lo sviluppo dei teatri permise all'attore, fino a quel momento acrobata di strada o giullare di corte, di esibirsi in trame più complesse e di iniziare a costituire delle compagnie.

A tale proposito, non è un caso che alcuni vedano nella parola "arte" della Commedia dell'Arte un significato più ampio, che rimanda a quello di "corporazione" e che designa un teatro costituito da attori di professione, riuniti in compagnie al fine di tutelare i propri interessi economici. Con questa accezione del termine "fare teatro" significava, pertanto, appartenere ad una precisa corporazione ed esercitare una vera attività economica. La Commedia dell'Arte è alla base del teatro moderno, in quanto *è creazione italiana. Agli italiani, autori della Commedia dell'Arte, risalgono non solo la costituzione industriale e commerciale, ma la moltiplicazione dei teatri stabili, l'introduzione delle donne come attrici, i congegni dei mutamenti di scena, le analoghe formazioni del teatro musicale.*⁶⁴

Per il suo carattere comico essa venne definita anche con altri nomi: Commedia dei Buffoni o degli Istrioni (definizione che contrappone gli attori professionisti da quelli dilettanti delle corti); Commedia delle Maschere; Commedia all'Improvvisto o a Soggetto. Tutte queste dizioni evidenziano che tipico di questo genere era il fatto di recitare sulla base di una traccia (il canovaccio), che si mirava a coinvolgere il pubblico con frizzi⁶⁵ e lazzi⁶⁶ comici, che non si imparava il testo a memoria e che la buona riuscita della rappresentazione era legata ad un lavoro di *équipe*.

2.1.1 Il ruolo della maschera nella Commedia dell'Arte

La Commedia dell'Arte è definita anche Commedia delle Maschere, in quanto una delle sue caratteristiche più importanti è quella di essere rappresentata con le maschere, aspetto

⁶⁴ Benedetto Croce, 1932, 32, fonte "*Le arti della scena. Lo spettacolo in Occidente da Eschilo al trionfo dell'opera*".

⁶⁵ Il "frizzo" è una battuta pungente e arguta, che deriva dal verbo "frizzare", a sua volta derivato dalla forma latina non attestata "frictiare", intensivo di "frigere".

⁶⁶ Il termine "lazzo" indicava una breve scena di carattere mimico, predisposta nel canovaccio per interrompere la monotonia del dialogo e si riferisce al mondo della Commedia dell'arte e della farsa. In seguito, indicò poi un atto o motto buffonesco in genere, spesso sguaiato. L'etimologia della parola è tuttora oggetto di discussione, perché c'è chi rintraccia una derivazione dal latino "actio" (azione, atto); chi vede in essa un'origine spagnola della voce e chi ancora pensa all'italiano settentrionale *lazzo* (laccio) o al germanico *hazus* (istrione).

tipico del teatro greco antico. La parola “maschera” deriva, infatti, dal longobardo “*maska*”, il cui significato rimanda ad uno spirito maligno che divora uomini vivi.

Il successo ottenuto con la Commedia dell’Arte permise ai comici italiani di rivoluzionare i luoghi comuni del teatro. Ad ogni attore spettò, infatti, una parte ben precisa, un personaggio sul quale si doveva specializzare, affinandone sempre di più i contorni e le sfumature. Nacquero in questo modo le maschere della Commedia dell’Arte, che altro non erano che personaggi dalle caratteristiche fisse, alle quali poteva essere continuamente cambiata la fisionomia: il volto, l’abito, gli ornamenti, il modo di parlare e di comportarsi era considerato e sperimentato. Le maschere si andarono sempre più perfezionando: dai primi tratti iniziali, schematici ed essenziali, esse si caratterizzarono per una definizione non solo esteriore, ma anche psicologica.

La maschera costituì, dunque, un fattore di grande attrazione che conferì alla Commedia dell’Arte una sensazione di proibito, di anti-sacro e di anti-convenzionale: pertanto, essa attrasse sia le classi più popolari sia quelle agiate che trovavano in questo genere il contrario di tutto ciò che era preconstituito. L’infrazione alla regola diventa la regola. Pertanto, le maschere della Commedia dell’Arte facevano ridere, ma il loro fascino conteneva anche note di inquietudine, che andavano a colpire ciò che di convenzionale, di preconstituito è in tutti noi.

2.1.2 Il “non scritto” della Commedia dell’Arte

Nel 1545 si ha testimonianza della nascita della prima compagnia professionista: fu anche l’anno del Concilio di Trento e dell’inizio della Controriforma, la quale ostacolò ogni tipo di sperimentazione teatrale. La Commedia dell’Arte fu fortemente attaccata per due motivi: sia per lo stile di vita condotto dagli attori, considerati uomini immorali; sia per il fatto che sui canovacci era impossibile applicare qualsiasi forma di censura. La relazione della Curia, stilata nel 1578 dal Cardinale Paeolotti, né è prova, in quanto evidenzia l’irritazione dei censori di fronte alle tracce dei comici e il carattere fondante della Commedia dell’Arte, ovvero l’improvvisazione.

Il comico dell’arte si riservava, infatti, la possibilità, una volta in scena, di modificare qualche battuta rispetto alla tradizione consolidata per rendere più attuale la commedia da lui rappresentata. Se nei teatri francesi, inglesi, spagnoli, il teatro si legava saldamente alla figura del drammaturgo, la Commedia dell’Arte ne era priva, poiché quello a cui essa si riferiva era semplicemente un elenco di cose da fare al momento di entrare in scena.

La recitazione della Commedia dell'Arte si modificò in improvvisazione. E questa, sebbene per tutti non fosse il suo tratto distintivo, ebbe comunque il merito di mettere l'attore al centro, che diventava un vero e proprio istrione, in grado di attirare, mantenere, accentrare tutta l'attenzione del pubblico.

2.1.3 Recitare all'improvviso nella Commedia dell'Arte

“Non ci si deve aspettare di trovare in questo *Recueil* delle Commedie intere, poiché le *Pièces Italiennes* (le opere teatrali italiane) non possono essere stampate. La ragione è che i *Comédiens Italiens* (i comici italiani) non imparano niente a memoria a loro basta, per metter in scena un'opera, di averne visto il soggetto un momento prima di entrare in scena. Così la più grande bellezza delle loro *Pièces* è inseparabile dall'azione, il successo delle loro Commedie dipende assolutamente dagli Attori, che assegnano più o meno fascino a seconda che abbiano più o meno esprit (spirito) e a seconda della buona o cattiva situazione in cui si trovano mentre recitano. È questa necessità di recitare sul campo che fa sì che si abbia tanta difficoltà a sostituire un buon *Comédien Italien*, quando sfortunatamente egli viene meno. Chiunque può imparare a memoria e recitare sul Teatro ciò che avrà appreso: ma ci vuole ben altro per il *Comédien Italien*. Chi dice buon *Comédien Italien* intende un uomo che ha spessore, che recita più con l'immaginazione che con la memoria; che compone, recitando, tutto ciò che dice; che sa assecondare colui con cui si trova sul Teatro; vale a dire che sposa così bene le sue parole e le sue azioni con quelle del suo Compagno, che scende in campo con tutto il gioco e con tutti i movimenti che l'altro richiede, in modo da far credere a tutti che erano già d'accordo. Non si può dire lo stesso per un Attore che recita semplicemente a memoria; egli entra in Scena solo per recitare il prima possibile ciò che ha imparato a memoria ed è talmente preso da questo che, senza fare attenzione ai movimenti e ai gesti del suo Compagno, va per la sua strada, in una sorta di furiosa impazienza di liberarsi del proprio ruolo come di un fardello che lo affatica molto. Possiamo dire che questi Attori sono come Scolari, che vengono tremolanti a ripetere una lezione imparata con cura: o meglio essi somigliano agli eco che parlano solo se altri hanno parlato prima di loro. Sono attori di nome, ma inutili e a carico della loro Compagnia. Paragono un Attore di questo tipo a un Braccio paralitico che, benché inutile, porta ancora il nome di braccio. L'unica differenza che trova tra il braccio morto e il membro inutile della Commedia, è che se il primo non serve a niente al corpo è altresì vero che non ne riceve alcun nutrimento e che questo si divide tra i membri che fanno il loro dovere: ma l'ultimo (benché assolutamente inutile alla *Comédie*) non si stanca di ricevere tanto cibo

quanto gli Attori che faticano maggiormente, e che sono i più necessari. Questo vale per tutti quegli Attori inutili di cui quasi tutte le compagnie sono piene; Gente senza nature (ingegno) e senza arte, che una protezione capricciosa o una fortuna straordinaria ha elevato fino alla Parte intera, e che da allora guardano alla *Comédie* soltanto dal lato della Cassetta, e non da quello dell'impiego che essa esige da parte loro: facendo una differenza netta tra questi Attori di nome e quelli di fatto, questi Attori illustri che imparano a memoria la verità, ma che come Pittori eccellenti, sanno nascondere l'arte con l'arte e che affascinano gli Spettatori con la bellezza della voce, la verità del gesto, la giusta flessibilità del tono, e una certa aria graziosa, a proprio agio e naturale con cui accompagnano tutti i loro movimenti, e che diffondono su tutto quello che pronunciano".⁶⁷ Così scrisse Evaristo Gherardi nel suo *Recueil* a proposito dei comici dell'arte, offrendo una testimonianza fondamentale delle caratteristiche del genere: la preponderanza dell'attore sull'autore, la prevalenza del gesto e dell'azione sui momenti di riflessione, la grande preparazione professionale e il lavoro di *équipe*. La capacità di improvvisare dipendeva esclusivamente dal repertorio del singolo attore, che raccoglieva, in contesti sempre diversi, una serie di battute da applicare alle varie situazioni comiche. Non c'era spazio per la riflessione, che avrebbe inficiato il valore della battuta basata sul *calembour*, sul fraintendimento e sulla messa "in ridicolo" dell'eventuale antagonista. Fare questo implicava una grande preparazione professionale, una cultura non indifferente, una capacità di intuire gli umori del pubblico per adattarsi al meglio, un calcolo dei tempi di recitazione che non dovevano mai essere troppo lunghi onde evitare la noia e la distrazione. Un'altra caratteristica preponderante della Commedia dell'Arte era l'interazione fra comici, i quali proprio perché legati all'improvvisazione, dovevano stare molto attenti a quanto faceva o diceva l'altro personaggio sulla scena, perché in base ad esso dovevano variare le proprie battute.

Infine, l'attore era un personaggio poliedrico, poiché doveva saper fare tutto: cantare, ballare e suonare.

La Commedia dell'Arte era molto "meno improvvisata" di quanto si pensasse.

⁶⁷ La fonte di quanto inserito nelle virgolette è frutto di una nostra traduzione di servizio, tratta dall'*Avertissement qui'l faut* (Avvertimento per chi legge), inserita da Evaristo Gherardi nel I volume come peritesto dell'intera collezione.

2.3 La Commedia dell'Arte in Francia

La spontaneità e la vitalità della Commedia dell'Arte rese questo genere ben accetto alle corti tardo-rinascimentali dei regni italiani e l'eco della loro bravura raggiunse ben presto anche gli altri stati europei, i quali iniziarono ad invitare le migliori compagnie italiane presso le proprie corti. Tanto successo era dovuto sia alle innovazioni introdotte dal genere della Commedia dell'Arte sia alle qualità degli attori che vi appartenevano: uomini che, per la cultura raffinata e per una profonda conoscenza della mente umana, erano capaci di improvvisare. Alcuni commedianti provenivano dal mondo dell'aristocrazia, altri dalle prestigiose Accademie della Crusca e degli Spensierati, altri ancora dal popolo. Per motivazioni di carattere morale e di interesse economico, essi non ebbero mai rapporti stabili con le autorità religiose o politiche, tali da generare l'aspetto itinerante del loro teatro e la ricerca di un pubblico sempre nuovo.

Per questo motivo, i commedianti italiani iniziarono a viaggiare per tutta Europa, ma, in particolare, in Francia, dove le testimonianze della loro presenza sono molte. A partire dal 1500 arrivò, alla corte di Carlo IX, la compagnia del capo-comico "Zan Ganassa", ovvero Alberto Naselli. Egli si era esibito a Ferrara nel ruolo dello "zanni" durante un matrimonio estense alla presenza del Re francese, il quale lo volle a Parigi per celebrare il suo ingresso nella capitale. Carlo IX rimase a tal punto colpito dalla bravura dell'attore italiano che lo invitò a rimanere a corte per quasi un anno. In seguito, sebbene Zan Ganassa e la sua compagnia avessero avuto il permesso del sovrano per rappresentazioni pubbliche, il Parlamento fermò ogni attività. Tale fatto può essere spiegato sia nella mancata adesione dei commedianti italiani alle richieste della Confraternita della Passione, la quale aveva il privilegio di gestire le rappresentazioni teatrali e con esse i suoi incassi o nel fatto che era in atto il tentativo di neutralizzare Caterina de' Medici. Secondo alcuni, infatti, la regina aveva assoldato i comici italiani per distrarre il figlio Carlo IX dalle sue congiure con la parte protestante. Tale posizione potrebbe essere avvalorata dal fatto che le tensioni religiose, sfociate nel massacro di San Bartolomeo, avevano seguito di pochi giorni una rappresentazione della Compagnia di Zan Ganassa organizzata per festeggiare il matrimonio della figlia di Caterina de' Medici, Margherita di Valois, con il protestante Enrico di Navarra.

A partire dal 1571, fu invitata in Francia la compagnia dei Gelosi per rappresentare una commedia in occasione del battesimo di Charles Henry de Clermont. Essa, tuttavia, si esibì anche nel palazzo di Luigi Gonzaga, Duca di *Nevers*, alla presenza di Carlo IV, Duca di Lorena, e alla corte di Carlo IX e di Caterina de' Medici. Questa compagnia, attiva a Milano

e nel Nord-Italia per oltre 35 anni, fu protetta sia dai Duchi di Mantova che di Ferrara. Sebbene questo appoggio le avesse garantito una stabilità economica, la compagnia non si legò mai a nessuna corte, conservando la propria indipendenza che gli permise di recarsi nei luoghi francesi nei momenti che riteneva più opportuni. Ritornati in patria per evitare la guerra civile tra cattolici e ugonotti, la Compagnia continuò la propria attività teatrale presso la corte estense, per la quale mise in scena *L'Aminta* di Torquato Tasso alla presenza di Filippo II di Spagna. A Venezia furono ammirati da Enrico III, l'ultimo figlio di Caterina de' Medici, in procinto di essere incoronato sovrano per la morte di Carlo IX, il quale invitò i commedianti italiani a *Blois* (Parigi). L'invito fu accettato soltanto nel 1577, quando i conflitti interni del paese, legati a tensioni religiose, si furono placati. Tra i capo-comici alla guida della Compagnia dei Gelosi è utile ricordare Vittoria Piissini, attrice che entusiasmò Enrico III nelle recite veneziane, e che lasciò il gruppo per la guida della Compagnia dei Confidenti (1590); ad essa succedette Francesco Andreini (1578), inventore della maschera di Capitan Spaventa e marito di Isabella da Padova, colei che creò il personaggio di Isabella.

Tuttavia, è dopo il 1599 che i comici dell'arte ebbero una frequentazione assidua con la Francia, quando, in occasione delle nozze di Maria de' Medici con Enrico IV di Navarra, il re chiese al Duca di Mantova la Compagnia degli Accesi e l'Arlecchino Tristano Martinelli. In quegli anni le compagnie meno prestigiose recitavano nelle locande, nei cortili e nei palazzi signorili; affittavano di frequente le sale della Pallacorda che venivano trasformate in rudimentali teatri. Il successo dei comici italiani fu tale che prima nel 1603 e poi nel 1614, presero in affitto (in particolare Giovan Battista Andreini, figlio di Francesco, e Tristano Martinelli) una sala dell'*Hôtel de Bourgogne* al posto del *Palais-Royal*. Nel Sei/Settecento la loro emigrazione si fece sempre più assidua, diventando anche una sorta di scuola teatrale: ne è testimonianza la carriera di *Molière*, il quale apprese l'arte della recitazione dallo *Scaramouche* Tiberio Fiorilli. Nel 1653 una compagnia di attori italiani si insediò in modo stabile a Parigi con il nome di *Ancienne Troupe de la Comédie Italienne* sotto la direzione di Domenico Locatelli, in arte Trivellino e, a partire dal 1660, si stabilì al *Palais-Royal*, alternandosi alla Compagnia di *Molière* e influenzandola soprattutto nell'arte di disegnare i personaggi con brevi tratti. Nacque, pertanto, la *Comédie Italienne* che, sotto la protezione del Re, presentò al pubblico francese opere teatrali con attori di professione italiani. All'inizio, la *Comédie Italienne* mise in scena opere della Commedia dell'Arte nella loro lingua originale; in seguito, si fuse con i drammaturghi francesi dell'epoca (in particolare, con *Molière*), introducendo spesso personaggi nuovi rispetto alla tradizione (*Pierrot*, *Polichelle* e *Scaramouche*). Il personaggio preferito dai francesi fu comunque Arlecchino: i suoi interpreti più importanti, Domenico Biancolelli ed Evaristo Gherardi, recitarono esclusivamente in Francia e per essa

inventarono nuovi canovacci e repertori inediti in Italia. In questo modo, i comici italiani si francesizzarono e adattarono il proprio repertorio alla lingua francese.

La stabilità della *Comédie Italienne* non durò a lungo: il 4 maggio 1697, durante la rappresentazione de *La fausse prude* (“La Falsa Pudibonda”), ispirata alla *Reine Anonyme*, *Madame de Maintenon*, amante prima e moglie morganatica poi di Luigi XIV, il Mezzetin Angelo Costantini provocò, con la propria satira, l’ira del Re Sole, il quale chiuse l’*Hôtel de Bourgogne* ed espulse i comici italiani da tutta la Francia, dove ritornarono dopo il 1716.



Anonimo del Settecento (*Le Havre, Musée de la Ville*).



Incisione di Louis Jacob⁶⁸,
tratta dal disegno *Départ des Comédiens Italiens en 1697*,
di Jean Antoine Watteau⁶⁹, 1729

⁶⁸ Louis Jacob è stato un incisore francese vissuto nei secoli XVII e XVIII.

Per molti anni, la tradizione comica della Commedia dell'Arte sopravvisse, dunque, grazie all'attività dei comici attivi nei teatri montati nei mercati e nelle fiere, oggetto delle critiche di Evaristo Gherardi.

Gli attori italiani, ritornati a Parigi, nel 1762 si unirono all'*Opéra-Comique*, che fino al 1780 presenterà sempre il nome di *Comédie Italienne*.

