Augmenting Comprehension

Digital Tools and the History of Ideas

Proceedings of a conference at Bologna, 22–23 September 2002

Edited by Dino Buzzetti, Giuliano Pancaldi, and Harold Short

10

Diacritical ambiguity and markup

Dino Buzzetti

1. Adequacy of digital editions

The production of an edition in database form has recommended itself as a suitable solution for Gentile da Cingoli's *Commentary* to Porfiry's *Isagoge* (Buzzetti 1995: 145–54), a text written for didactic purposes at the University of Arts and Medicine in Bologna at around the turn of the thirteenth century (Buzzetti and Tabarroni; Buzzetti *et al.* 1992). All the examples of subsequent works of the same kind taken into consideration are characterized by a form of textual tradition that can be described as noticeably 'fluid', for reasons that apparently depend on the concrete forms of scholastic teaching practice. These very reasons in themselves make it historically preferable to reproduce the whole textual tradition than to reconstruct one single, improbable, authentic version of the text. Just as it is totally legitimate to say that the diffusion of certain doctrines was the outcome of the production of relevant texts, so it may equally well be maintained that the very form in which such texts were handed down was the outcome of their effective use in classroom teaching (Buzzetti 1997: 465). The choice of an edition in database form was therefore suggested by the need to obtain a form of representation suitable for textual traditions of a noticeably 'fluid' type (Buzzetti and Rehbein: 14–39). Generally speaking, it can be argued that only the analytical needs of research may provide, besides its heuristic motivations, the *criteria of adequacy* suitable to any possible digital representation of the text. A digital edition can only be wholly justified if it is capable of providing viable solutions to problems that can poorly be faced in any other way (Buzzetti 1999: 130).

2. Structure of information

In the case under consideration, a database not only provides an archive edition of all the existing witnesses of a certain textual tradition, but also organizes them within a coherent system of structured information that may afford precise answers to specific analytical needs. Similarly, inserting the digital images of the manuscripts offers not only a physical reproduction of the originals, which is pleasing to the eye, but also a different type of logical

representation of the textual information, which may be processed by means of appropriate automatic procedures. Each specific form of representation of the original reproduces the information contained in the document as *information structured* in different ways, which make it possible to carry out different analytical operations as required by each particular case.

The image of a document, for example, does not make its diplomatic transcription superfluous. The graphic information conveyed by an image is not the textual information conveyed by a transcription. Different ways of representing information afford different types of processing. A digital image (bitmap) may be considered a binary transcription of its visual content, and allows its graphic elements to be analysed. A diplomatic transcription explicitly represents some structural elements of the document and supports analytical procedures operating on these elements. Each distinct form of transcription may be considered a distinct form of analysis (Buzzetti 1995: 145-8). Each form of transcription picks out different structural elements, enabling different forms of analysis to be applied. Generally speaking, the possibility of applying certain analytical procedures depends on the way in which the information is organized by the form and structure of its representation.

For different reasons, both a philological reconstruction and a literary interpretation of textual production require *non-linear structural representations* of the text. A critical reconstruction of the text operates on the stratifications and multiplicity of the variant readings handed down in the course of its material transmission. The interpretation of any given edition of a text operates on the different ways in which its content can be understood and on the variety of its possible structural reconstructions. A philologist may reduce the variety of different linear textual readings to one single, complex, structural representation, whereas a literary critic may disclose a complex, non-linear whole of different structural interpretations of its content from one single linear representation of a text (Buzzetti 1996a: 87-9; Buzzetti 1996b: 225). Only a non-linear organization of the information can answer the analytical needs of textual and literary criticism (Buzzetti and Rehbein: 36-7). As highly abstract as it may be, this generalization is no less a basic and necessary condition of the adequacy of a digital representation of the text.

However, a philologist is concerned with variant readings, that is to say with structural elements of the *expression* of the text, while a literary critic deals with interpretative variants, that is to say with structural elements of its *content*. It is, therefore, necessary to consider how the structure assigned to the expression of a text can be related to the structure assigned to its content.

3. Processing structured information

The functionality of the digital representation of a text depends on the structure assigned to the information and on the operations that may be applied to it. Textual information is represented by means of linear sequences (or strings) of characters codified in binary form. *Markup*, that is to say the insertion of markers (or tags), makes it possible to assign a structure to the representation of the text (Raymond *et al.* 1992: 1–4), distinguishing different parts with different functions in the stream of codified characters; in their turn, the constituents of textual information may be organized in a *database* that describes and analyses their structural relationships. The structure assigned to the textual information, represented by the sequence of codified characters, will depend on the markup system. The operations that can be carried out on the elements of such a structure will depend on the database management system (DBMS). In order to verify the adequacy of a digital edition it is, therefore, necessary to consider how the markup system and the database management system are related to the digital representation of the text.

