A Note on the Structure and Contents of TEI P2

C.M. Sperberg-McQueen and Lou Burnard

Document Number: TEI EDJ6

16 March 1993

Version 3, March 26, 1993

TEI P2 is the first draft of one of the final work products of the TEI; the others will be a set of discipline-specific tutorials (TEI U1 et seq.) and one or more 'case-books' (TEI T1 et seq.) giving extended examples of TEI tags being used in various applications.

It contains three major parts: a prose specification, a reference section and full document type definitions (DTDs).

The prose specification is divided into 6 major sections, which are briefly summarized below. Each section contains brief descriptions of the tags proposed in a particular area, a formal summary of their proper usage with simple illustrative examples, and DTD fragments. The reference section is an alphabetic list of all elements, attributes and element classes defined in the TEI DTDs, giving a full definition for each together with a formal description of its usage. The whole is extensively cross referenced.

There follows a summary table of contents, as currently envisaged. Each chapter has an alphabetical abbreviation (shown in parentheses after the chapter title) that is useful in obtaining it from the TEI-L fileserver (on the fileserver, "P2" will precede the abbreviation.)[1] Chapters which have been published are marked with an asterisk and include section headings.

- Front Matter
- Part I. Introduction
  - About These Guidelines (AB)*
    - Texts and their Electronic Representation
    - Intended Applications
    - Structure of This Document
    - Notational Conventions Used in This Draft
    - Historical Background
  - Concise Summary of SGML (SG)
  - Structure of the TEI Document Type Declarations (ST)
    - Main and Auxiliary DTDs
    - Core, Base, and Additional Tag Sets
    - The TEI2.DTD File
    - Global Attributes
    - Element Classes
    - Other Parameter Entities in TEI DTDs
    - Invocation of TEI DTDs

---

[1] The first three chapters released use numerical instead of alphabetical identifiers in their file names: Character Sets (P221), the TEI Header (P222), and Base Tag Set for Transcription of Spoken Texts (P234).

-- Combining TEI DTD Fragments

- Part II: Core Tags and General Rules
  - — Characters and Character Sets (CH; also, P221)*
    - -- Local Character Sets
    - -- Shifting among Character Sets
    - -- Character Set Problems in Interchange
    - -- Writing System Declaration
  - — The TEI Header (HD; also, P222)*
    - -- Organization of the TEI Header
    - -- The File Description
    - -- The Encoding Description
    - -- The Profile Description
    - -- The Revision Description
    - -- Minimal and Recommended Headers
    - -- Note for Library Cataloguers
  - — Elements Available in All TEI DTDs (CO)*
    - -- Paragraphs
    - -- Ambiguous Punctuation
    - -- Highlighting and Quotation
    - -- Names, Numbers, Dates, Abbreviations, and Addresses
    - -- Simple Editorial Changes
    - -- Simple Links and Cross References
    - -- Lists
    - -- Notes, Annotation, and Indexing
    - -- Reference Systems
    - -- Bibliographic Citations and References
    - -- Passages of Verse or Drama
    - -- Arbitrary Segments
  - — Default Text Structure (DS)*
    - -- Divisions of the Body
    - -- Elements Common to All Divisions
    - -- Groups of Texts
    - -- Front Matter
    - -- Title Pages
    - -- Back Matter
    - -- DTD Fragment for Default Text Structure
- Part III: Base Tag Sets
  - — Base Tag Set for Prose (PR)*
    - -- Divisions of the Body
    - -- Contents of Prose Divisions
    - -- Front Matter
    - -- Title Pages
    - -- Back Matter
    - -- Specifying the Prose Base
    - -- Overall Structure of the Prose DTD
  - — Base Tag Set for Verse
  - — Base Tag Set for Drama
  - — Base Tag Set for Transcriptions of Spoken Texts (TS; also, P234)*
    - -- General Considerations and Overview
    - -- Overall Structure of Spoken Texts
    - -- Basic Structural Elements
    - -- Segmentation and Alignment

-- Recommended Transcription Practice
— Base Tag Set for Letters and Memos (LM)
— Base Tag Set for Printed Dictionaries (DI)
— Base Tag Set for Terminological Data (TE)*
-- The Terminological Entry
-- Tags for Terminological Data
-- Basic Structure of the Terminological Entry
-- Overall Structure of Terminological Documents
-- Additional Examples of Term Entries
— Combining Base Tag Sets (CB)

Part IV. Additional Tag Sets
— Segmentation and Alignment (SA)*
-- Pointers and Links
-- Multi-headed Pointers
-- External Pointers and References
-- Correspondence and Alignment
-- Aggregation and Virtual Elements
— Simple Analytic Mechanisms (AI)
— Feature Structure Analysis (FS)
— Manuscripts, Analytic Bibliography, and Physical Description of the Source Text (PH)
— Critical Editions (TC)
— Additional Tag Set for Language Corpora (CC)*
-- Varieties of Composite Text
-- Contextual Information
-- Associating Contextual Information with a Text
-- Linguistic Annotation of Corpora
-- Recommendations for the Encoding of Large Corpora

• Part V: Auxiliary Document Types
— Structured Header (SH)
— Writing System Declaration (WD)
— Feature Structure Declaration (FD)
— Tag Set Documentation (TD)
• Part VII: Alphabetical Reference List of Tags and Attributes
• Part VIII: Reference Material
— Full TEI Document Type Declarations (TD)
— Formal Grammar for the TEI-Interchange Format Subset of SGML (GR)*

-- Notation
-- Grammar for SGML Document (Overview)
-- Grammar for SGML Declaration
-- Grammar for DTD
-- Grammar for Document Instance
-- Common Syntactic Constructs
-- Lexical Scanner
-- Differences from ISO 8879