⁶⁹ Jean Antoine Watteau (Valenciennes, 1684 – Nogent-sur-Marme, 1721) è stato un pittore francese. Presso i mercanti d'arte e di stampe di Rue Saint-Jacques e, in particolare, da Pierre II e da suo figlio Jean studiò le opere di Tiziano, Rubens, Jacques Callot. E forse proprio in questo contesto che entrò in contatto con il pittore di costumi e di scene teatrali Claude Gillot: entrò nella sua bottega nel 1703 e vi rimase fino al 1707/1708. In questo periodo l'artista lavorò sui soggetti della *Comédie Italienne* (ne è un esempio il disegno *Arlecchino imperatore della luna*, 1708 (?), ispirato probabilmente dal disegno di Gillot *Per amore di una bella*, 1706 e conservato oggi al Musée des Beaux-Arts). Prendendo spunto dal teatro creò nuove forme iconografiche (ne sono esempio *I Piccoli Comédiens*, 1706-08, Parigi, Museo Carnevale).

2.4 Evaristo Gherardi



**Arlecchino raffigurato in due atteggiamenti significativi (XVIII secolo):
Arlecchino sospirante e Arlecchino ghiottone.
Incisioni tratte dalle *Scènes comiques* di Claude Gillot
(Roma, Biblioteca e Raccolta teatrale del Burcardo).
L'Arlecchino ritratto è forse l'attore Evaristo Gherardi.**

Figlio del comico Giovanni Gherardi da Spoleto, conosciuto anche con il soprannome di “flautino” per aver inventato questa maschera, Evaristo Gherardi nacque a Prato l'undici novembre 1663. Trasferitosi a Parigi nel 1683, dopo aver studiato filosofia al *Collège de la Marche*, fu invitato a proseguire la carriera comica paterna piuttosto che gli studi umanistici, rispettando il motto dell'antico proverbio *è meglio un ottimo Arlecchino in libertà che un filosofo alla Pastiglia*. Nel 1689, anno in cui un documento ufficiale lo qualifica come *professeur ès langues étrangères* (“Professore di Lingue Straniere”), si mostrò per la prima volta sulle scene nel ruolo di Arlecchino, libero prima per la morte del comico Domenico Biancolelli e poi per il passaggio di Angelo Costantini al ruolo di *Mezzetin*. Fu anche la fame, come ricorda Umberto Cecchi, in un suo saggio, a supportare questa decisione perché prima o poi sarebbe arrivata *a travolgere e a stravolgere le cose del mondo e a dimostrare che il potere dei re non viene da Dio, ma dalla remissività, dalla magnanimità, dalla stupidità degli uomini nei confronti di un loro simile*.⁷⁰ Soltanto l'arte improvvisata, una professione senza scrittura e di satira, la denuncia

⁷⁰ Umberto Cecchi, *L'Arlecchino del Re Sole. La vita e il teatro di Evaristo Gherardi*, Edizioni del Palazzo, Prato 1986.

di un lazzo, poteva garantirgli la possibilità di una vita alternativa. Nel 1697, a seguito della cacciata degli italiani da Parigi, Evaristo Gherardi si dedicò a raccogliere le migliori opere e scene francesi recitate nel *Théâtre Italian* in un volume che uscì con il titolo de *Le Théâtre Italian*. Il suo *Recueil* ebbe una prima pubblicazione tra il 1691 e il 1697 a Bruxelles senza che venisse presentato il nome del curatore; nel 1700, invece, vennero pubblicati a Parigi i sei volumi definitivi con il nome di Evaristo Gherardi. Benché la raccolta fosse stata censurata, venne ristampata più volte e spesso addirittura contraffatta, come testimonia lo stesso curatore nel suo primo volume. Morì all'improvviso il trentuno agosto 1700 (si racconta che avesse battuto la testa mentre recitava a *Saint-Maur* e che avesse trascurato questo incidente), tornando da un viaggio a Versailles, luogo in cui si era recato per presentare la propria opera al Delfino di Francia. Evaristo Gherardi scrisse una sola commedia, *Il ritorno dalla fiera di Bezons*, inserita nel suo *Recueil*.

Il merito di Gherardi fu, tuttavia, anche quello di aver reso il personaggio di Arlecchino meno grottesco, più raffinato e colto, giocando più sulle sfumature del linguaggio che sulla mimica del corpo. Arrivato sulle scene dopo anni di studi di filosofia, promosse nuove sperimentazioni teatrali e un diverso atteggiamento nel mestiere dell'attore.

2.5 Il *Théâtre Italien* di Evaristo Gherardi⁷¹

Il *Théâtre Italien* di Evaristo Gherardi è una raccolta di *Scenes Françaises* recitate a Parigi. Come afferma lo stesso *queste scene sono opera di diverse persone di spirito e di merito, composte per la maggior parte nelle ore di ricreazione e date da alcuni gratuitamente alla Troupe. Erano come incastrate nei nostri soggetti. Tutta Parigi le ha ammirate quando le abbiamo messe in scena, e tutta Parigi le rimpiange oggi che non sono più in scena.*

Essa è da considerarsi, pertanto, una testimonianza fondamentale per conoscere l'evoluzione del comico italiano in Francia nel Sei/Settecento. Lo scopo non fu soltanto quello di raccogliere al suo interno le cinquantacinque commedie rappresentate all'*Hôtel de Bourgogne* tra il 1680 e il 1697, quanto piuttosto quella di raccontare l'origine dei canovacci, i quali, a seguito della cacciata degli italiani da Parigi per volontà di Luigi XIV, vennero talvolta rappresentati alla *Foire* ("spettacoli della fiera"). E' difficile fornire informazioni precise sull'attività dei teatri di fiera fino al 1713 circa e delle rappresentazioni che avvenivano in provincia dopo il 1697. Quello che è quasi sicuro è che sia i gestori dei teatri *forains* sia le compagnie sfruttarono l'ampio repertorio su cui, per anni, era vissuto il *Théâtre Italien*, un patrimonio spesso inedito e non ancora tutelato da alcun *privilege*. Per questo motivo, Evaristo Gherardi, nel 1700, con il suo *Recueil*, denunciò il saccheggio degli impresari e dei capo-comici *forains*, colpevoli, secondo lui, di aver plagiato i canovacci della Commedia dell'Arte e le conquiste tecniche, scenografiche e comiche della *Comédie Italienne* in Francia. Ventuno anni dopo l'edizione del 1700, a seguito della pubblicazione della raccolta di Gherardi, uscì quella de *Le Théâtre de la Foire*, nella quale il commediografo francese *Alain-René Lesage* (1688-1747) ammise che, almeno nella prima fase della produzione forense, si era attinto dal patrimonio italiano.

La prima pubblicazione del *Recueil* risale al 1694 e, come ricorda il suo curatore, *provocò l'invidia dei miei Compagni*. Questa raccolta, sebbene fosse stata censurata, ebbe grande successo da essere contraffatta e venduta in Olanda, a Bruxelles, a Liegi e in tutte le Province del Regno.

Evaristo Gherardi non trascurò *niente di tutto ciò che poteva abbellire l'opera e dare piacere al Lettore* : sono presenti scene in italiano, in francese e in entrambe le lingue (circa quaranta), di cui il curatore dichiara di aver corretto soltanto il dialetto di *Pasquariel*. Tali commedie, definite Commedie Francesi adattate al *Théâtre Italien*, non appartenevano al repertorio delle

⁷¹ Il materiale utilizzato per scrivere questo paragrafo è stato attinto dall'*Avertissement* (Avvertimento per chi legge) di Evaristo Gherardi, conservato nel primo volume del *Recueil*.

Pièces Italiennes, ma a quello della *Troupe*, obbligata ad uniformarsi al gusto e all'intelligenza della maggior parte dei suoi spettatori e ad inserire molto più francese che italiano. I dialoghi presenti nel *Recueil* sono ricchi di una satira fine e delicata, di una conoscenza perfetta dei costumi del secolo e di nuove espressioni.

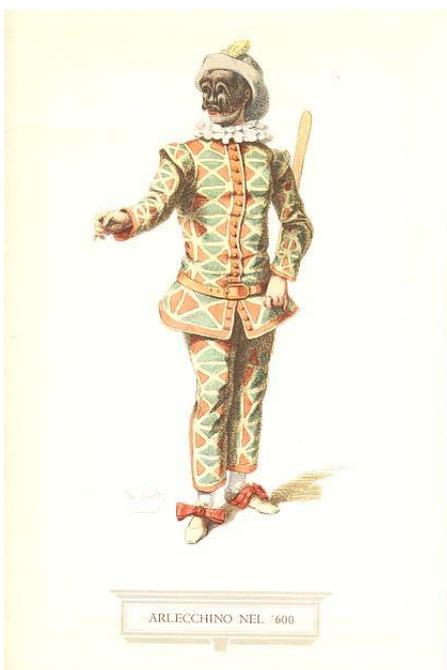
2.6 Le maschere de *Le Théâtre Italian*

Le maschere presenti nelle commedie del *Recueil* di Evaristo Gherardi hanno spesso una caratterizzazione non italiana, seppure da essa abbiano preso spunto. La maggior parte di loro sono nate e si sono sviluppate in Francia, paese nel quale la Commedia dell'Arte italiana fu apprezzata più che in altre parti d'Europa. Qui le maschere italiane si sono evolute, facendo proprie le caratteristiche del luogo che le ha accolte: prima fra tutti, la lingua e il modo di esprimersi; ma ancora gli atteggiamenti, il carattere, certi requisiti e vizi. Ad esempio, *Pierrot*, francesismo derivato dal personaggio italiano Pedrolino, uno dei primi “zanni”, interpretato da Giovanni Pellesini (Compagnia dei Gelosi) come un servo pacato, ingenuo e remissivo, marito fedele, che all'azione preferiva la considerazione di sé come uno sciocco. In Francia, esso perse i suoi tratti originari e la sua ingenuità si trasformò in onestà, nota dominante del suo carattere. *Pierrot* diventò languido e coraggioso, pigro e sfacciato, misterioso e mai prevedibile.

Quello che testimoniano queste caratterizzazioni ed evoluzioni è che i personaggi di *Pierrot*, *Polichinelles* (Pulcinella), *Scaramouche* (Scaramuccia) e tante altri di origine italiana, hanno saputo adattarsi a luoghi diversi, assumendo ogni volta caratteristiche sempre nuove.

Per questo motivo, si è voluto inserire una breve rassegna delle principali maschere italiane presenti nella raccolta di Evaristo Gherardi per capirne l'importanza dell'influsso all'interno del panorama francese.

2.6.1 *Arlequin*



Racconta la tradizione che la maschera di Arlecchino sia nata a Bergamo e che il suo costume sia stato cucito dalla madre, poverissima, con gli scampoli di vari colori. Secondo un'altra versione pare che Arlecchino fosse stato al servizio di uno speciale avaro, il quale lo vestiva con le toppe degli abiti consumati. L'origine del nome è tuttora oggetto di molte ipotesi: c'è chi vede in essa una derivazione della parola francese *Herlequins* o *Hellequins* (“diavolo” o “buffone”), rintracciabile in leggende folcloristiche francesi (*Chasse Arthur* o *Chasse Saint Hubert*), in cui si narrano cavalcate

notturne di demoni e di dannati, i *Mesnie Herlequin* (“la famiglia degli arlecchini”), alla ricerca di prede animali e umane. C’è poi chi vede l’origine nobile nel Conte di *Boulogne*, il quale, dopo essere caduto da cavallo combattendo contro i normanni, per vendicarsi avrebbe dato origine ad una serie di scherzi, o in *Achille da Harlay*, un gentiluomo alla corte di Enrico III, che aveva protetto un saltimbanco italiano. Il suo nome potrebbe anche derivare dalla saga di *Erl - Koenig* (“Re degli elfi”), in quanto Arlecchino era solito recarsi nei boschi, insieme ai suoi compagni, per stendere incantesimi e fare scherzi. L’ultima ipotesi è quella che ricollega il mimo comico italiano ad una possibile prosecuzione di quelli greci e latini, ipotizzando un passaggio da *Erculicunus* (“Erocole”) a *Herluquin*, in quanto il primo era spesso rappresentato come personaggio comico. Aldilà della provenienza, il nome di Arlecchino fu comunque diffuso da Alberto Naselli, il quale, dopo aver inventato il personaggio di “Zan Ganassa”, ebbe l’opportunità, lavorando in Francia nel XVI secolo, di mescolare il mito del diavolo buffone con quello del contadino rozzo di Bertoldo.

La maschera che indossa Arlecchino ha lo scopo di esaltare i difetti dei volti umani, anche quelli che non visibili: la sagoma nera accentua, infatti, due enormi archi orbitali e una fronte prominente. Il colore nero potrebbe poi trarre spunto dalla professione dei facchini o dei servi di piazza, incaricati al trasporto del carbone.

*Maurice Sand*⁷² fornisce poi delle spiegazioni riguardanti il costume. Egli parla, infatti, di una prima versione (1570), formata da una giacca e da pantaloni gialli, coperti entrambi da pezze di stoffa verde e blu; di una seconda, a cui viene dato il nome di Arlecchino (1671), costituita da una giacca e da un pantalone giallo chiaro ricoperto da triangoli di stoffa rossi e verdi; e di una terza (1858) che, a differenza della precedenti, mostrerebbe una certa eleganza. Composta da una giacca e da un pantalone a losanghe verdi, rossi, gialli e blu, si caratterizza per la presenza di bottoni di stoffa, di calze bianche, di una mezza maschera e di un paio di scarpe nere.

Con l’arrivo in Francia per opera del capo-comico Alberto Naselli, Arlecchino, ormai *Arlequin*, si raffina: considerato fino ad allora un servo sciocco, acquista, con il cambiamento dell’abito, una nuova malizia.

La caratteristica più evidente di Arlecchino è l’agilità, tanto che per poterlo rappresentare bisognava essere quasi atleti. Lo furono Domenico Biancolelli, il quale inventò l’inchino

⁷² Jean-François-Maurice-Arnauld, conosciuto con il nome di Maurice Sand (Parigi, 30 Giugno 1823 – Nohant-Vic, 4 Settembre 1889), era un illustratore e scrittore francese. Figlio del barone Casimir Dudevant e di George Sand, scrittrice femminista francese, Maurice ebbe l’opportunità di studiare presso Eugène Delacroix (artista francese del periodo romantico). Egli viene ricordato non solo per i suoi romanzi, ma anche per i suoi studi sulla Commedia dell’Arte: ne è esempio *Masque set bouffons (Comédie Italienne) texte et dessins* del 1860.

arlecchinesco e che morì per una congestione polmonare dopo una danza acrobatica; Visentini, che nella commedia *Il Convitato di Pietra* eseguiva un capitombolo in aria tenendo in mano un bicchiere pieno di vino; Tristano Martinelli, il quale faceva in aria un doppio salto mortale; Evaristo Gherardi, anche lui deceduto a seguito di un colpo violento sul capo, battuto sull'assito durante una catena di piroette. Tutti questi attori recitarono soprattutto in Francia, lasciando testimonianze di un *Arlequin* eclettico, acrobata, elegante nei movimenti e raffinato nelle parole: il protagonista di un gioco teatrale che voleva essere prima di tutto visivo.

2.6.2 *Pasquariel*

Se l'origine del personaggio di Pasquariello è certa, in quanto maschera napoletana, non si può affermare lo stesso sulla derivazione del suo nome, la quale è tuttora oggetto di diverse ipotesi. C'è chi in essa vede, infatti, il nome di Pasquino, spirito mordace e satirico dei romani, oppure il nome del suo interprete principale, Giovan Pietro Pasquarello, comico della Compagnia di "Zan Ganassa", il quale lo rappresentò più volte come servo sciocco e funambolo.

La maschera di Pasquariello o Pascarello fu introdotta in Francia intorno al 1685 con il nome di *Pasquariel* dall'attore Giuseppe Tortoriti, conosciuto anche con il soprannome di "Truonno" ("tuono"), poiché alla grande agilità fisica univa una voce profonda.

La caratteristica principale del personaggio è la strafottenza e la furbizia: abile in ogni situazione, egli costringe spesso Arlecchino a sposare la parente Colombina, ma si propone anche come ballerino o pittore quando l'amico esprime il desiderio di farsi fare un ritratto.

All'inizio portava una maschera nera con il naso lungo, tipica di tutti gli zanni, e indossava un abito bianco che, più tardi, fu cambiato in nero. Da quel momento, che pare coincida con l'ingresso in Francia, esso ebbe una giacca e dei calzoni di velluto nero, un colletto bianco, delle calze rosse e un cappello fatto a calotta. A testimonianza di questa evoluzione vi è una raffigurazione (*Diversarum Nationum Habitus*, 1596), nella quale è rappresentato con un abito bianco tipico degli zanni, e un'incisione del 1700, in cui è raffigurato con un paio di calzoni e di calze bianche.



Pasquariello o *Pasquariel* venne quasi sempre rappresentato con il suo carattere napoletano, sebbene nel periodo francese abbia subito una profonda evoluzione del carattere, diventando meno tollerante e più impertinente.

2.6.3 *Colombine*



Nata a Venezia, il suo nome compare, per la prima volta, nella Compagnia degl'Intronati Senesi nel 1530. A renderla, tuttavia, un personaggio a tutto tondo è l'attrice italiana Carla Biancolelli, moglie dell'attore della compagnia di *Molière*, *Pierre Lenoir*. L'esistenza di questa maschera, tuttavia, è testimoniata già nelle commedie di Plauto, dove sono presenti ancelle ciniche e adulatrici. Ella è la personificazione della schiava antica, sempre pronta a suggerire malizie alla propria padrona e sua complice nei segreti domestici e amorosi. Conservando questo carattere, Colombina diventerà la “servetta” della

Commedia dell'Arte, dove è spesso associata ad Arlecchino, di cui ne è la fidanzata.

All'inizio Colombina aveva il volto coperto da una mezza maschera e indossava un abito bianco, un grembiule verde e una cuffia tra i capelli. È stata raffigurata anche con una gonna celeste di pizzo e con un grembiule bianco ricamato, con una camicetta bianca con sopra un corpetto rosso. In una rappresentazione, avvenuta a Parigi nel 1695, Colombina avrà addirittura una giubba e una gonna a toppe colorate, tipica del suo compagno, assumendo il nome di “Arlecchina”. Con questa veste fu interpretata da Isabella Tessandri, da Isabella Biancolelli Franchini e da sua nipote Caterina. Tuttavia, nelle prime rappresentazioni francesi della Commedia dell'Arte non è presente il nome Colombina o *Colombine*, testimoniato, invece, alla fine del Seicento nel teatro parigino della *Comédie Italienne* e dal testo di Carlo Cantù *Cicalamento in canzonette ridicolose, o vero Trattato di matrimonio tra Buffetto, e Colombina comici* (1646).

Colombina è un personaggio grazioso e civettuolo: in ogni sua trasformazione rimane maliziosa e scaltra, come nel repertorio di Evaristo Gherardi e in quello di Carlo Goldoni. È Gherardi a parlare dell'adattamento di Colombina al gusto francese: questo si percepisce

bene nel repertorio de *Le Théâtre Italien* e, in particolare, nelle commedie di *Nolant De Fatonville Colombine Avocat Pour Et Contre* (1675) e *Colombine Femme Vengée* (1689), interpretate da *Madame Maillard* e *Mademoiselle Delisle*. Per la prima di queste due attrici, nel 1715, *Alain René Lesage* scrisse la commedia *Colombine Arlequin*.