By necessity, markup must act on the expression of the text, while a database must proceed, by necessity too, quite independently of it, if it is to apply a totally formal and abstract operational model to the constituents of the content of the text. Hence it is necessary to avoid any discrepancy between the structure of the content and the structure of the digital expression of the text. In ther words, the *semantic model* of the content of the text must be applicable to the *syntactic structure* of its linear representation.

This prerequisite is not normally satisfied by markup systems that conform to SGML (Standard Generalized Markup Language). The formal grammar that regulates the use of tags admits only a sequential segmentation of the text; moreover, it admits only a hierarchy of relationships of subordination of its various elements to more comprehensive units, which are in their turn ordered in succession; thus it basically makes it possible to represent the structure of the text only as a *hierarchical tree structure*, whose elements, like the nodes subsuming them, are ordered in a linear way. For example, in a poem it is not possible to represent at the same time a segmentation by lines and a segmentation by enunciations or other grammatical constructions that may extend beyond the line itself: the syntax of an SGML-conformant system does not allow one to mark the several structural elements of an *enjambement* in a straightforward way, necessitating circuitous ad hoc solutions. In addition to such limitations in expressive power, there is a further peculiarity with much more serious and compromising consequences. SGML systems assume as a constraint language, i.e. as a formal language defining operational restrictions on the elements of the textual structure, the context-free grammar defined by

the DTD (Document Type Definition) of the SGML document, that is the formal syntax that regulates the use of its tags (Sperberg-McQueen and Huitfeldt 1999:-30). Thus, the restrictions placed on the syntax are, for no good reason, extended to the semantics by default: the need to assign only linear and hierarchical relationships to the structure of the expression leads to the inability to operate on the non-linear relationships of the structure of the content. Documents in SGML format may avail themselves of automatic procedures to check syntactic congruence, but do not make it possible to define operational procedures that can be applied to non-linear textual relationships (Sperberg-McQueen and Huitfeldt 1999: 41).

4. Markup and the structure of expression

What is the relationship between the structure of the expression and the structure of the content? What connection is there between the syntax and the semantics of the digital representation of the text? The analysis of the formal status of markup and of a database in relation to the digital form of the text makes it possible to outline a semiotics of its machine-readable representation and to establish the first elements of a *theory of the digital text*.

Markup is essentially notational (Buzzetti 1999: 146) and acts directly on the expression of the text; indeed, it may be considered a part of the very expression of the text, with respect to which it carries out a proper *diacritical* function. Markup is 'simultaneously embedded and seperable' from the text; it assigns a structure to the expression of the text and is itself a structure belonging to the text (Raymond *et al.* 1992: 1–4). Like punctuation or any other diacritical sign, it can be considered a metalinguistic description of the structure of the text, or an extension of the writing system itself, which makes it possible to render explicit those characteristics of the text that would otherwise be implicit. Markup can therefore be considered, respectively, as a form of metalinguistic notation or as an extension of the expressive resources of the object-language, i.e. of the very language constituting the text. Assuming the two different forms of notation to be effectively equivalent, one can conclude that markup is endowed with a *self-reflexive* function in relation to the text, and that it can be considered, respectively, an extension of the expression that makes its structure and rules of use entirely explicit, or an external form of reference to its functional and structural features. In the one case, the logical form of the language constituting the text must include forms of predication of the second order, that is forms of expression capable of representing structural aspects of its very working; in the other, the representation of its structural features is kept separate from the text, and a different language is used in order to be able to describe them.

The essentially notational nature of markup, furthermore, does not prevent it from representing structural features of both the content and the expression of the text. However, it is important to bear in mind how a form of representation necessarily connected with the structure of the expression of the text can describe structural features of the text that belong to its content and which do not generally depend on the structure of its expression. In fact, due to its very nature as a basically notational system, markup necessarily assigns a structure to the expression of the text even when it refers to structural features of its content.

5. Structure of content and database

A structural representation of the content of the text, independent of the structure of its expression, may, on the other hand, also be obtained by means of a database. A database is normally used as a form of abstract representation of our 'knowledge about the world' (Goldschlager and Lister: 248). It can, therefore, be taken as a structural and abstract description of the specific contents spoken about in the text; in short, as a structural *model* of its content. A database does not directly take into consideration the expression of the text and its structure. What addresses itself primarily to the expression of a document is markup; it is markup that expresses the abstract structural form of its format. A database, in contrast, does not take into consideration the formal structure of the representation of data, but describes the formal structure of its semantics (Raymond *et al.* 2004: 3–6). In other words, markup provides an abstract representation of the syntactic structure, or the format of data, whereas a database provides an abstract representation of the semantic structure, or the model of data, and implements a formalism operating on it (Joloboff: 75–6). If we refer, as in our case, to data sets comprised of sequences of codified characters, that is to a digital representation of textual information, it is possible to say that markup exhibits the abstract structure of the expression of the text and makes it explicit, whereas a database provides the model, or abstract representation, of the structure of its content.