2.6.4 *Scaramouche*

L'origine del nome Scaramuccia si rintraccia nella parola “piccola battaglia” e allude al carattere millantatorio e pauroso della maschera. Nata a Napoli, venne in seguito esportata in Francia dal comico italiano Tiberio Fiorilli, il quale, durante una rappresentazione francese nel 1640, diede vita al personaggio francese *Scaramouche*. Essa è comunque già presente in un dipinto del 1572, nel quale sono raffigurati alcuni cortigiani di Carlo IX nei costumi delle maschere della Commedia dell'Arte. Ebbe grande popolarità nell'Europa del Cinque/Seicento, in quanto erede, nelle parole e nelle gesta, della figura del Capitano, ma anche dei paladini, dei crociati e dei condottieri di ventura. Adattato alla corte di Francia, *Scaramouche* veniva rappresentato come un signorotto di contrade inesistenti, ricco e non meno nobile di Carlo Magno, ma ladro e furfante e, per questo, rivale dichiarato di *Polichinelle*.



E' merito di Tiberio Fiorilli, attore e maestro di *Molière*, aver reso questa maschera popolare in Francia, in quanto ne modificò il carattere originario: preferì, infatti, rappresentarlo con la chitarra piuttosto che con la spada e dotato di una arguzia e di una psicologia ancora più complessa dello Scaramuccia italiano. Fiorilli, inoltre, privandosi della maschera, era solito imbiancarsi il volto con la farina quando lo metteva in scena, poiché riteneva che la maschera dovesse essere il volto stesso dell'attore: si recitava meglio, secondo l'attore, con i gesti e con il volto, con le parole e con il silenzio.

Il suo costume era tutto nero: solo il cappello era ornato da una grande piuma variopinta.⁷³

⁷³ Un aneddoto riporta che *Molière*, una volta, riferendosi al cielo, abbia detto: “*Questa sera il cielo si è vestito come Scaramouche*”.

La maschera di *Scaramouche*, dopo Fiorilli, fu indossata nel 1680 da Giuseppe Tortoriti, già interprete di Pasquariello. Dopo di lui, nella compagnia di Riccoboni entrò Giacomo Raosini (1716), Bonaventura Benozzi (1732) e, infine, Daniele Gandini.

2.6.5 *Cinthio*



La maschera di *Cinthio*, conosciuta anche con il nome di Flavio, Orazio, Lelio, Leandro, rappresentava la figura del giovane innamorato. Per questo motivo, impose ai suoi interpreti un aspetto gradevole, una certa prestanza fisica e una voce suadente, in quanto costituiva l'ideale maschile. Il personaggio, seppure spesso caratterizzato da una nota di ridicolo e di comicità, vestiva secondo la moda del tempo. Da questo derivò che il giovane amoroso, più che una maschera, diventasse un carattere tipico dell'epoca.

Le sue doti erano quelle di un uomo colto, il quale dava prova di conoscere sonetti e

madrigali e di avere gusto musicale.

Come quasi tutte le maschere della Commedia dell'Arte, anche questa ebbe un grande successo in Francia, sia con il nome di Flavio sia con quello di *Cinthio*. Uno dei suoi più importanti interpreti fu Giovanni Battista Andreini, figlio di Francesco e di Isabella Andreini. Nato a Firenze nel 1579, rappresenterà il giovane amoroso soprattutto con il nome di Lelio, mentre Marco Romagnesi, considerato uno dei più affascinanti uomini del suo tempo, salirà sul palco con il nome di Orazio. Da questo momento, la figura del giovane innamorato sarà conosciuta soprattutto con il nome di Leandro, la quale, rispetto alle altre, avrà anche una vena di comicità per il destino poco felice.

Rappresentato come amante perfetto nelle compagnie italiane del Seicento e nelle commedie di *Molière*, acquistò, con la fine del secolo, una nota di ridicolo, in quanto, seppur figlio di un capitano e uccisore di Saraceni, non riesce a collezionare altro che rifiuti. Maldestro e permaloso, non ammette che altri abbiano più fortuna di lui e, per questo, mette mano più volte alla spada che nessuno, tuttavia, ha mai visto. Questo giovane

amoroso che per lungo tempo era stata la raffigurazione dell'ideale maschile, con il tempo va sempre di più assomigliando alla figura del Capitano, oggetto di diletto delle corti francesi.

2.6.6 Mezzetin

L'origine del nome Mezzettino potrebbe derivare dalla parola "boccale di mezza misura", a testimonianza della quale si ha una stampa francese del XVIII secolo, nella quale si rintraccia la scritta *Mezzettin Boccal*. Tuttavia, c'è anche chi vede in questo nome un riferimento ad una persona che, per raggiungere i propri obiettivi, si serve di piccoli e poveri mezzi.

La maschera, adottata soprattutto nel XVII secolo dall'attore Ottavio Onorati della Compagnia dei Confidenti, raggiunse il successo con il comico Angelo Costantini alla fine del Seicento, quando fu chiamato (1681) al Teatro Italiano di Parigi per lavorare con l'*Arlequin* Domenico Biancorelli.

Mezzettino, nato come variante di Arlecchino e di Brighella, si caratterizzò subito per il carattere più dinamico rispetto allo zanni tradizionale e per la spregiudicatezza in ogni situazione.

Il costume di Mezzettino, all'inizio, era quello tipico degli zanni: un vestito bianco che cascava morbido, un cappello a punta ed una maschera bruna. Costantini, una volta entrato nella *Troupe Italienne*, ne modificò sia il carattere che l'abito: lo caratterizzò, infatti, per un'astuzia quasi diabolica⁷⁴ e lo mise in scena a viso scoperto, con un abito a righe rosse e bianche, con i calzoni al ginocchio, con un colletto bianco, e con un paio di scarpe e calze bianche. Nell'ultima versione (1632), per adattarsi al gusto dell'epoca, i pantaloni vennero scorcianti.

Il successo di questo personaggio crebbe sia in Francia che fuori: nel 1697, quando fu chiuso il teatro italiano da Re Luigi XIV per una satira di Angelo Costantini, l'attore si trasferì alla corte di Augusto il Forte, elettore di Sassonia e Re di Polonia. È in questo



⁷⁴ Angelo Costantini, nelle commedie di *Molière*, interpretò anche il personaggio di Sganarello, il cui nome pare derivare da una storpiatura del verbo "disingannare". Questo personaggio, ispirato fortemente da quello di *Mezzetin*, ha il ruolo del marito ingannatore e ingannato, debole in ogni circostanza, orgoglioso e sciocco.

periodo che l'attore veronese cambiò il carattere burlone di *Mezzetin*, contaminandolo col personaggio di Scapino. Tuttavia, il repertorio delle commedie raccolte da Evaristo Gherardi, testimoniano la natura di questa maschera come variazione del carattere di Brighella.

2.6.7 *Isabelle*



Isabella come figura dell'innamorata nacque con Isabella Andreini, la quale mise in scena questo personaggio femminile che trae il proprio nome da quello dell'attrice. Nata a Padova nel 1562 da una famiglia modesta, a 16 anni sposò Francesco Andreini, inventore del personaggio di Capitan Spaventa di Valle Inferna. Isabella intraprese la carriera teatrale assieme al marito, iniziando a recitare nella Compagnia dei Comici Gelosi diretti da Flaminio Scala. Prima attrice della compagnia inventò la maschera di essa stessa.

All'inizio indossò un abito di raso rosa, decorato con preziosi pizzi alle maniche e alla scollatura e arricchito

da un cappellino con piume variopinte. Le caratteristiche umane del personaggio erano le stesse di Isabella Andreini: la maschera, infatti, era romantica e brillante allo stesso tempo. Recitò per un lungo periodo alla corte francese, in quanto la Compagnia dei Gelosi fu la prima ad insediarsi in Francia. Nel giugno del 1604, mentre stava tornando in Italia, morì, lasciando scritti, ma anche una nuova maschera alla Commedia dell'Arte.

Tra la fine del Seicento e l'inizio del Settecento, il personaggio di Isabella fu rappresentato da Isabella Biancolelli, figlia di Domenico, la quale entrò nella Compagnia degli Intronati a partire dal 1631.

È merito di Francesca Biancolelli, quarta figlia di Domenico e conosciuta anche con il nome di *Dominique*, mettere in scena un'Isabella più maliziosa, ardita e sicura di sé.

2.6.8 Pierrot



Pierrot trae la sua origine dalla maschera di Pedrolino, un “pagliaccio” bianco che già nel 1576 diede il nome alla Compagnia del Pedrolino, molto attiva presso la Corte di Ferrara. All’inizio Pedrolino fu interpretato da Giovanni Pellesini, la cui carriera si sviluppò sia in Italia che in Francia: nel 1613 quest’attore, all’età di 87 anni, recitava a Parigi. Le notizie della Compagnia di Pedrolino sono poche e frammentarie: consapevoli della sua esistenza, non abbiamo, però, una visione completa degli attori che ne facevano parte. L’unica testimonianza risale ad un documento dell’epoca nel quale ricaviamo che *gran parte della invernata in Firenze et di poi in Pisa, et doppo certe settimane essersene andata a Luccha di dove avrebbe voluto ritornare a Pisa una seconda volta se il commissario Capponi, per timore di certi scandoli, non glie l’avesse impedito*. Si può immaginare quali fossero gli “scandoli”: artisti della Commedia dell’Arte e, per questo, intemperanti in episodi forse eccessivi. Quello che è sicuro è l’importanza di Flaminio Scala all’interno della compagnia. Nel 1611, Francesco Andreini scrisse la presentazione a *Il Teatro delle Favole Rappresentative* di Flaminio Scala, con lo scopo di tramandare l’arte del capo-comico e per poter rappresentarsi come simbolo e portavoce degli attori comici. Questa pubblicazione è fondamentale, in quanto è una delle poche raccolte degli scenari teatrali, in cui non vengono raccontate soltanto le caratteristiche di singoli attori, ma viene data una visione generale degli spettacoli in cui essi recitavano. Negli scenari della raccolta di Flaminio Scala, la presenza di Pedrolino è attiva e frequente.

Sulla scena Pedrolino è di solito rappresentato come un servo fedele al servizio del proprio padrone che si fida totalmente di lui, ma anche come un uomo che si fa prendere dall’umorismo e che inganna gli altri solo per il piacere di farlo. I suoi intrighi hanno, però, sempre lo scopo di servire gli interessi del padrone.

Una xilografia del 1621 lo rappresenta con un vestito simile ad un altro zanni, ma non così sformato. Indossa, infatti, un abito bianco con il cappello dello stesso colore. *Molière*, notando il successo di questo personaggio, lo modificò in *Pierrot* che, come Pedrolino, è la personificazione del contadino francese. In Francia, al suo costume tradizionale vengono aggiunti enormi bottoni neri e il suo carattere diventa pigro, statico, lunare. *Pierrot*

commette *gaffes* che, seppure appaiono assurde, sono espressioni del buon senso. Non è sciocco e si accorge bene delle stravaganze del padrone, nelle quali commette errori in modo consapevole. Un particolare di questo aspetto lo troviamo in *La Critique de la Caluse des Femmes*, commedia di *Losme de Montchesnay* e contenuta nel *Recueil* di Evaristo Gherardi, nella quale *Pierrot* è il servo di un marito giovane, *Cintbio*.

Da quel momento *Pierrot* ebbe continue trasformazioni: in particolare, il comico francese *Jean Baptist Debureau* (1796-1846) eliminò ogni grossolanità, affinò ulteriormente il costume, infondendo a questo personaggio una sensibilità morbosa e un fondo costante di malinconia.

Conclusioni

La ragione del successo della Commedia dell'Arte risiede prima di tutto nel suo linguaggio, nelle sue modalità di comunicare concetti semplici e nell'utilizzo di forme espressive sempre diverse. Un comico dell'arte possedeva una grande gestualità ed era capace di inventare ed evocare nuovi suoni; i costumi colorati alludevano a contesti differenti, ad azioni pantomimiche che spesso sconfinavano nella danza. L'uso della maschera annunciava la tipologia del personaggio e del suo carattere. Elementi questi interpretabili facilmente da persone colte e non colte.

Il merito dei Comici dell'Arte fu, tuttavia, anche quello di aver liberato l'attore dai vincoli del testo scritto al fine di spingerlo verso una nuova scrittura scenica che nascesse dall'improvvisazione e dalla fantasia dell'attore, il quale metteva in scena se stesso, ma anche il suo corpo, la sua voce, la sua energia, il suo pensiero.

I Comici dell'Arte hanno consegnato alla storia del teatro un patrimonio nuovo e liberatorio, nel quale la comicità, l'ilarità, il sarcasmo, l'ironia sono gli elementi che attirano lo spettatore di ogni classe sociale. Questi attori hanno superato anche i condizionamenti di uno spazio scenico definito, in quanto, rispetto ai teatri tradizionali, la loro arte era rappresentabile ovunque, sia in spazi liberi sia in spazi chiusi. Innovativi perché indifferenti alle macchine teatrali, essi invitavano lo spettatore a immaginare, a partecipare e spesso cercavano di coinvolgerlo con le proprie satire nell'analisi oggettiva e disincantata della società a cui apparteneva.

Attualmente la Commedia dell'Arte mantiene inalterato il tono mordace e satirico con il quale affronta gli argomenti di vita quotidiana. Non è particolarmente rappresentata, ma trova comunque ampi squarci per sottolineare e denunciare un certo "perbenismo borghese", fatto di regole e convenzioni che essa rifiuta. Si riconosce oggi in attori poliedrici ed eclettici: ad esempio, Roberto Benigni e Dario Fo, attori e autore comici dei propri testi che prendono spunto dalla realtà per denunciare potere e sopraffazione; senza trascurare Totò e tutti quelli attori o quei comici che hanno fatto dell'improvvisazione un mezzo per dare vita alle perplessità e ai quesiti della gente comune, la quale reagisce al potere in modo semplice, ma non per questo meno importante. La Commedia dell'Arte, oltre a genere teatrale, è un movimento artistico, che ha inciso profondamente nella creatività di artisti: i pittori del Settecento e quelli contemporanei come Picasso, Guttuso, De Chirico hanno dedicato gran parte della propria produzione a personaggi della Commedia dell'Arte, in particolare ad Arlecchino.

Scopo di questo capitolo è stato quello di approfondire la presenza e l'influenza della tradizione italiana nel teatro francese, nel quale il comico italiano fu, senza dubbio, una tappa fondamentale nell'evoluzione della *Comédie Italienne* in *Comédie Française*. In questo quadro, infatti, le maschere italiane hanno dato vita ad una metamorfosi rispetto alla tradizione dell'Arte. Nella Francia del Settecento la comicità mimico-gestuale si spostò progressivamente verso una francesizzazione delle maschere e del loro linguaggio. Nessuna opera può introdurci meglio nell'universo comico francese del *Recueil* di Evaristo Gherardi

CAPITOLO III

La codifica TEI-XML

La scrittura rappresenta la sostanza linguistica, l'inchiostro è il supporto della manifestazione grafemica (vista come una forma), la pergamena è il supporto della disposizione dell'inchiostro (vista come una forma), i tratti fisici-chimici della pergamena sono il supporto delle sue qualità formali, e così di seguito.⁷⁵

⁷⁵ U. Eco, *I limiti dell'interpretazione*, Milano, Bompiani, 1994, pag. 184.

Introduzione

In questa terza parte illustrerò le procedure seguite per codificare la raccolta *Le Théâtre Italien* di Evaristo Gherardi. Sono state affrontate problematiche di diversa natura, alle quali si è cercato di fornire una soluzione, in termini di rappresentazione, per non allontanarsi troppo dall'edizione originale.

E' utile subito sottolineare che il lavoro svolto è stato pensato per essere inserito nella sezione "Teatro" della Piccola Biblioteca Digitale Romanza, a cui si è fatto riferimento nel primo capitolo.

Una volta identificati gli obiettivi da seguire (digitalizzazione e codifica) con le finalità principali di preservarlo, di analizzarlo e di distribuirlo in Rete per mezzo della Piccola Biblioteca Digitale Romanza, si è proceduto ad individuare gli strumenti con cui poter raggiungere gli scopi prefissati.

L'elaborazione digitale implica una corretta comprensione del testo sulla base di criteri metodologici fissati fin dall'inizio del lavoro: pertanto, essa è stata fondamentale, in quanto non ha precluso la libertà d'interpretazione e ha permesso di analizzare l'originale in modo dettagliato. In questo modo abbiamo garantito al lettore digitale la possibilità di comprendere il testo mediante opportune marcature, in quanto *i limiti all'interpretazione del messaggio sono posti, non dall'emittente, ma dall'esistenza di un nesso tra questo e il ricevente*⁷⁶. Sulla base di questo la riflessione si è ampliata alla "comunità interpretante", in modo tale da divenire luogo di conservazione e di dibattito, così come avviene dalla lettura delle pagine materiali di un'edizione cartacea. La codifica elettronica dell'opera di Evaristo Gherardi vuole, infatti, dare spazio alla comunità al fine di permettere una comprensione consapevole dell'edizione elettronica. Per raggiungere questo obiettivo abbiamo pensato ad una codifica che unisse i lettori, gli editori, la comunità sulla base di un programma condiviso, di principi omogenei, di definizioni scientifiche e culturali coerenti. Abbiamo, pertanto, definito criteri di analisi testuale, capaci di descrivere e, quindi, di codificare ogni aspetto del testo nel suo più piccolo dettaglio, anche interpretativo. In questo modo, sono stati identificati fattori interni ed esterni, linguistici, strutturali e semantici *che rendono il testo un unicum, una testimonianza originale e irripetibile, una tessera insostituibile nella lunga durata della sua tradizione, nella storia delle sue successive metamorfosi e trasposizioni.*⁷⁷ In particolare, nel momento in cui si è deciso di codificare il *Recueil* di Evaristo Gherardi abbiamo pensato alla possibilità che esso fosse

⁷⁶ U. Eco, *I limiti dell'Interpretazione*, Milano, Bompiani, 1994, pag. 186.

⁷⁷ Istituto Universitario di Lingue Moderne di Milano, *Internet e le muse. La rivoluzione digitale nella cultura umanistica*, 1997, cit. pag. 27.

rappresentato con le sue note e con le sue immagini senza manipolarlo nella forma e nel contenuto e riflettendo sul significato che esso avrebbe dovuto avere se collocato in Rete.

Sulla base di questo sono sorte le prime problematiche: ci siamo posti la domanda se era effettivamente possibile rappresentare il testo in modo esaustivo, conservando tutte le informazioni presenti, e, se nel caso lo fosse stato, quali erano le informazioni da codificare e quali, invece, non risultavano significative. Come testimonia la Storia questi sono problemi antichi. La Biblioteca di Alessandria (305 a.C.), voluta dal suo fondatore come strumento di potere da contrapporre alla Biblioteca Liceo, mise in evidenza modi diversi di concepire il sapere: se la prima, infatti, voleva un continuo aumento di libri come forma di conoscenza universale e, quindi, di potere, l'altra si organizzava sulla base di un programma intellettuale preciso, quello di Aristotele. Anche l'introduzione della stampa a caratteri mobili di Gutenberg (1450) evidenziò la necessità di rigore scientifico nella comprensione di quali testi stampare con i caratteri mobili e quali non ne avevano bisogno.

Questo breve accenno alla storia del testo inteso come strumento di sapere dimostra con chiarezza come sia complessa la rappresentazione di un documento, in quanto essa è chiamata a soddisfare due diverse esigenze: la riproduzione delle informazioni del testo originale e quella della sua interpretazione mediante la codifica e la successiva distribuzione.

Il momento attuale non è molto diverso da quello passato, poiché quando codifichiamo un'opera in formato digitale altro non è che quello che fecero un tempo gli Umanisti con le loro attività di traduzione e di normalizzazione dei testi, momento nel quale venne cercata una lingua dotta e comune (il latino), paragonabile alla attuale necessità di trovare un linguaggio comune e condivisibile (la TEI). Al carteggio degli umanisti si sostituisce oggi quello delle Biblioteche Digitali, nelle quali, così come nell'Umanesimo, si cerca di rispettare i concetti di universalità e di libertà di accesso. Quello che più conta in questo paragone è, tuttavia, il ruolo degli intellettuali e degli Enti preposti alla ricostruzione dei testi, alla conservazione della memoria storica e culturale, alla critica dei testi come forma di costruzione del pensiero e giudizio e come acquisizione del sapere.

3.1 La TEI

Come si è avuto modo di accennare nel secondo capitolo, la diffusione delle applicazioni computazionali nel campo umanistico ha implicato la definizione di schemi di codifica per rappresentare in modo digitale testi cartacei.

Uno degli schemi di codifica più importanti è la *Text Encoding Initiative*, un progetto di ricerca internazionale e oggi modello di codifica standard per rappresentare le informazioni testuali e per gestire i dati umanistici in formato elettronico di un testo.

Sviluppata a partire dal 1986, essa fu supportata da tre delle associazioni più importanti del settore dell'Informatica Umanistica e della Linguistica Computazionale: l'*Association for Computers and the Humanities* (ACH), l'*Association for Computational Linguistics* (ACL) e l'*Association for Literary and Linguistic Computing* (ALLC). In seguito, presero parte al progetto anche l'*U.S. National Endowment for the Humanities*, *Directorate General XIII of the Commission of the European Communities*, l'*Andrew W. Mellon Foundation*, e il *Social Science and Humanities Research Council of Canada*.

Gli obiettivi principali della TEI vennero esposti durante una conferenza programmatica svoltasi al *Vassar College* (New York) nel novembre del 1987, dove venne affermato che lo scopo della TEI era quello di *essere in grado di rappresentare le caratteristiche testuali necessarie per la ricerca; essere semplici, chiare e concrete; essere di semplice utilizzazione per i ricercatori senza il ricorso a software specializzati; permettere una definizione rigorosa e un'efficiente elaborazione dei testi; consentire estensioni definite dall'utente; essere conformi agli standard esistenti o in procinto di essere adottati. Affinché le Norme godessero della più ampia accoglienza, è stato importante assicurare che: il nucleo comune delle caratteristiche testuali fosse facilmente condiviso; le caratteristiche specialistiche supplementari fossero facili da aggiungere (o da rimuovere) da un testo; fossero possibili molteplici codifiche parallele della stessa caratteristica; la ricchezza della codifica potesse essere definita dall'utente, con una soglia minima molto bassa; fosse fornita un'adeguata documentazione del testo e della sua codifica.*⁷⁸ Nella stessa sede, fu garantita la piena indipendenza della TEI rispetto ad un software in particolare, in quanto standard di codifica universale in grado di rappresentare tutte le categorie di testi, ognuno con le proprie caratteristiche.

All'inizio la *Text Encoding Initiative* utilizzò lo *Standard Generalized Markup Language* (SGML) e, quindi, uno schema di codifica di tipo dichiarativo. Le specifiche di questa prima versione furono pubblicate nel 1991 con il titolo *Guidelines for Electronic Text Encoding and Interchange, TEI P1*. Nel 1992 si arrivò alla TEI P2, in quanto il linguaggio da poco sviluppato subì delle ulteriori revisioni; nel 1994 si ebbe poi la TEI P3.

⁷⁸ Si consulti il materiale al seguente indirizzo internet: http://www.tei-c.org/release/doc/tei-p4-lite/teiu5_it.html#Heading2

Nel 2000 i membri della *Text Encoding Initiative* costituirono il *TEI Consortium* (organizzazione internazionale, senza fine di lucro, con lo scopo di rendere permanente il progetto di ricerca), il quale ha sede ufficiale a Bergen (Norvegia), ma è ospitato presso la Brown University (Providence, Rhode Island, Stati Uniti), Oxford University (Oxford, Gran Bretagna), e l'University of Virginia (Charlottesville, Virginia, Stati Uniti). Il *TEI Consortium* è guidato da un Consiglio di Direzione per la parte amministrativa e da un Consiglio Elettivo per la parte della ricerca.

I gruppi di ricerca presenti in Nord-America e in Europa hanno pubblicato, nel giugno del 2002, una nuova versione dello schema di codifica, la TEI P4, che si distingue dalle precedenti versioni per la completa corrispondenza con il linguaggio di *markup* XML. Grazie a questa corrispondenza, la TEI può essere associata a documenti XML e, quindi, essere proiettata in Rete. Attualmente la TEI è arrivata alla versione P5 utilizzata per la codifica della raccolta di Evaristo Gherardi.

3.1.1 Lo schema di codifica TEI

La TEI offre un vocabolario standardizzato XML. Necessita, pertanto, di una DTD, i cui elementi sono contenuti nei documenti ufficiali delle *Guidelines*, scaricabili dal sito del Consorzio TEI.⁷⁹

I principi base dello schema di codifica TEI sono quelli di una codifica di tipo dichiarativo, capaci di descrivere le strutture morfo-sintattiche e semantiche del testo. La DTD TEI traduce e traspone i formalismi strutturali e funzionali presenti nel testo letterario nel rispetto del linguaggio di marcatura utilizzato. A tale proposito, Fabio Ciotti afferma che *il modello descrittivo dei testi che sottende la TEI è basato su una formalizzazione delle convenzioni nella produzione di documenti testuali che sono state definite a partire dalla diffusione della stampa, e che sono state codificate fino al punto di divenire un vero e proprio schema di argomentazione (la divisione in capitoli, parti, paragrafi, etc., è un tipico esempio di questo fenomeno di determinazione dialettica tra modelli del pensiero e schemi imposti dagli strumenti di produzione intellettuale).*⁸⁰ L'adozione di uno schema di codifica di tipo dichiarativo capace di identificare i rapporti strutturali e funzionali presenti nel testo comporta, come sottolinea Fabio Ciotti, *un apporto assolutamente soggettivo dello studioso-codificatore, che necessita di interpretare la funzione delle varie componenti*

⁷⁹ L'indirizzo da cui è possibile prelevare o consultare una copia delle direttive per la codifica è: <<http://www.tei-c.org/P4X/>>. Ne esiste anche un'edizione a stampa, per i tipi della University of Virginia Press (<<http://www.upress.virginia.edu/books/tei.html>>).

⁸⁰ Il documento è consultabile al seguente sito internet: http://www.bibliotecaitaliana.it/doc/tei_intro.php

*strutturali tipografiche, o manoscritte, per essere in grado di impiegare in modo adeguato i marcatori per gli elementi previsti nella DTD della TEI.*⁸¹

La DTD, tuttavia, non può prescindere da una preparazione filologica/umanistica del codificatore, poiché la TEI è un linguaggio di *markup* realizzato da umanisti per umanisti.

Gli elementi definiti dalla DTD TEI sono più di 500: per selezionarli è possibile, e consigliabile, scegliere tra i diversi modelli di codifica che la *Text Encoding Initiative* offre a seconda della tipologia del testo da codificare. Inoltre, lo schema TEI è modulare e personalizzabile.

La DTD è costituita da una serie di sotto DTD, definite *DTD Fragments*, le quali, se combinate tra loro, riproducono uno schema di codifica che ben si adatta alle caratteristiche del testo da codificare. I *DTD Fragments* definiscono i *Tag Sets*, gruppi diversi di tag, che si dividono in tre insiemi:

- i *Core Tag Set* : contengono l'insieme minimo degli elementi comuni ad ogni tipo di testo e, pertanto, devono essere inseriti nella DTD;
- i *Base Tag Set* : contengono i *Tag Set* specifici per ogni tipo di testo (prosa, poesia, dramma);
- gli *Additional DTD Tag Set* : contengono l'insieme di elementi utili a rappresentare le caratteristiche del testo.

Se questi elementi non sono sufficienti, la TEI permette di personalizzare la DTD standard: è possibile, infatti, modificando sia i tag sia le entità per mezzo degli *Auxiliary Tag Set*, i quali gestiscono particolari applicazioni (la produzione dello schema di codifica e la sua documentazione). A tale proposito, il *TEI Consortium*, al fine di generare lo schema di codifica, ha creato prima l'applicazione *Web Pizzeria Chef*⁸² e, in seguito, *Roma*, alla quale siamo ricorsi noi.

Il *TEI Consortium*, inoltre, *per rispondere al 90% delle esigenze del 90% della comunità di utenti della TEI*⁸³ e per rendere più semplice la realizzazione di testi in formato elettronico compatibili con l'intero schema, ha prodotto anche una versione ridotta e semplificata dello schema di codifica TEI. Questa versione, detta *TEI-Lite*, contiene un sottoinsieme degli elementi (circa 170) dell'intero sistema TEI.

⁸¹ Si consulti il documento al seguente sito internet: http://www.griseldaonline.it/informatica/manuale_parte6_3.htm

⁸³ Si consulti il documento al seguente sito internet: http://www.bibliotecaitaliana.it/doc/tei_intro.php

3.1.2 La versione TEI scelta

Per il nostro progetto si è adottato la versione TEI P5, rilasciata nel 2007, in quanto risponde meglio all'esigenza di costruire e gestire un *corpus* testuale per i nuovi elementi introdotti (ne sono esempio, “xml:lang” e “xml:id”, attributi del W3C). Seppure questi cambiamenti abbiano avvicinato ulteriormente la TEI all'XML, la versione TEI P5 non è ancora stata completata.

Dal momento che si è reso necessario scegliere un modulo di codifica in particolare, prima di iniziare la codifica del testo originale si è analizzato il testo nella sua struttura basilare, riflettendo, in particolare, su quali caratteristiche si intendessero marcare. Trattandosi, infatti, di una raccolta di opere teatrali che non mirasse alla sola conservazione, ma anche alla sua diffusione, si è da subito abbandonata l'idea di utilizzare uno schema di codifica ridotto (la *TEI-Lite*), privilegiando piuttosto approfonditi criteri di codifica. Questo è stato possibile sia mediante la personalizzazione degli elementi della DTD sia attraverso l'ampia scelta di moduli di codifica. Scegliere il modulo di codifica ha implicato comunque una serie di riflessioni, come l'individuazione delle caratteristiche del testo che si intendevano marcare, la natura del testo (prosa, poesia, drammatico, storico), le caratteristiche stilistiche e l'apparato critico.

3.2 Gli strumenti utilizzati

Al fine di concretizzare questa serie di riflessioni si è dovuto ricorrere ad altre due riflessioni prima di iniziare la codifica vera e propria, ovvero come produrre uno schema di codifica e che tipo di strumento software utilizzare per codificare il testo.

3.2.1 Produzione dello schema di codifica

Come accennato precedentemente, per produrre lo schema di codifica si è utilizzato il software online Roma, visualizzabile all'indirizzo internet <http://www.tei-c.org/Roma/>, realizzato da Arno Mittelbach e da Sebastian Rahtz, entrambi docenti dell'Università di Oxford.



These pages will help you design your own TEI validator, as a DTD, RELAXNG or W3C Schema.

Create a new or upload existing customization

- Build schema (Create a new customisation by adding elements and modules to the smallest recommended schema)
- Reduce schema (Create a new customization by removing elements and modules from the largest possible schema)
- Create customization from template
- Open existing customization

Home Page di Roma

Su questa pagina abbiamo selezionato la voce *build schema*, in modo tale da settare i parametri inseriti. La schermata che abbiamo ottenuto è stata la seguente.

Set your parameters

Title	<input type="text" value="My TEI Extension"/>
Filename	<input type="text" value="myTEI"/>
Namespace for new elements	<input type="text" value="http://www.example.org/ns/nonTEI"/>
Prefix for TEI pattern names in schema	<input type="text" value="tei_"/>
Language	<input checked="" type="radio"/> English, <input type="radio"/> Deutsch, <input type="radio"/> Italiano, <input type="radio"/> Español, <input type="radio"/> Français, <input type="radio"/> Portugues, <input type="radio"/> Russian, <input type="radio"/> Svenska, <input type="radio"/> 日本語, <input type="radio"/> 中文
Author name	<input type="text" value="generated by Roma 3.12"/>
Description	<input type="text" value="My TEI Customization starts with modules tei, core, textstructure and header"/>

Una volta fatto questo abbiamo scelto il modello TEI P5 che più soddisfaceva alle nostre esigenze. Trattandosi di un *corpus* teatrale abbiamo selezionato il modulo *drama* e soltanto dopo abbiamo generato la rispettiva DTD.

Module name	A short description	Changes
add analysis	Simple analytic mechanisms	
add certainty	Certainty and uncertainty	
add core	Elements common to all TEI documents	
add corpus	Corpus texts	
add dictionaries	Dictionaries	
add drama	Performance texts	
add figures	Tables, formulæ, and figures	
add gajji	Character and glyph documentation	
add header	The TEI Header	
add iso-fs	Feature structures	
add linking	Linking, segmentation and alignment	
add msdescription	Manuscript Description	
add namesdates	Names and dates	
add nets	Graphs, networks, and trees	
add spoken	Transcribed Speech	

Elenco Moduli previsti dalla TEI

TEI Roma: generating validators for the TEI

You are currently working on **My TEI Extension**

Time to give you a schema

Which format do you prefer?

Generate

Progress: 100%

Roma was written by Arno Mittelbach and is maintained by Sebastian Rahtz. Sanity check written by Ioan Bernevig. Documentation language en. Please direct queries to the [TEI @ Oxford](#) project. This is Roma version 3.12, last updated 2009-07-13. Using TEI P5 version 1.8.0. Last updated on November 5th 2010.

[If your browser did not redirect you, please click here](#)

Dal momento che Roma offre al codificatore anche la possibilità di generare la documentazione relativa alla DTD in formato HTML e di salvare in formato XML le informazioni relative al materiale salvato, abbiamo scelto di generare anche questo materiale.

TEI Roma: generating validators for the TEI

You are currently working on **My TEI Extension**

Documentation?

Getting some nice documentation

Which output would you prefer?

Generate

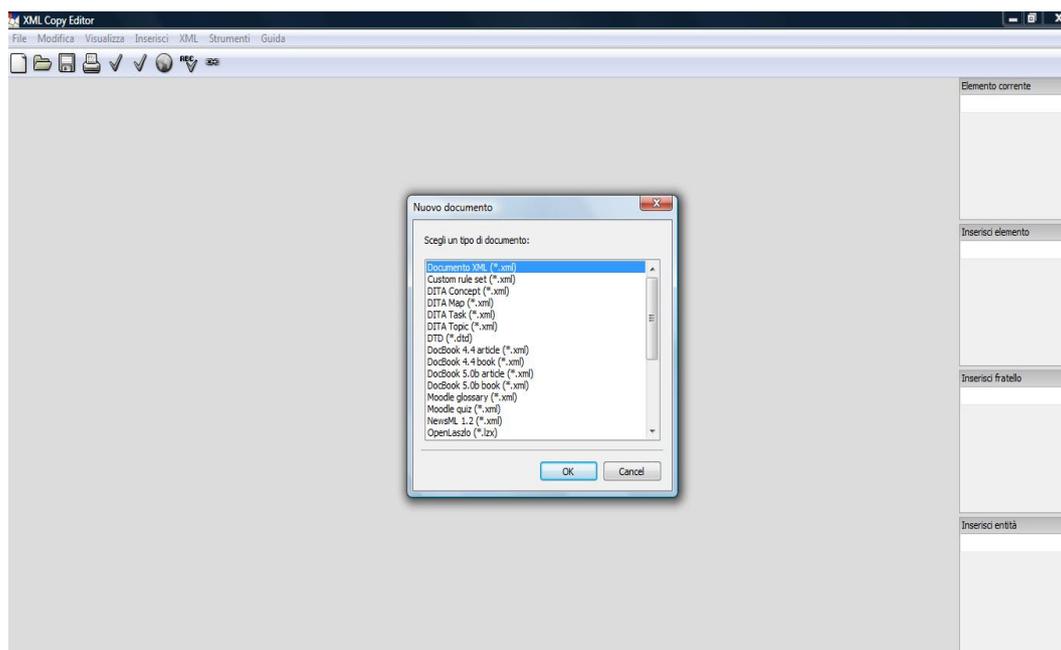
3.2.2 Il software *XML Copy Editor*

Prima di partire con la codifica effettiva del testo è stato indispensabile, inoltre, selezionare e provare gli strumenti software capaci di associare i tag al testo da codificare.

Dopo un'attenta valutazione la nostra scelta si è concentrata su due programmi: *Oxygen* e *XML Copy Editor*. Dal momento che il primo è un programma a pagamento, che concede una licenza *trial*⁸⁴ soltanto per la durata di un mese, la scelta definitiva è caduta sul secondo programma, in quanto editor gratuito.

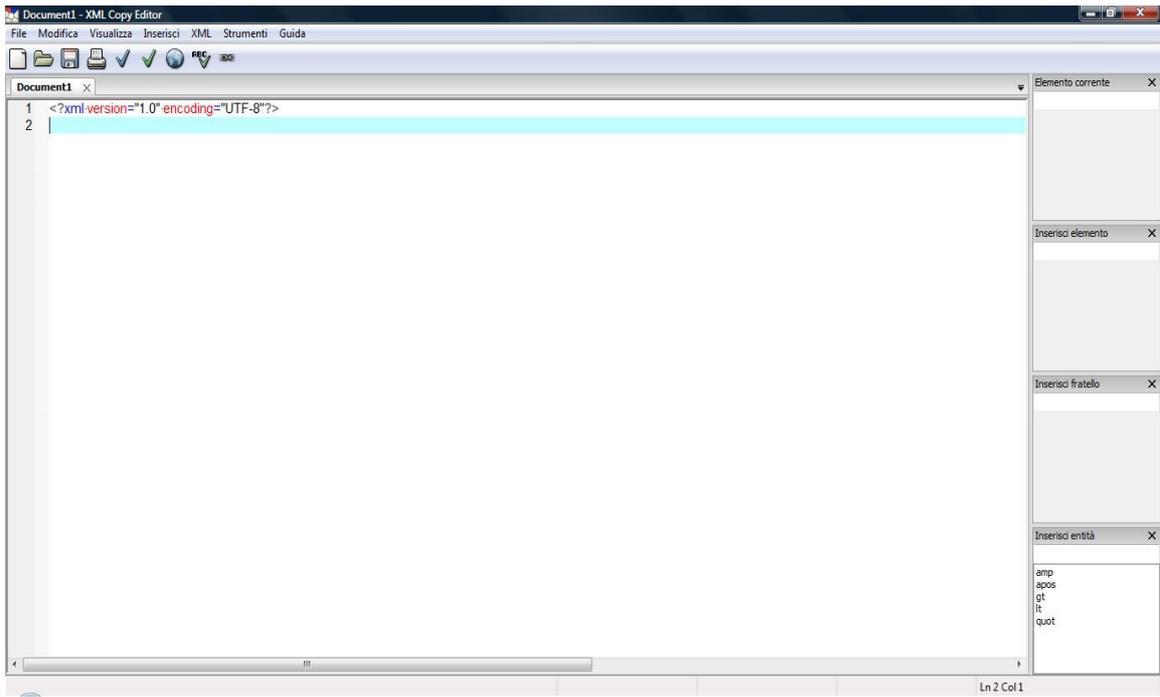
XML Copy Editor è uno dei più completi editor XML al momento disponibili, in quanto le sue principali caratteristiche sono: supporto file XSLT; visualizzazione colorata della sintassi; controllo della sintassi inserita; auto completamento dei tag.

Abbiamo, dunque, adoperato questo software per codificare il testo originale, per verificare la parte digitalizzata, per validare sintatticamente i file XML e, nella fase più avanzata del lavoro, per preparare i fogli di stile XSLT.



Schermata Iniziale di XML Copy Editor

⁸⁴ Il termine *trial* indica le versioni di programmi distribuiti sotto la formula “prova”, al fine di permettere all’utente di provare e valutare le potenzialità del software prima di procedere all’acquisto.



Inizializzazione XML Copy Editor

3.3 La preparazione del testo elettronico

Per codificare al meglio il *corpus* teatrale abbiamo seguito una serie di fasi. Dopo aver generato lo schema di codifica mediante il software gratuito Roma e dopo aver provato *XML Copy Editor*, abbiamo preparato la struttura gerarchica dell'intero testo.

Sono state individuate a tale scopo sia le informazioni interne sia quelle esterne. Le prime si riferiscono al corpo del documento e, quindi, ai dati testuali: ne sono esempio le informazioni di natura sintetica (titoli, capitoli, paragrafi), analitica (corpo del documento), ipertestuale (citazioni, note, epigrafi). Le seconde si basano, invece, sulle informazioni che ruotano intorno al testo vero e proprio (autore, luogo, anno di pubblicazione). Grazie a questa divisione iniziale è stato possibile rappresentare l'originale in MRF.

Quando una sequenza di caratteri in MRF diviene un testo elettronico, abbiamo di fronte un prodotto complesso che, per sua natura, presenta davvero un valore aggiunto per la comunità.

3.3.1 Acquisizione del testo originale

La raccolta *Le Théâtre Italien* di Evaristo Gherardi è attualmente proprietà della *Bibliothèque Nationale de France*, la quale possiede i sei volumi originali cartacei editi da J.-B. Cusson e P. Wiite a Parigi nel 1700. Tuttavia, tre di questi sono in libero accesso su Gallica⁸⁵, la Biblioteca Digitale della *Bibliothèque Nationale de France*.

Trattandosi di una Tesi interdisciplinare ho potuto fare affidamento sulla Dottoressa Barbara Sommovigo, responsabile della Piccola Biblioteca Digitale Romanza e Docente presso il Dipartimento di Lingue e Letterature Romanze, la quale ha ordinato e acquistato i formati PDF dei tre volumi non disponibili su Gallica. Tuttavia, per poter passare alla fase seguente, ovvero alla normalizzazione del testo, è stato necessario, una volta ottenuti i PDF, ricorrere a software di riconoscimento ottico dei caratteri (OCR)⁸⁶, in modo tale da poter avere le opere nel formato “.doc”. Anche di questo aspetto si è occupata la Dottoressa Sommovigo, in quanto è stata lei a curare la fase successiva della normalizzazione dei testi. Ella è ricorsa al programma *Fine Reader*, sebbene abbia ottenuto scarsi risultati a causa sia della qualità della stampa (essa presenta, infatti, ombre e macchie

⁸⁵ Si consulti il seguente sito *Internet* www.gallica.bnf.fr all'indirizzo

http://gallica.bnf.fr/Search?ArianeWireIndex=index&lang=EN&q=gherardi+theatre+italien&p=1&f_century=17

⁸⁶ Col nome di OCR (Optical Character Recognition, traduzione: riconoscimento ottico dei caratteri) si indica una categoria di programmi la cui funzione è quella di trasformare una immagine grafica, generalmente acquisita con uno scanner, contenente del testo in un file di testo modificabile col computer.

sullo sfondo) sia per la lingua antica che non permette di avere grande precisione (caratteri che si confondono e punteggiatura non riconosciuta).

3.3.2 La normalizzazione del testo

Dal momento che la normalizzazione del testo è quel processo per mezzo del quale un testo viene trasformato al fine di renderlo costante e coerente, esso si è reso necessario prima di passare alla codifica vera e propria.

È stata conservata l'ortografia dell'edizione originale, oscillazioni grafiche comprese, ad eccezione dei casi in cui c'è stata la necessità di uniformare l'uso del *trait d'union* secondo l'utilizzo moderno e di razionalizzare l'impiego degli accenti (distinzione tra a e à, ou e où etc.) e degli apostrofi (distinzione tra ma e m'a, ny e n'y, etc.).

È stata rispettata anche la punteggiatura, eccezion fatta in alcuni casi, come, ad esempio, quando ad una domanda esplicita non segue il punto interrogativo. Questo aspetto della normalizzazione ha richiesto particolare attenzione, in quanto per le opere teatrali spesso non è possibile risalire ai manoscritti originali. Il copione - o ciò che più si avvicina all'idea moderna di "copione" - veniva distribuito alla *troupe* e letto nei "salotti". Per questo motivo, oltre al veloce decadimento di queste copie d'uso e alla scomparsa dei manoscritti originali, contribuì anche la tendenza a sostituirli con le prime edizioni a stampa, che seguivano a breve scadenza le prime serie di rappresentazioni. Trattandosi di opere teatrali, i diversi segni della punteggiatura corrispondevano ad altrettante indicazioni di regia come, ad esempio, le pause da rispettare nel corso della rappresentazione.

Una volta compiute questa serie di operazioni preliminari si è passati alla fase successiva, ovvero alla codifica vera e propria.

3.4 La struttura di un documento TEI

Una volta terminate queste fasi iniziali di analisi, si è proceduto a codificare il *corpus*. Tuttavia, prima di entrare nel dettaglio della codifica TEI-XML è opportuno soffermarsi sulla struttura che devono avere tali documenti.

Ogni documento TEI presenta l'elemento radice `<TEI>`, dal quale discendono i due elementi figli `<teiHeader>` e `<text>` che dividono il testo in due parti.

Nella `<TEI Header>` si trovano i metadati (informazioni bibliografiche ed editoriali relative al documento TEI e alla sua fonte cartacea), per mezzo dei quali è possibile rappresentare il frontespizio dell'edizione digitale. Esso contiene quattro elementi: `<fileDesc>`; `<encodingDesc>`; `<profileDesc>`; `<revisionDesc>`.

L'elemento `<text>` riproduce, invece, il contenuto testuale dell'opera originale e contiene, al suo interno, altri quattro elementi: `<front>`; `<group>`; `<body>`; `<back>`.

Dal momento che la raccolta da codificare può essere classificata come testo composito, noi abbiamo seguito una struttura di questo tipo:

```
<TEI>
<teiHeader> [ intestazione del tomo ] </teiHeader>
<text>
<front> [ front matter del tomo ] </front>
<group>
<text>
<front> [ front matter della prima commedia ] </front>
<body> [ body della prima commedia ] </body>
<back> [ back matter della prima commedia ] </back>
</text>
<text>
<front> [ front matter della seconda commedia ] </front>
<body> [ body della seconda commedia ] </body>
<back> [ back matter della seconda commedia ] </back>
</text>
...
[ tutte le altre commedie ]
...
</group>
<back> [ back matter del tomo ] </back>
</text>
</TEI>
```

3.5 La *TEI Header*

Dopo aver accennato alla struttura di un documento *TEI*, è opportuno soffermarsi sul ruolo della `<teiHeader>`, la quale fornisce le informazioni relative alla bibliografia dell'edizione elettronica e dell'originale cartacea. Essa risulta essenziale, in quanto permette ai documenti digitali di essere *Tei Conformed*. Dentro la `<teiHeader>` si trovano, infatti, le informazioni relative alla fonte, alla tipologia di codifica adottata, al responsabile della codifica, alla documentazione, alla conservazione del testo codificato, alla revisione del testo, alla sua catalogazione e indicizzazione e sulla sua provenienza.

Come si è già accennato, la descrizione di questi dati avviene per mezzo di quattro parti.

3.5.1 La descrizione del file

Il primo elemento obbligatorio della `<teiHeader>` è `<fileDesc>`. Esso contiene informazioni analoghe a quelle contenute nel frontespizio di un testo cartaceo. Presenta sette elementi figli, che forniscono una descrizione bibliografica completa del file e della sua fonte: troviamo, infatti, il `<titleStmt>`, l'`<editionStmt>`, l'`<extent>`, la `<publicationStmt>`, l'`<seriesStmt>`, il `<notesStmt>`, il `<sourceDesc>`. Di questi soltanto tre sono obbligatori: il `<titleStmt>`, la `<publicationStmt>` e il `<sourceDesc>`.

Il primo elemento obbligatorio contenuto all'interno di `<fileDesc>` è `<titleStmt>`, il quale contiene il titolo del documento elettronico. All'interno di questo le informazioni sono codificate con i tag `<title>`, il quale riporta il titolo dell'opera (nel nostro caso, è stato riportato il titolo originario della raccolta insieme alla specifica di "edizione elettronica"), e il tag `<author>`, il quale presenta il nome dell'autore dell'opera. Per distinguere il nome dal cognome viene utilizzato il tag `<name>`, seguito dall'attributo `<type>`, il quale marca questa distinzione con due ulteriori elementi `<surname>` e `<forename>`. L'elemento `<respStmt>` fornisce, invece, le informazioni riguardanti il responsabile dell'edizione digitale. Il ruolo e il nome della persona sono indicati per mezzo dei tag `<resp>` e `<name>`: il primo elemento descrive il compito del codificatore, mentre il secondo fornisce il nome.

Mostriamo di seguito una porzione di codice:

```
<titleStmt>
<title>
<hi rend="corsivo">
Le Théâtre Italien:
</hi>
```

```

edizione elettronica
</title>
<author>
<name type="forename">Evaristo</name>
<name type="surname">Gherardi</name>
</author>
<respStmt>
<resp>
<lb></lb>Digitalizzazione e Codifica elettronica a cura di:
</resp>
<name>Roberto Rosselli, Professore di Codifica dei Testi;
<lb></lb>
Barbara Sommovigo, Ricercatrice di Letteratura Francese;
<lb></lb>
Laura Giannini, Laureanda in Informatica Umanistica</name>
</respStmt>
</titleStmt>

```

Il secondo elemento figlio di `<fileDesc>` è `<editionStmt>`, il quale fornisce le informazioni relative all'edizione digitale di un testo.

Il numero e la data dell'edizione sono codificati con il tag `<edition>`, mentre la data è inserita all'interno del marcatore `<date>`.

Mostriamo di seguito una porzione di codice:

```

<editionStmt>
<edition>Prima edizione elettronica
<date>2011</date>
</edition>
</editionStmt>

```

Il secondo elemento obbligatorio di `<fileDesc>` è `<publicationStmt>`, il quale fornisce le informazioni relative alla pubblicazione e alla distribuzione del testo in versione digitale. Esso contiene una serie di sotto – elementi. Il primo è `<publisher>`, il quale marca il nome della persona o dell'Ente responsabile della pubblicazione del testo in versione digitale. In questo caso il responsabile della pubblicazione è Dipartimento di Lingue e Letterature Romanze in collaborazione con il Corso di Laurea in Informatica Umanistica, i cui referenti sono la Dottoressa Barbara Sommovigo e il Professore Roberto Rosselli Del Turco. L'elemento `<availability>` contiene le informazioni relative alla disponibilità del testo. Il marcatore `<date>`, con cui si conclude l'elemento `<publicationStmt>`, contiene la data di pubblicazione del documento elettronico.

Mostriamo di seguito una porzione di codice:

```
<publicationStmt>
<publisher>Università degli studi di Pisa;<lb></lb>
Facoltà di Lettere e Filosofia;<lb></lb>
Corso di Studi in Informatica Umanistica;<lb></lb>
Corso di Codifica di Testi</publisher>
<availability status="restricted">
<p>Questo progetto è protetto dal diritto d'autore</p>
</availability>
<date>2011</date>
</publicationStmt>
```

La DTD TEI prevede anche l'elemento obbligatorio *<sourceDesc>*, il quale contiene la descrizione della fonte. Il riferimento bibliografico all'originale è organizzato in modo gerarchico, in quanto gli elementi sono inseriti all'interno del marcatore *<biblFull>*. All'interno di esso compare per primo l'elemento *<titleStmt>*, i due elementi *<title>* caratterizzati dall'attributo *<type>* con valore *<main>* e *<sub>*, i quali distinguono il titolo dal sottotitolo. Il tag *<extent>* fornisce una descrizione fisica dell'opera originale (dimensioni e numero delle pagine) e l'elemento *<seriesStmt>*, con cui si conclude la descrizione bibliografica della fonte, descrive il titolo della collana di appartenenza dell'opera originale.

Mostriamo di seguito una porzione di codice:

```
<sourceDesc>
<biblFull>
<titleStmt>
<title type="main">Le Théâtre Italien
</title>
<title type="sub">Tome II</title>
<author>de Gherardi</author>
</titleStmt>
<editionStmt>
<edition>edizione</edition>
</editionStmt>
<extent>556 pagine;</extent>
<publicationStmt>
<publisher></publisher>
<pubPlace>Parigi</pubPlace>
<date>2005</date>
</publicationStmt>
```

```

<seriesStmt>
<title>Le Théâtre Italien</title>
</seriesStmt>
</biblFull>
</sourceDesc>
</fileDesc>

```

3.5.2 Descrizione della codifica

Dopo l'elemento *<fileDesc>* la TEI prevede l'elemento *<encodingDesc>*, il quale descrive le metodologie e i principi editoriali utilizzati per la codifica elettronica del testo. Tale marcatore non è obbligatorio, ma il suo uso è consigliato poiché espone le linee guida alla base del progetto di codifica.

La descrizione della codifica si apre con il tag *<projectDesc>*, il quale contiene le informazioni riguardanti lo scopo per cui il testo cartaceo è stato codificato. All'interno del marcatore *<editorialDecl>* è contenuta, invece, la spiegazione dei principi editoriali seguiti per codificare il testo: essi riguardano sia la normalizzazione del testo sia la correzione apportata alla versione digitale in base al testo di riferimento. Nell'elemento *<refsDecl>* sono contenute poi le informazioni dello schema di riferimento standard introdotto nella codifica. Nel nostro caso illustra il ruolo dell'attributo *<n>* presente sia nel tag *<div>* sia in *<pb>*.

Mostriamo di seguito una porzione di codice:

```

<encodingDesc>
<projectDesc>
<p>Per quanto riguarda la visualizzazione del testo in rete, ho cercato di inserire delle scelte visive che mettessero in evidenza gli elementi fondamentali della marcatura. Sono state inserite le immagini presenti in ogni commedia della raccolta, ricollocate nella loro posizione originaria e sono state marcate in particolare le indicazioni di scena. </p>
</projectDesc>
<editorialDecl>
<p>Al momento di sviluppare i criteri di codifica ho scelto di privilegiare lo statuto narrativo del testo. Lo scopo principale è stato quello di mettere in rilievo gli aspetti tipici di testi teatrali, come le indicazioni di scena, le quali, anche nel caso non fossero esplicite, sono marcate con una generica indicazione. Si è cercato in tal modo di predisporre una fruizione del testo codificato che ne mettesse in risalto il funzionamento narrativo,

```

senza trascurare la ricchezza espressiva e discorsiva dell'opera. La codifica del testo nel formato XML-TEI ha permesso la creazione di un ipertesto che rendesse possibile la fruizione delle parti desiderate.</p>

</editorialDecl>

<refsDecl>

<p>L'attributo N indica il numero della commedia all'interno della raccolta e quello delle scene e i cambi di pagina sono codificati mediante "pb" che, accompagnati dall'attributo n, indicano il valore corrispondente al numero della pagina nel testo cartaceo di riferimento. Inoltre, per mantenere i dialoghi nella loro forma originaria è stato utilizzato l'elemento "lb".</p>

</refsDecl>

</encodingDesc>

3.5.3 Descrizione del profilo

L'elemento <profileDesc> compare subito dopo <encodingDesc> e ha lo scopo di fornire una serie di informazioni non bibliografiche che caratterizzano, sotto vari aspetti descrittivi, un testo. Esso possiede, infatti, due ulteriori elementi opzionali: <creation>, il quale riporta l'anno dell'edizione digitale, e <langUsage>, che specifica le lingue presenti nel testo (francese, italiano, latino). Ad esso è associato l'attributo <ident> che specifica il valore secondo lo standard ISO 639 (sono fornite, infatti, le due lettere iniziali in modo da identificare il linguaggio utilizzato).

Mostriamo di seguito una porzione di codice:

```
<profileDesc>
  <creation>
    <date when="xx-xx-2010">2011</date>
  </creation>
  <langUsage>
    <language ident="FR">Francese</language>
    <language ident="IT">Italiano</language>
    <language ident="LA">Latino</language>
  </langUsage>
</textClass>
</profileDesc>
```

3.5.4 Descrizione della revisione

L'ultimo elemento del `<teiHeader>` è `<revisionDesc>`, il quale ha lo scopo di fornire informazioni relative alle revisioni compiute rispetto al testo originale. Ogni modifica compiuta sulla versione elettronica è racchiusa all'interno dell'elemento `<change>`, nel quale sono forniti sia la data (`<date>`), sia il nome ed il ruolo del responsabile della modifica (`<respStmt>`), e ancora una breve descrizione dell'intervento (`<desc>`). Tuttavia, è opportuno precisare che la digitalizzazione del *Le Théâtre Italien* non ha comportato dei veri cambiamenti, in quanto si è cercato di rispettare la natura originale dell'opera. Con questi elementi la `<teiHeader>` viene chiusa.

Mostriamo di seguito una porzione di codice:

```
<revisionDesc>
  <change>
    <date when="xx-xx-2010">2011</date>
    <name>Barbara Sommovigo</name>
    <desc>Correzione in base al testo di riferimento:
    <bibl>
      <author>Evaristo Gherardi</author>
      <title>Le Théâtre Italien</title>
      <publisher></publisher>
      <date>2011</date>
    </bibl>
    </desc>
  </change>
  <change>
    <date when="xx-xx-2011">2011</date>
    <name>Roberto Rosselli Del Turco</name>
    <desc>Supervisione Tecnica</desc>
  </change>
  <change>
    <date when="xx-xx-2011">2011</date>
    <name>Laura Giannini</name>
    <desc>Codifica XML del testo de Le Théâtre Italien de Gherardi Tome
    II</desc>
  </change>
</revisionDesc>
```

3.6 Il *text*

Come si è già avuto modo di accennare, l'elemento `<text>` si suddivide in ulteriori quattro elementi: `<front>`, `<group>`, `<body>` e `<back>`.

Vediamo nel dettaglio il compito di ogni elemento.

3.6.1 Il `<front>`

L'elemento `<front>` è un elemento opzionale, contenente tutti i materiali di tipo avantestuale che precedono il corpo del testo nelle edizioni a stampa (pagina del titolo, frontespizio, introduzioni o prefazioni) ed è posizionato all'inizio del corpo del testo. Si trova subito dopo l'elemento `<text>`, di cui è figlio.

Grazie al `<front>` è stato possibile inserire le informazioni riguardanti il frontespizio dell'opera originale.

Mostriamo di seguito una porzione di codice:

```
<text>
<front>
<head type="title">
<hi rend="corsivo">
Le<lb></lb>
Théâtre <lb></lb>
Italian<lb></lb>
</hi>
</head>
<head type="subtitle">
De <lb></lb>
Gherardi</head>
<head type="subtitle">TOME I</head>
</front>
```

3.6.2 Il `<group>`

Il secondo elemento figlio opzionale di `<text>` è `<group>`, il quale raggruppa una sequenza di testi o gruppi appartenenti alla medesima collezione e che contiene, a sua volta, gli elementi `<front>`, `<body>`, e `<back>`.

All'interno del `<front>`, presente in `<group>`, sono stati inseriti elementi come `<castList>`, il quale apre l'elenco dei personaggi presenti in una determinata commedia, seguito poi

dall'elemento `<castItem>`, che fornisce una singola voce all'interno della lista dei personaggi. Per essere più precisi, si è ricorsi poi all'elemento `<role>`, il quale fornisce il nome di ogni singolo personaggio, mentre `<roleDesc>` permette di identificarne il ruolo all'interno della commedia.

All'interno di questo `<front>` è stato inserito il marcatore `<set>`, il quale contiene la descrizione dell'ambientazione, della scena, dell'azione di un'opera teatrale, tradizionalmente inserita nel peritesto iniziale.

Là dove questi marcatori non sono presenti è perché nell'edizione originale erano assenti.

3.6.3 Il `<body>`

L'elemento obbligatorio `<body>`, secondo e più importante figlio di `<text>`, contiene l'intero corpo di un singolo testo, con l'esclusione di ogni elemento preliminare o di appendice, contenuti invece all'interno dei tag `<front>` e `<back>`. Come il `<front>`, anche questo elemento è contenuto all'interno di `<group>`, in modo tale da rappresentare i contenuti testuali di ogni commedia.

All'interno di `<body>` sono presenti una serie di suddivisioni, necessarie a definire la struttura interna del testo da codificare: esse, presenti anche nell'edizione originale mediante opportuni segni tipografici (spaziature, pagine bianche, stili di carattere), rappresentano le partizioni strutturali (atto o scena nel nostro caso). Per rappresentare queste divisioni testuali è stato utilizzato l'elemento `<div>`, al cui interno sono presenti le varie diramazioni che rappresentano le diverse scene presenti di una determinata commedia. All'interno del `<div>` sono, inoltre, presenti i seguenti elementi: `<head>`, `<sp>`; `<speaker>`; `<p>` e `<stage>`.

Abbiamo, pertanto, scelto di codificare tutte le informazioni che caratterizzassero il testo nella sua natura. Trattandosi, tuttavia, di un testo drammatico abbiamo posto la principale attenzione alle indicazioni di scena, che abbiamo diversificato a seconda della loro funzione nel rispetto delle norme TEI. Le *Guidelines* offrono, infatti, al codificatore la possibilità di utilizzare diversi attributi.

Guardiamo ora nel dettaglio i singoli elementi.

3.6.3.1 L'elemento `<head>`

Come è possibile osservare nella riproduzione grafica della pagina, ogni volume si apre con i titoli in corsivo. Nella codifica TEI titoli ed intestazioni dei blocchi di tipo `<div#>`

vengono marcati grazie all'elemento `<head>`, dove mediante l'attributo *type* si indica il tipo di titolo. Possiamo avere, infatti sia `<head type="title">` sia `<head type="subtitle">` per distinguere il titolo principale da un sotto-titolo.

Mostriamo di seguito una porzione di codice:

```
<head type="title">La Cause des Femmes</head>
<head type="subtitle">
SCENES FRANÇOISES
<lb></lb>
DE
<lb></lb>
LA CAUSE
<lb></lb>
DES FEMMES.
<lb></lb>
</head>
```

3.6.3.2 L'elemento `<sp>`

Nella codifica TEI-XML, per codificare una singola battuta siamo ricorsi all'elemento `<sp>`, il quale permette di marcare l'intervento di ogni personaggio. Tale elemento può contenere ulteriori tre elementi: `<speaker>`, `<stage>`, `<p>`.

Mostriamo di seguito una porzione di codice:

```
<sp>
<speaker>COLOMBINE.</speaker>
<p>Ah pour le coup, Monsieur, j'y pers mon Latin. Votre femme morte
depuis six mois, vous a laissé tout au moins deux cent mille livres,
et pour plus d'un million de repos; et cependant, malgré ce grand
crespe, et ce deuil qui ne devrait pas passer l'habit, je vous
trouve un esprit aussi lugubre, que si l'on vous menaçoit de
ressusciter la défunte. Je vous avoüe que cela me passe, et je
n'aurois jamais crû qu'il y eust aucun chagrin assez bourru, pour
oser s'attaquer à la personne d'un homme veuf.
</p>
</sp>
```

3.6.3.3 L'elemento `<speaker>`

L'elemento `<sp>` è seguito da `<speaker>`, il quale identifica il nome del personaggio che recita una determinata battuta.

Mostriamo di seguito una porzione di codice:

```
<sp>
<speaker>COLOMBINE.</speaker>
[...]
</sp>
```

3.6.3.4 L'elemento `<stage>`

Il marcatore `<stage>` fornisce le informazioni riguardanti le didascalie e le direttive di scena presenti all'interno di un testo o di un frammento di testo drammatico. Il valore di questo marcatore viene specificato mediante l'attributo `<type>`, il quale possiede una serie di valori a seconda della natura dell'indicazione di scena. Troviamo, infatti, valori come `<setting>` (descrive l'ambientazione); `<entrance>` (descrive un ingresso); `<exit>` (descrive un'uscita); `<business>` (descrive l'attività sul palco); `<novelistic>` (rappresenta un testo narrativo che indica le direzioni di scena); `<delivery>` (descrive il modo in cui un personaggio parla); `<modifier>` (fornisce dettagli sul personaggio); `<location>` (descrive l'ambientazione); `<mixed>` (descrive più di uno dei valori descritti).

Mostriamo di seguito una porzione di codice:

```
<sp>
<speaker>COLOMBINE,</speaker>
<stage type="delivery">
en riant.
</stage>
</sp>
```

Come si può osservare dal codice, l'elemento `<stage>` è caratterizzato anche dall'attributo `<type="delivery">`, il quale permette di associare un valore specifico all'indicazione di scena contenuta all'interno.

3.6.3.5 L'elemento `<p>`

I paragrafi sono la parte più piccola di un'unità strutturale di un testo in prosa. Nella codifica i paragrafi sono delimitati dall'elemento `<p>`. Dal momento che nell'originale essi

hanno un carattere proprio, nell'edizione elettronica abbiamo cercato di rispettare le caratteristiche della fonte, riproducendoli esattamente.

Mostriamo di seguito una porzione di codice:

```
<sp>
<speaker>COLOMBINE,</speaker>
<p>
Ah, par ma foy, voilà du fruit nouveau; un Mary qui pleure sa femme !
Hé fy, Monsieur, ne faites pas cette sottise-là devant le monde, vous
feriez crier les petits enfans après vous.
</p>
</sp>
```

3.6.3.6 L'elemento **<pb>**

Nella codifica delle opere di Evaristo Gherardi abbiamo rispettato anche le interruzioni di pagina, codificate mediante l'elemento `<pb n="" />`, elemento essenziale per segnalare nella codifica il passaggio da una pagina ad un'altra. L'elemento `<pb>` appartiene alla classe degli elementi *milestone*, poiché esso non marca delle porzioni di testo, ma identifica dei punti di riferimento all'interno del contenuto testuale. Nella loro codifica abbiamo rispettato le *Guidelines* TEI. I `<pb>` sono stati inseriti nella stessa posizione della pagina a cui fanno riferimento. Inoltre, mediante l'attributo `<n>` è stato indicato il numero esatto della pagina che segue il punto di inserimento.

Mostriamo di seguito una porzione di codice:

```
<lb></lb>
<pb n="5">
</pb>
<lb></lb>
```

Conclusioni

Dopo aver completato la codifica delle opere, siamo passati alla revisione del lavoro compiuto: lo scopo è stato quello di cercare errori sia nella trascrizione del documento, sia nella marcatura del testo prima di procedere alla creazione dei fogli di stile XSLT.

Per fare questo abbiamo prima verificato la correttezza del contenuto testuale in base all'edizione originale di riferimento con lo scopo di individuare probabili errori (si è riletto, pertanto, il testo codificato per confrontare l'edizione elettronica con quella cartacea) e poi revisionato la marcatura TEI-XML con lo scopo di riscontrare il massimo rispetto alle *Guidelines* TEI.

Il nostro modo di organizzare e di produrre il sapere dipende da esperienze culturali, ma anche corporee e gestuali, che trasformano il pensiero in atto concreto. Oggi l'esperienza è orientata verso la scrittura elettronica, dove ogni presentazione del testo, sia testuale che grafica, non è altro che una possibilità offerta al lettore capace, in questo modo, di produrre autonomamente sapere, perché *la nostra memoria è per essenza ricostruttiva*.

In questo scenario di cambiamento sono mutati i criteri della creazione e della comunicazione della conoscenza che si offre adesso come atto produttivo interattivo. In questo modo il virtuale diventa reale.

CAPITOLO IV

Le Théâtre Italien online

La disumanità del computer sta nel fatto che, una volta programmato e messo in funzione, si comporta in maniera perfettamente onesta⁸⁷.

⁸⁷ Aforisma di Isaac Asimov.

Introduzione

La codifica elettronica è stata svolta con la consapevolezza che, una volta terminato questo lavoro, sarebbe stata resa graficamente. Questo aspetto è stato necessario per due motivi: il primo riguarda la natura estetica del progetto; il secondo riguarda la volontà di spiegare meglio gli elementi – molti browser attualmente disponibili non sono in grado di gestirgli e rappresentano soltanto l'albero del documento XML. È un limite tanto più evidente in testi lunghi e complessi come quello che è stato preso in esame per questa Tesi.

La visualizzazione grafica è stata possibile utilizzando il linguaggio *Extensible Stylesheet Language* (XSL)⁸⁸ e uno dei suoi componenti, l'*Extensible Stylesheet Language Transformations* (XSLT) e i fogli di stile CSS.

⁸⁸ L'*Extensible Stylesheet Language* è il linguaggio che descrive i fogli di stile che supportano graficamente i documenti XML. I fogli di stile permettono, dunque, di rispettare quanto prescritto dall'*Extensible Markup Language*, il quale prevede che i contenuti di un documento siano separati dalla formattazione della pagina. È merito dell'XSL se è possibile visualizzare un file XML in formati diversi (pagina web e pagina stampabile). L'XSL incorpora, al suo interno, tre linguaggi: *XSL Transformations*, XSLT (linguaggio di trasformazione XML), da noi utilizzato; l'*XSL Formatting Objects*, XSL.FO (usato per l'applicazione degli stili e della visualizzazione a un documento XML); *XML Path*, XPath (usato nei fogli di stile XSLT per descrivere il percorso da seguire per visualizzare parti di un documento XML). Tutti questi linguaggi sono codificati e gestiti dal W3C.

Come vedremo successivamente, la grande flessibilità dei browser attualmente disponibili (in particolare, Internet Explorer e Mozilla) nel supportare la trasformazione di XML in XHTML ha permesso l'utilizzo dei fogli di stile XSLT.

4.1 Generare l'output del *Recueil*

I documenti XML costituiscono un formato generale per la rappresentazione della struttura logica dei dati: per questo motivo, una volta terminata la codifica elettronica del *Recueil*, si è proceduto a creare una rappresentazione più corretta del materiale prodotto. Per visualizzare le informazioni in un modo diverso dalla semplice rappresentazione ad albero è possibile ricorrere a due linguaggi: i fogli di stile CSS o i fogli di stile XSLT. Il primo, che vedremo meglio successivamente, non è adatto a rappresentare un documento XML, in quanto risulta poco flessibile sia nel gestire le informazioni sia nell'introduzioni di elementi specifici. Per questo motivo, siamo ricorsi all'utilizzo del linguaggio XSL (*Extensible Style Language*), linguaggio di descrizione dei fogli di stile per i documenti in formato XML. Esso contiene al suo interno tre sotto-linguaggi: l'*XSL Formatting Objects* (XSL:FO); l'*XML Path* (XPath); l'*XSL Transformations* (XSLT). Il primo è un linguaggio di tipo obiettivo che ben si presta alla realizzazione del layout; il secondo come accedere a specifiche parti di un documento XML; il terzo è un linguaggio di tipo dichiarativo per la specifica di trasformazioni di documenti XML. L'XSLT venne progettato con lo scopo di servire come linguaggio generalizzato per i fogli di stile, ma la sua flessibilità gli ha permesso adesso di essere utilizzato per molteplici applicazioni.

Per trasformare il *Recueil* da noi codificato in XML siamo ricorsi a questo linguaggio: esso ha permesso, infatti, di generare un file in formato XHTML partendo da quello originale mediante questa dichiarazione iniziale:

```
<?xml-stylesheet type="text/xsl" href="IIitome.xsl"?>
```

Usando questa dichiarazione è stato possibile avere una prima rappresentazione grafica di uno dei volume codificati. Riportiamo di seguito come essa appare:

[Il Progetto Il Recuil La Codifica de Le Theatre Italian ricerca](#)

Tome II

- [LA CRITIQUE DE LA CAUSE DES FEMMES.](#)
- [LE MARCHAND DUPPE.](#)
- [SCENES FRANCOISES DE LA DESCENTE DE MEZZETIN AUX ENFERS.](#)
- [LA CRITIQUE DE L'HOMME A BONNE FORTUNE.](#)

La Cause De Les Femmes

```

<?xml-stylesheet type="text/xsl"
href="C:/Users/Lalli/Desktop/progetto/opera.xsl"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
xmlns="http://www.w3.org/1999/xhtml" xmlns:tei="http://www.tei-
c.org/ns/1.0" version="1.0">
<xsl:output method="xhtml" encoding="UTF-8" doctype-public="-
//W3C//DTD XHTML 1.0 Strict//EN" doctype-
system="http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd"
indent="yes">
</xsl:output>
<xsl:template match="/">
<html>
<head>
<title> Le Théâtre Italien</title>
<link rel="stylesheet" type="text/css" media="screen, projection"
href="./style.css">
</link>
</head>
<body>
<div class="container">
<div class="menunav">
<a>
<xsl:attribute name="href">
<xsl:text>home.html</xsl:text>
</xsl:attribute>
<xsl:text>Il Progetto</xsl:text>
</a>
<a>
<xsl:attribute name="href">
<xsl:text>opera.html</xsl:text>
</xsl:attribute>
<xsl:text>Il Recueil</xsl:text>
</a>
<a>
<xsl:attribute name="href">
<xsl:text>codifica</xsl:text>
</xsl:attribute>
<xsl:text>La Codifica de Le Théâtre Italien</xsl:text>
</a>
<a>

```

```

<xsl:attribute name="href">
<xsl:text>ricerca.html</xsl:text>
</xsl:attribute>
<xsl:text>ricerca</xsl:text>
</a>
<div class="clearer">
<span>
</span>
</div>
</div>
<div class="main">
<div class="subnav">
<h2>Tome II</h2>
<ul>
<xsl:for-each select="//tei:body/tei:div/tei:head">
<li>
<a href="#{../@n}">
<xsl:attribute name="class">
<xsl:text>ancore</xsl:text>
</xsl:attribute>
<xsl:apply-templates>
</xsl:apply-templates>
</a>
</li>
</xsl:for-each>
</ul>
</div>
</div>
<div class="content">
<xsl:apply-templates select="tei:TEI//tei:group/tei:text/tei:front">
</xsl:apply-templates>
<xsl:apply-templates select="tei:TEI//tei:body">
</xsl:apply-templates>
</div>
<div class="clearer">
<span></span>
</div>
</div>
<div class="bottom">
<div class="footer">
<p><xsl:apply-templates select="tei:TEI//tei:respStmt">
</xsl:apply-templates><br></br>

```

```

<xsl:text>Sito valido</xsl:text>
<a href="http://validator.w3.org/">XHTML</a><xsl:text> |
</xsl:text><a href="http://jigsaw.w3.org/css-validator">
<xsl:text>CSS</xsl:text></a>
</p>
</div>
</div>
</body>
</html>
</xsl:template>
<xsl:template match="tei:div/tei:head">
<h4>
<a name="{../@n}">
<xsl:apply-templates>
</xsl:apply-templates>
</a>
</h4>
</xsl:template>
<xsl:template match="tei:text/tei:group/tei:text/tei:front/tei:head">
<xsl:choose>
<xsl:when test="@type='title' ">
<h1><xsl:apply-templates/>
</h1>
</xsl:when>
<xsl:when test="@type='subtitle' ">
<h2><xsl:apply-templates/>
</h2>
</xsl:when>
<xsl:otherwise>
<h4 class="auteurs">
<xsl:apply-templates/></h4>
</xsl:otherwise>
</xsl:choose>
</xsl:template>
<xsl:template match="tei:div/tei:p/tei:graphic">
<img class="immagine" alt="immagine">
<xsl:attribute name="src">
<xsl:value-of select="@url">
</xsl:value-of>
</xsl:attribute>
</img>
</xsl:template>

```

```

<xsl:template match="tei:lb">
<br></br>
</xsl:template>
<xsl:template match="//tei:stage">
<p class="corsivo">
<xsl:apply-templates/>
</p>
</xsl:template>
<xsl:template match="//tei:sp">
<br/><xsl:apply-templates/>
</xsl:template>
<xsl:template match="//tei:set">
<p class="corsivo">
<xsl:apply-templates/>
</p>
</xsl:template>
<xsl:template match="//tei:roleDesc">
<i><xsl:apply-templates/>
</i>
</xsl:template>
<xsl:template match="tei:pb">
<br></br>
</xsl:template>
</xsl:stylesheet>

```

Nei fogli di stile XSLT viene fornito un'insieme di norme, dette regole di *template* (sono così chiamate in quanto contengono al loro interno questo elemento), le quali, mediante l'attributo *match*, passano al pattern l'istruzione da eseguire. Per questo motivo, il linguaggio XSL risulta più complesso di quello alla base dei fogli di stile CSS, in quanto essi non si rivolgono direttamente al documento.

Dunque, la trasformazione del file ha seguito una serie di fasi: il foglio di stile è stato applicato alla radice del documento⁸⁹, a cui, in seguito, vengono applicate le regole specifiche. Mediante, infatti, l'istruzione `<apply-templates></apply-templates>` si applica ricorsivamente il foglio di stile al nodo selezionato.

Pertanto, quando un processore XSLT applica ad un foglio di stile il file XML originale, cerca una regola di trasformazione adeguata all'output da produrre ed esegue il corpo del documento. L'XSLT 1.0 usa l'*Xpath* 1.0 come linguaggio per scrivere le espressioni, le quali

⁸⁹ Trattandosi della parte iniziale del documento, dove non è ancora stata specificata alcuna regola di trasformazione, al nodo viene applicata una regola di default predefinita che gli permette di procedere. Questa regola è data dall'elemento "match".

hanno l'obiettivo di specificare i pattern delle regole di trasformazione; di scegliere i nodi da trasformare; di calcolare eventuali condizioni booleane; di generare il contenuto testuale in output XHTML.

Una regola di trasformazione corrisponde al documento o ad una parte di esso in input e costruisce una parte del documento in output. Un esempio di regola di trasformazione può essere la seguente:

```
<xsl:template match="//tei:stage">
  <i>
    <xsl:apply-templates/>
  </i>
</xsl:template>
```

L'attributo *match* contiene un pattern e il contenuto dell'istruzione *template* è un costruttore di sequenza. Pertanto, un'insieme di regole di trasformazione, identificato con il sotto-insieme di queste norme che il pattern associa al nodo corrente, produce un risultato nel momento in cui viene valutato il costruttore di sequenze della regola selezionata.

4.2 La trasformazione della *TEI Header*

Iniziamo un'analisi più dettagliata dei fogli di stile con la descrizione di quanto fatto per creare l'output HTML per la *<teiHeader>* del *corpus* codificato in TEI-XML.

Il linguaggio XSLT permette la trasformazione di un file XML in un altro formato (XML, HTML e TXT) e, quindi, è capace di estrarre i dati: l'applicazione, infatti, sulla base alle istruzioni fornite all'interno del documento, seleziona i dati del file di origine e li posiziona all'interno di una nuova struttura che sarà un documento HTML. Un documento XSLT si apre con la dichiarazione XML:

```
<?xml version="1.0" encoding="UTF-8"?>
```

Questa, oltre a fornire la descrizione sulla versione XML utilizzata, fornisce informazioni in merito al sistema di codifica utilizzato: in questo, l'UTF-8.

Ad essa segue la dichiarazione dell'elemento radice del foglio di stile XSLT, che specifica la versione utilizzata, 1.0, e i criteri di trasformazione.

```
<xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
```

A cui segue subito dopo l'indicazione della versione del formato HTML desiderato per l'output:

```
<xsl:output method="html" encoding="iso-8859-1" version="4.01"
doctype-public="-//W3C//DTD HTML 4.01 Transitional//EN"/>
```

Dopo questa prima fase, si apre la sintassi del foglio di stile XSLT basata sui *template*, modelli o regole, all'interno delle quali, sono contenute le istruzioni che mettono in relazione i nodi (elementi) del file XML con le istruzioni definite nel foglio di stile XSLT.

Mostriamo di seguito una porzione di codice:

```
<xsl:template match="/">
<html>
<head>
<title>Le Theatre Italian</title>
</head>
<body>
<h1>
<xsl:value-of select="/TEI/teiHeader"/>
</h1>
<xsl:apply-templates/>
```

```

</body>
</html>
</xsl:template>

```

Questa porzione di codice, che segue un percorso di tipo assoluto che parte dall'elemento radice e segue l'albero del documento XML, consente al processore di mandare in output le istruzioni presenti al suo interno. In questo modo, viene creato il file HTML con gli opportuni tag: `<html>` `<head>` `<title>` contenuti all'interno di `<fileDesc>`. Questo criterio viene seguito per alcune parti che compongono la `<teiHeader>`: il `<fileDesc>`; l'`<encodingDesc>`; il `<profileDesc>`. Mostriamo di seguito una porzione di codice che permette la trasformazione del file:

```

<xsl:template match="tei:teiHeader">
  <xsl:apply-templates>
</xsl:apply-templates>
</xsl:template>
<xsl:template match="tei:filedesc">
  <xsl:apply-templates>
</xsl:apply-templates>
</xsl:template>
<xsl:template match="tei:titleStmt">
  <xsl:apply-templates>
</xsl:apply-templates>
</xsl:template>

```

Con queste istruzioni si procede all'elaborazione del percorso fino al momento in cui non si incontra `</xsl:apply-templates>`, istruzione che impone al processore di fermarsi e di richiamare i modelli per gli elementi figli dell'elemento selezionato. Solo dopo aver effettuato queste operazioni, il *template* ritorna attivo ed esegue le operazioni che rimangono. In particolare, il contenuto del primo *template* si limita a dire al processore di chiamare gli altri modelli per gli elementi figli del nodo corrente: l'elemento `<teiHeader>` è, infatti, una sorta di contenitore per la serie di sotto-elementi contenenti i dati veri e propri. Una impostazione analoga al modello adottato per il `<teiHeader>` è presente anche nel suo primo figlio, `<fileDesc>` e, a sua volta, per il figlio di questo, `<titleStmt>`: elementi contenenti un gruppo di tag che marcano informazioni strutturate e semantiche. È opportuno fare una precisazione in merito al comportamento di `</xsl:applytemplates>`. Ecco l'output prodotto da questa istruzione.

```

<html>
  <head>

```

```

<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-
1">
<title>Le Théâtre Italien: edizione elettronica</title>
</head>
<body>

```

Il foglio di stile prevede, inoltre, altre istruzioni, come:

```

<xsl:template match="tei:fileDesc/tei:titleStmt">
<br></br><br></br><h2>
<strong>
<a name="descrizioneEdizioneElettronica">
<xsl:text>Descrizione Dell'Edizione Elettronica</xsl:text>
</a>
</strong>
</h2>
<xsl:apply-templates>
</xsl:apply-templates>
<br></br>
</xsl:template>
<xsl:template
match="tei:teiHeader/tei:fileDesc/tei:titleStmt/tei:title">
<br></br><br></br>
<strong>
<xsl:text>Titolo dell'Opera:</xsl:text>
</strong>&#160;
<xsl:apply-templates>
</xsl:apply-templates>
</xsl:template>
<xsl:template match="tei:fileDesc/tei:titleStmt/tei:author">
<br></br><br></br>
<strong>
<xsl:text>Autore:</xsl:text>
</strong>&#160;
<xsl:apply-templates>
</xsl:apply-templates>
<br></br><br></br>
</xsl:template>
<xsl:template
match="tei:fileDesc/tei:titleStmt/tei:respStmt/tei:resp">
<xsl:text disable-output-escaping="yes"></xsl:text>
<strong>&#160;

```

```

<xsl:apply-templates>
</xsl:apply-templates>:</strong>
</xsl:template>
<xsl:template
match="tei:fileDesc/tei:titleStmt/tei:respStmt/tei:name">
<xsl:apply-templates>
</xsl:apply-templates>
<xsl:text disable-output-escaping="yes">
</xsl:text>
</xsl:template>

```

Questo *template*, per ogni elemento contenuto all'interno di *<fileDesc>*, crea un paragrafo che scrive il “titolo della serie”; il “titolo dell’opera”; il nome dell’“autore”, i quali, racchiusi dentro il tag **, saranno riprodotti in grassetto nell’output HTML.

La medesima impostazione del *template* appena utilizzato si ritrova anche per gli altri elementi presenti all'interno della *<teiHeader>*. Si discostano dal precedente modello, invece, le dichiarazioni di responsabilità, le quali, nel sorgente XML, sono racchiuse all'interno del tag *<respStmt>* che, a sua volta, contiene gli attributi *<resp>* e *<name>*: il primo descrive la natura della responsabilità; il secondo il nome del responsabile della codifica. Nella *<teiHeader>* questa coppia è ricorsa due volte: la prima per registrare la responsabilità della digitalizzazione e codifica nella persona di Laura Giannini; la seconda per la dichiarazione di responsabilità della supervisione linguistica e tecnica del Professore Roberto Rosselli Del Turco e della Dottoressa Barbara Sommovigo. Riportiamo di seguito la porzione di codice corrispondente:

```

<xsl:template
match="tei:fileDesc/tei:titleStmt/tei:respStmt/tei:resp">
<xsl:text disable-output-escaping="yes">
</xsl:text>
<strong>&#160;
<xsl:apply-templates>
</xsl:apply-templates>:</strong>
</xsl:template>
<xsl:template
match="tei:fileDesc/tei:titleStmt/tei:respStmt/tei:name">
<xsl:apply-templates>
</xsl:apply-templates>
<xsl:text disable-output-escaping="yes">
</xsl:text></xsl:template>

```

Il *template* del primo elemento `<respStmt>` richiama i modelli degli elementi figli per mandarli in output HTML mediante l'espressione `<xsl:text disable-output-escaping="yes"><p></xsl:text>`; successivamente, mediante l'elemento `</xsl:apply-templates>`, scrive nell'output il valore dell'elemento `<resp>` racchiudendolo dentro il tag ` ... ` e facendolo seguire dal segno di due punti ed uno spazio bianco. Il *template* dell'elemento `<name>`, invece, subito dopo aver trasmesso nell'output il proprio valore scrive il tag HTML di chiusura paragrafo `</p>` sempre grazie all'espressione: `<xsl:text disable-output-escaping="yes"></p></xsl:text>`.

Mostriamo di seguito la porzione di codice HTML prodotto dai modelli appena illustrati: Anche l'elemento `<availability>` ha un *template* più particolare, che mostriamo di seguito:

```
<xsl:template match="tei:publicationStmt/tei:availability">
<br></br><br></br>
<strong>
<xsl:text>Copyright:</xsl:text>
</strong>&#160;
<xsl:apply-templates>
</xsl:apply-templates>
</xsl:template>
```

Il primo *template* chiama, mediante l'istruzione `</xsl:apply-templates>`, i modelli da seguire per gli elementi figli di `<availability>` che, in questo caso, è soltanto uno, il copyright, contenente le dichiarazioni di fruizione del documento.

Il linguaggio XSLT offre, inoltre, la possibilità di stabilire che determinate operazioni devono essere svolte solo al verificarsi di condizioni specifiche. Nel *template* che segue sono fornite le informazioni inerenti alla citazione bibliografica del testo originale e i criteri di revisione seguiti.

```
<xsl:template
match="tei:revisionDesc/tei:change/tei:desc/tei:bibl/*">
<xsl:choose>
<xsl:when test="position() = 1">: &#160;
<xsl:apply-templates>
</xsl:apply-templates>,
</xsl:when>
<xsl:when test="position() != last()"> &#160;
<xsl:apply-templates>
</xsl:apply-templates>,
</xsl:when>
<xsl:otherwise> &#160;
```

```
<xsl:apply-templates>
</xsl:apply-templates>.
</xsl:otherwise>
</xsl:choose>
</xsl:template>
```

In questa porzione di codice si trova l'espressione XPath, la quale seleziona tutti gli elementi figli di `<bibl>`. Il comando `<xsl:choose>` ordina al processore, ogni qual volta si verificano determinate condizioni, di eseguire azioni diverse: quando il nodo elaborato è il primo (la condizione è data dall'espressione `test="position() = 1"`) il processore scrive nell'output il carattere due punti e successivamente due spazi bianchi e, dopo aver fatto visualizzare il contenuto dell'elemento (`<xsl:apply-templates/>`), scrive una virgola; quando l'elemento selezionato non è l'ultimo (`test="position() != last()"`), allo stesso modo della condizione precedente, scrive due spazi bianchi e poi una virgola. L'ultimo comando è quello di `<xsl:otherwise>`, il quale stabilisce che in tutti gli altri casi si generino in output due spazi bianchi ed il contenuto dell'elemento seguito dal punto.

Concludendo, possiamo dire che queste sono dimostrazioni pratiche di alcune delle innumerevoli potenzialità che si hanno nella rappresentazione di dati identificati con l'utilizzo di una marcatura di tipo dichiarativo. Al momento della codifica TEI-XML, infatti, abbiamo racchiuso le varie porzioni di testo all'interno di tag con funzioni specifiche: dunque, se con la codifica abbiamo espresso cosa un tag contenesse, soltanto con i fogli di stile XSLT abbiamo selezionato i contenuti del blocco testuale da rappresentare. In questo modo, un testo codificato non è più vincolato ad un'unica rappresentazione, bensì si presta alle varie esigenze degli utenti.

4.3 La trasformazione del *Recueil*

Come nel caso della *<teiHeader>*, anche per generare l'output dell'intero *corpus* si è ricorsi ai fogli di stile XSLT.

Prima di entrare nel vivo dei fogli di stile, è opportuno ricordare due criteri seguiti per la codifica della raccolta: in primo luogo, trattandosi di volumi, si è deciso di considerare i testi come testi compositi; in secondo luogo, trattandosi di testi teatrali, le parti marcate sono state soprattutto quelle inerenti alla lista degli attori; alle ambientazioni; alle didascalie di scena. Per far sì che il testo codificato fosse rappresentato nella sua interezza e, quindi, corredato sia dal proprio frontespizio che dal blocco testuale, la prime istruzioni che abbiamo fornito sono state la seguenti:

```
<xsl:for-each select="//tei:body/tei:div/tei:head">
  <li>
    <a href="#{../@n}">
      <xsl:attribute name="class">
        <xsl:text>ancore</xsl:text>
      </xsl:attribute>
      <xsl:apply-templates>
    </xsl:apply-templates>
    </a>
  </li>
</xsl:for-each>
</ul>
</div>
</div>
<div class="content">
  <xsl:apply-templates select="tei:TEI//tei:group/tei:text/tei:front">
</xsl:apply-templates>
  <xsl:apply-templates select="tei:TEI//tei:group/tei:text/tei:body">
</xsl:apply-templates>
```

La prima istruzione (*<xsl:for-each select="//tei:body/tei:div/tei:head">*) calcola la sequenza specificata dall'attributo "select" e applica il costruttore di sequenza definito nel suo corpo ad ogni componente. Questa istruzione consente al parser di rendere cliccabili gli elementi "head" ogni qual volta essi si trovino all'interno dell'elemento *<div>*, il quale deve essere contenuto dentro *<body>*.

Le istruzioni (*<xsl:apply-templates select="tei:TEI//tei:group/tei:text/tei:front">* e *<xsl:apply-templates select="tei:TEI//tei:group/tei:text/tei:body">*) forniscono al parser le informazioni

inerenti al percorso da seguire: la prima si riferisce al frontespizio del volume; la seconda al contenuto testuale. A proposito della prima si è proceduto a fornire un'ulteriore istruzione:

```
<xsl:template match="tei:text/tei:group/tei:text/tei:front/tei:head">
<xsl:choose>
<xsl:when test="@type='title' ">
<h1>
<xsl:apply-templates/>
</h1>
</xsl:when>
<xsl:when test="@type='subtitle' ">
<h2>
<xsl:apply-templates/>
</h2>
</xsl:when>
<xsl:otherwise>
<h4 class="auteurs">
<xsl:apply-templates/>
</h4>
</xsl:otherwise>
</xsl:choose>
</xsl:template>
```

Con questo *template* vengono fornite le istruzioni per rappresentare i titoli presenti nel <front>: ogni tipologia avrà, infatti, una rappresentazione specifica. Il tipo “title” avrà la rappresentazione in <b1>; il tipo “subtitle” <b2> e, infine, la lista degli attori, se presente, in <b4>. Questa istruzione è possibile grazie all'utilizzo dell'elemento <choose>, istruzione condizionale che rappresenta in ordini diversi i rami di esecuzione.

Dopo questa istruzione, il foglio di stile si concentra sulle istruzioni date ai marcatori presenti nel blocco testuale. Mostriamo di seguito alcune porzioni di codice:

```
<xsl:template match="//tei:stage">
<i>
<xsl:apply-templates/>
</i>
</xsl:template>

<xsl:template match="//tei:set">
<p class="corsivo"><xsl:apply-templates/></p>
</xsl:template>
```

```

<xsl:template match="//tei:roleDesc">
<p class="corsivo"><xsl:apply-templates/>
</xsl:template>

```

Questi *template* contengono le istruzioni per rappresentare in HTML l'elemento `<stage>`, `<set>`, `<roleDesc>`. Questi sono ulteriormente caratterizzati dalla presenza dell'elemento `<i>`. Quest'ultimo elemento, definito nel CSS, permette di rappresentare le parole, marcate con questi elementi, in corsivo.

Sono stati definiti *template* anche per altri elementi

```

<xsl:template match="//tei:sp">
<br/><xsl:apply-templates/>
</xsl:template>
</xsl:stylesheet>

```

```

<xsl:template match="tei:div/tei:p/tei:graphic">
<img class="immagine" alt="immagine">
<xsl:attribute name="src">
<xsl:value-of select="@url">
</xsl:value-of>
</xsl:attribute>
</img>
</xsl:template>

```

```

<xsl:template match="tei:lb">
<br/>
</xsl:template>

```

Il primo modello contiene le istruzioni per rappresentare l'elemento `<sp>`; il secondo consente di inserire, all'interno dell'output HTML, anche le immagini; il terzo trasforma gli elementi `<lb>` in `
` per garantire l'uso dell'alinea.

4.4 I *Cascading Style Sheet*

Dal momento che i fogli di stile XSLT permettono di trasformare un documento XML in un formato diverso, nel nostro caso XHTML, essi si sono rivelati insufficienti per rappresentare in una forma grafica gradevole il documento originale. Per questo motivo, siamo ricorsi ai fogli di stile CSS (*Cascading Style Sheet*) per gestire il layout del lavoro codificato. Essi sono meno complessi dei fogli di stile XSL, ma, allo stesso tempo, risultano meno flessibili nel momento in cui si vogliamo manipolare dati (note e immagini). Per questo motivo, là dove si voglia rendere un documento XML gradevole dal punto di vista estetico possono essere utilizzati senza difficoltà.

Esistono due versioni di CSS, la CSS1 e la CSS2, evoluzione della precedente, in quanto integra maggiori funzioni⁹⁰: per questo motivo, ancora oggi la maggior parte dei browser supporta meglio la prima specifica. Per questo motivo, noi siamo ricorsi alla versione CSS1.

Per allegare un foglio di stile CSS al documento XML si è dovuti ricorrere ad una dichiarazione: posta subito dopo la dichiarazione della versione XML utilizzata, essa si presenta in questo modo:

```
<?xml-stylesheet type="text/xsl" href=""?>
```

Il file del foglio di stile non è un altro che un file di testo, la cui estensione sarà, però, quella di .css. Tuttavia, affinché il CSS sia associabile al documento XML, è opportuno che ogni elemento presente nel file originario segua una regola, composta da un “selettore”, espressione che specifica a quali oggetti vada applicata la norma, e dalla “specifica di una regola”, l’insieme delle proprietà di stile che ne costituiscono la definizione e che è racchiusa tra parentesi graffe. Lo pseudo-codice è il seguente:

```
selettore {specifica};
```

4.4.1 La sintassi dei fogli di stile CSS

Sebbene collegati al linguaggio HTML, i CSS presentano una sintassi completamente diversa, in quanto si basa sulla dichiarazione di selettori, proprietà e valori.

⁹⁰ CSS 2.1 è un linguaggio di fogli di stile che consente agli autori e agli utenti di collegare gli stili (font e spaziature) a documenti strutturati (documenti HTML ed applicazioni XML). Supporta fogli di stile media-specifici in modo tale che gli autori possano adattare la presentazione dei propri documenti ai browser visuali; ai dispositivi di fonetica; alle stampanti; ai dispositivi braille; ai dispositivi palmari. Gestisce, inoltre, il posizionamento dei contenuti; il layout delle tabelle; alcune proprietà dell’interfaccia utente.

I selettori specificano a quali marcatori HTML devono essere applicate le dichiarazioni di stile. Per questo motivo, possono presentarsi in tre forme diverse: i selettori HTML; i selettori id; i selettori di classe.

I selettori HTML indicano gli elementi HTML a cui applicare le dichiarazioni. Ne è un esempio la seguente porzione di codice:

```
h1, h2 {
font-family:"Times New Roman", Times, serif; color: #000000;
font-weight:normal;
}
```

Questa regola applica agli elementi marcati dall'attributo "h2" il font *Times New Roman* e il colore nero.

I selettori "id" applicano gli stili agli elementi dichiarati mediante l'attributo "id". Ne è esempio la seguente porzione di codice HTML i cui elementi sono richiamati nel foglio di stile CSS:

```
<div id="quickSummary">
<div class="menunav">
<a href="index.html">Il Progetto</a>
<a href="opera.html">Il Recueil</a>
<a href="frontespizio.html">Il Frontespizio Elettronico</a>
<a href="codifica.html">La Codifica de Le Théâtre Italien</a>
<div class="clearer">
<span></span>
</div>
</div>
```

```
#quickSummary {
background: no-repeat url("top2.jpg");
width: 811px;
height: 85px;
margin: 0px;
padding: 0px;
}
```

I selettori di classe applicano gli stili a quegli elementi che presentano la classe corrispondente. Mostriamo di seguito le porzioni di codice corrispondenti:

```
<xsl:template match="//tei:set">
<p class="corsivo">
```

```

<xsl:apply-templates/>
</p>
</xsl:template>

.corsivo{
font-style: italic;
}

```

Se con i selettori è possibile associare degli stile agli elementi presenti nell'HTML, è con le proprietà che vengono definiti gli aspetti legati alla presentazione di un documento: le proprietà, dunque, servono a specificare il colore, la dimensione di un font, il criterio di impaginazione.

Se riutilizziamo, ad esempio, le prime due porzioni di codice mostrate, e che illustriamo nuovamente adesso,

```

h1, h2 {
font-family:"Times New Roman", Times, serif; color: #000000;
font-weight:normal;
}

#quickSummary {
background: no-repeat url("top2.jpg");
width: 811px;
height: 85px;
margin: 0px;
padding: 0px;
}

```

noteremo come la proprietà siano l'elemento "font-family" nel primo caso, mentre *width*, *height*, *margin*⁹¹, *padding*⁹² nel secondo. L'utilizzo della prima regola implica la visualizzazione della famiglia dei font *Times*, il cui valore iniziale è determinato dal browser. La seconda permette di definire la larghezza, l'altezza, il margine e lo spazio intorno all'elemento. Nel CSS progettato per *Le Théâtre Italien* siamo ricorsi anche alla proprietà *border*, la quale permette di definire i bordi⁹³. Mostriamo di seguito la porzione di codice corrispondente:

```

#intro acronym {
color: #5d5144;
}

```

⁹¹ La proprietà "margin" rappresenta l'involucro esterno del box e, per questo, è molto simile al "padding", in quanto anche in questo ci si limita a definire la grandezza.

⁹² Il pudding è costituito dai singoli lati di un box: per questo motivo è possibile avere padding-top; padding-right;padding-bottom; padding-left. Lo sfondo occupa proprio lo spazio definito da questo elemento.

⁹³ Come per il padding, anche la proprietà border permette di definire i bordi di un box ricorrendo agli elementi color e style.

```

font-style: italic;
text-decoration: none;
border-bottom: none;
}

```

L'attributo "bottom" permette di definire la misura dei singoli lati del box di riferimento.

All'interno di questo foglio di stile è, inoltre, presente anche la proprietà "display", la quale, se si guarda il codice corrispondente,

```

#quickSummary p.p2 {
display: block;
[...]
}

```

consente di rappresentare il contenuto senza un'interruzione di linea. I box, infatti, possono essere di due tipi: blocco oppure in linea. Il primo fa riferimento agli elementi di blocco (la porzione di codice ne è esempio con l'elemento "p") e permette che via sia un'interruzione di linea predefinita tra la riga precedente e la successiva, definita *block box*; il secondo fa riferimento agli elementi *inline* (come *em* e *spam*) visualizzati all'interno di una riga, definita *inline box*.

```

h3 span {
display: none;
}

```

Questa regola, costituita dalla proprietà "display" seguita dal valore "block", permette che i contenitori di blocco (ad esempio, l'elemento "span") che iniziano una nuova riga siano adattati al box corrispondente, in quanto la larghezza di un box *inline* è pari alla larghezza del contenuto.

Ad ogni proprietà vengono assegnati dei valori. Essi possono essere specifici per ognuna, come nel nostro caso, dove per costruire una pagina si è ricorsi ad unità di misura; oppure comuni a più proprietà. Mostriamo di seguito due porzioni di codice già utilizzate precedentemente:

```

#quickSummary {
background: no-repeat url("top2.jpg");
width: 811px;
height: 85px;
margin: 0px;
padding: 0px;
}

```

```
h1, h2 {  
font-family:"Times New Roman", Times, serif; color: #000000;  
font-weight:normal;  
}
```

Come risulta evidente i valori espressi dalla prima regola sono i pixel assegnati alla larghezza; all'altezza; al margine; allo spazio intorno all'elemento. I valori inseriti, invece, nella seconda norma sono codice esadecimale del colore del testo, in questo caso nero.

4.5 Il collegamento dei CSS ai file HTML

Per comprendere appieno quanto fatto per progettare il CSS *Le Théâtre Italien* di Evaristo Gherardi, è opportuno sottolineare come essi sono stati inseriti nel file HTML. Per applicare questi fogli di stile al documento da formattare vi sono tre modi. Esistono, infatti, tre tipi di CSS: i CSS *inline*; i CSS *embedded*; i CSS esterni.

I CSS *inline* hanno una progettazione più semplificata rispetto alle altre due versioni, in quanto prevedono l'inserimento dell'attributo *style* all'elemento HTML⁹⁴; i CSS *embedded* si basano sempre sull'attributo precedente, ma il loro utilizzo è diverso – è previsto infatti l'elemento *style* mediante una dichiarazione esterna all'elemento HTML a cui applicarlo⁹⁵. I CSS esterni sono, dunque, la migliore versione da utilizzare⁹⁶.

Come esemplifica il loro nome, essi si trovano in un file esterno – la sua estensione è “.css” – a cui le pagine HTML sono collegate mediante la seguente dichiarazione:

```
<link rel="stylesheet" type="text/css" media="screen, projection"
href="stili/css.css"/>
```

⁹⁴ Questa versione di CSS attualmente risulta superata, in quanto l'elemento *style* su cui si basa risulta attualmente deprecato. Inoltre, essa non permette la separazione del contenuto dalla forma, ovvero dell'HTML dalla sua resa grafica.

⁹⁵ Nel CSS *embedded* l'attributo *style* compare insieme all'elemento *text*, che indica al browser quale contenuto contenga al suo interno. Questa versione è migliore della precedente, in quanto garantisce una maggiore separazione del contenuto dalla forma.

⁹⁶ È questa versione che abbiamo utilizzato per la codifica de *Le Théâtre Italien* di Evaristo Gherardi.

4.6 La grafica finale

La grafica del *Recueil* è sicuramente uno degli aspetti che, durante la progettazione di questa Tesi, è stato realizzato con maggiore attenzione.

Trovare l'unione tra funzionalità e creatività, tenendo ben presente l'originale cartaceo, è stata forse una delle problematiche maggiori, in quanto se da una parte la grafica deve suscitare attenzione nell'utente, dall'altra essa non deve prevalere sui contenuti. Pertanto, ho cercato di mantenere un certo equilibrio anche nella realizzazione del layout.

4.6.1 Le immagini

L'immagine scelta per l'intestazione è un dipinto di Sergio Fasolini ed è intitolata *Maschere*. Essa è idonea a rappresentare le figure principali che compaiono nelle commedie de *Le Théâtre Italien* di Evaristo Gherardi. Sono state cambiate le dimensioni⁹⁷ e l'impostazione⁹⁸, come si può vedere confrontando le due immagini.



Maschere

⁹⁷ da 1771 * 1351 dell'immagine di partenza, siamo arrivati ad avere 990 * 379.

⁹⁸ Sull'immagine *Maschere* è stato applicato il filtro "affresco", utilizzando il pennello alla sua grandezza massima. E' stato allargato sfondo nero, in modo tale da rendere l'immagine nitida anche se ingrandita.



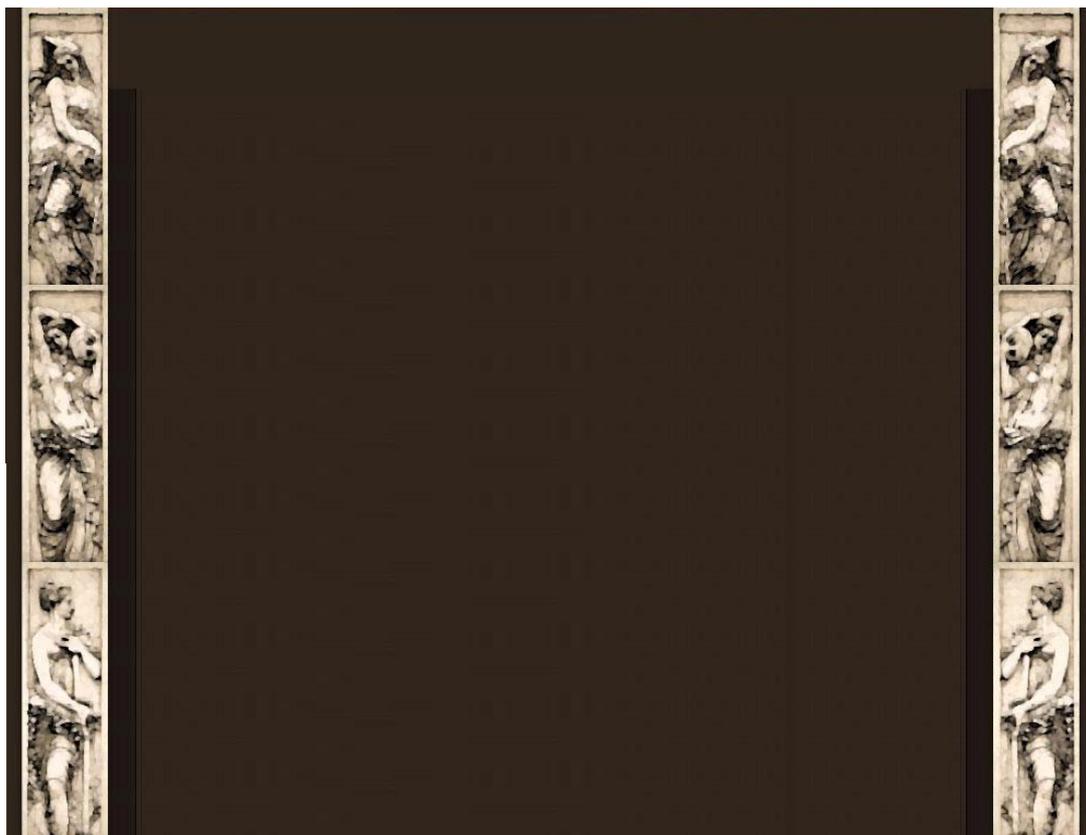
Le stesse operazioni compiute sull'immagine precedente sono state poi applicate a quella successiva. Ecco ciò che troviamo all'interno de *Le Théâtre Italien* online di Evaristo Gherardi.



Mediante l'applicazione di questa regola,

```
background: repeat-y url(sculptures.jpg);
```

definita nel CSS si è potuta arricchire la grafica con un decoro laterale alla pagina, il quale è stato ottenuto grazie alla ripetizione in verticale di una stessa scultura. Il risultato ottenuto è stato il seguente:



Conclusioni

Il tirocinio che richiede il Corso di Studi in Informatica Umanistica mi ha permesso di dedicarmi a questo lavoro, la cui validità consiste nella possibilità di rendere fruibile, in una biblioteca digitale, un'opera che altrimenti sarebbe difficilmente accessibile. Questa caratteristica rappresenta proprio il fine che si propone la Piccola Biblioteca Digitale Romanza, nata con l'intenzione di favorire la divulgazione di testi difficilmente reperibili, anche perché poco conosciuti.

Il mio lavoro è consistito nello scegliere una grafica accattivante, che induca l'utente a continuare la propria lettura o la propria ricerca. A tal fine, il contenuto della *Recueil* è stato riportato in modo che il fruitore potesse visualizzare immediatamente le battute della commedia distinguendole dalle parti introduttive e didascaliche, dalle indicazioni di scena e dall'elenco dei personaggi.

Dunque, il lavoro si è spostato lungo due linee conduttrici: la prima è stata determinata dal desiderio di arricchire la Piccola Biblioteca Digitale Romanza e la seconda dalla volontà di applicare la tecnica informatica alle discipline umanistiche, dimostrando che anche i mezzi più moderni servono a riscoprire le opere del passato. In tal modo, la mia Tesi ha rappresentato un piccolo, ma utile apporto allo studio della letteratura francese del Sei/Settecento.

Bibliografia

- Apollonio Mario, *Storia della Commedia dell'Arte*, Firenze, Sansoni, 1982.
- Basili Carla, Pettenati Corrado, *La biblioteca virtuale*, Milano, Bibliografica, 1994.
- Cecchi Umberto, *L'arlecchino del Re Sole*, Prato, Edizioni Del Palazzo, 1986.
- Ciotti Fabio, Roncaglia Gino, *Il mondo digitale. Introduzione ai nuovi media*, Roma-Bari, Editori Laterza, 2001.
- Genette Gérard, *Soglie. I dintorni del testo*, trad. it., Torino, Einaudi, 1989.
- Gigliozzi Giuseppe, *Il testo e il computer. Manuale di informatica per studi letterari*, Milano, Bruno Mondadori, 1997.
- Francesco Fiorentino, *Il teatro francese del Seicento*, Roma, Laterza, 2008.
- Griffith Patrick, *XHTML & CSS. Il web secondo HTML dog*, trad.it Pelleri R., Pearson Education Italia, 2007.
- Lenci Alessandro, *Testo e Computer. Elementi di Linguistica Computazionale*, Roma, Carocci, 2005.
- Pierazzo Elena, *La codifica dei testi*, Roma, Carocci, 2005.
- Salarelli Alberto, Tammaro Anna Maria, *La biblioteca digitale*, Milano, Editrice Bibliografica, 2000.
- Sand Maurice, *Le maschere italiane*, Roma, Editalia, 1994.
- Segre Cesare, *Avviamento all'analisi del testo letterari*, Torino, Einaudi, 1999.
- Shepherd D., *XML. Guida completa* (trad. it.), Milano, Apogeo, 2002, (ed. orig. *Teach yourself XML in 21 Days second edition*, Sams, 2001).
- Tessari Roberto, *La maschera e l'ombra*, Milano, Mursia, 1989.
- Testaverde Matteini Annamaria, Evangelista Anna, *I canovacci della commedia dell'arte*, Torino, Einaudi, 2007.
- Todarello Nazzareno, *Le arti della scena. Lo spettacolo in Occidente da Eschilo al trionfo dell'opera*, Padova, La Torre, 2006.
- Valeriano Leo, *La tradizione delle maschere italiane*, Roma, Rai-Eri, 2004.

Webliografia

Bardi Luca, *Prende forma la digital library*, Biblioteche Oggi, n.10, 1998:
<http://www.bibliotecheoggi.it/1998/19981000601.pdf>

Celentano A., Cortesi P., Mastrandrea P., *Informatica Umanistica: una disciplina di confine*, Mondo Digitale, n.4, dicembre 2004: http://www.mondodigitale.net/Rivista/05_numero_uno/Celentano_def.p.44-55.pdf

César : calendrier électronique des spectacles sous l'ancien régime et sous la révolution :
<http://www.cesar.org.uk/cesar2/>

Ciotti Fabio, *Cosa è la codifica informatica dei testi. Atti del Convegno Umanesimo & Informatica* (Trento 24-25 maggio 1996), a cura di Daniela Gruber e Patrick Pauletto: [http://circe.lett.unitn.it/circe/\[...\]/ciotti.pdf](http://circe.lett.unitn.it/circe/[...]/ciotti.pdf)

La Commedia dell'Arte: http://it.wikipedia.org/wiki/Commedia_dell%27arte

La commedia dell'arte al Burcardo: <http://www.burcardo.org/commedia/index.htm>

I comici rivoluzionari: <http://www.sipario.it/mcommediaintro.htm>

La Commedia dell'Arte: genesi, sviluppo e influenze sul balletto:
<http://www.balletto.net/giornale.php?articolo=706>

La Piccola Biblioteca Digitale Romanza: <http://piccolabdr.humnet.unipi.it/>

La Text Encoding Initiative: <http://www.tei-c.org/index.xml>

Le immagini riportate in questa Tesi sono tratte da:

Jacques Callot, 1621: <http://it.wikipedia.org/wiki/Zanni>

Louis Jacob, 1729: <http://www.burcardo.org/museo/comart01.htm>

Sergio Fasolini, 2010: <http://sergiofasolini.wordpress.com/>