It is, therefore, clear that the model of the text provided by a database is not affected by restrictions depending on the linear structure of its representation. A database can somehow be considered as a sort of deep, essentially *non-linear* structure, from which the linear, or surface, structure of the text may be generated (Buzzetti 1999: 147–8). In order to ensure functionality adequate to the digital representation of a text, it is necessary to avoid any constraint or mutual dependency between the form of the structural representation of the content and the form of the structural representation of the expression of the text. The markup system must be capable of projecting the structure of the content, which is not necessarily linear, onto the linear structure of the

expression, and the database management system must be capable of providing an operational model that could suitably associate the linear expression of the text with the formal representation of the several, and not necessarily linear, structural relationships of its content. The *congruence* between the linear structural properties of the expression and the structural properties of the content cannot be ensured, in general, by strongly embedded forms of markup, that is to say markup systems in which 'the position' of the tags 'in the data is information bearing' (Raymond *et al.* 1992: 3–4). Markup systems based on SGML do not normally satisfy this condition.

Is it, then, possible to describe in some way the structural relationship between the form of the expression and the form of the content of a text?

6. Instability and structural compensation

There is a form of *compensation* between the structure of the expression and the structure of the content of a text (Buzzetti 1999: 151ff.). For a given expression it is possible to assign different interpretations and contents, and different expressions of the text may correspond to a given interpretation and a given content. The fixity and invariance of the expression (or the content) reciprocally entail the variance and the indeterminacy of the content (or the expression). To put it briefly, there may be different ways of understanding what is said and different ways of saying what is meant. The equivalence of different *synonymous expressions* presupposes the identity of the same interpretive content, and, conversely, the compatibility of different interpretive contents calls for the identity of the same *polysemic expression*. Hence, the variation range of the content is constrained by the identity of the expression, and the variation range of the expression is constrained by the identity of the content. The indeterminacy of this relationship confers mobility and dynamism to textual structures, which may become stable as definite structural forms either of the expression or of the content, and may, reciprocally, determine the instability either of the corresponding content or of the corresponding expression. Consequently, the set of internal relationships between the constituent parts of the text remains mostly implicit, and the structure of the text may be defined as the 'set of latent relations' among its structural elements (Segre and Kemeny 1988: 44).

The dynamic processes of a text are of a holistic nature. In the relationship of compensation between determinacy and indeterminacy, diversity applies to the parts and identity to the whole. It is the whole expression of the text that refers, polysemically, to its different interpretations, and the different expressions that describe the content of the text refer, synonymically, to the whole. Hence, the expression in its entirety is compatible with different types of analysis and with different collocations of the parts of the model, just as the model, in its entirety,

is compatible with different divisions of the structure of the expression and different rearrangements of its parts. What invariably comes into play is the *relationship between the parts and the whole,* although each and every time what is assumed as a definite whole identical to itself is the content or the expression of the text, and what is respectively assumed to be an indeterminate structural partitioning of the whole is the set of its textual variants or the set of its interpretative variants. Thus, the parts of the text in which the phenomena of dynamic *instability* are to be found, that is its *critical zones,* are to be sought respectively in those structural forms of the expression and the content where a determined relationship between the whole and its parts comes about. The areas where the *rules* of the structural partitioning of the whole operate and are directly applied may be considered critical zones of the text. In the case of expression, the critical zone, the area of instability and indeterminacy, is to be sought in that portion of the object-language which conveys self-reflexive information about its own structure. In the case of content, the critical zone is to be looked for in the set of complex structural relationships that the operational rules determine among the constituent elements of the model.

With respect to the elements of the structure they define, rules have an ambivalent position. They set up the relationship of the single elements with the whole system, but they can be expressed in different ways and their status is basically ambiguous. In the case of expression, they can be expressed either metalinguistically, or through second-order object-language assertions; in the case of content, they can be expressed either as a specification of higher-order structural components and of the complex structural relationships they establish among the elements of the model, or as constraints and restrictions applied to the operations they define on such elements. The different ways of expressing the rules are equivalent and alternative and determine the place where the *ambiguity* of the structure becomes explicit. The structural relations of compensation between the parts and the whole of the two subsystems, expression and content, that make up the text in its entirety depend on the rules that define the structure of the text and on their operational constraints. It is, therefore, the expression of these rules and of their operational constraints that produces phenomena of structural instability and of dynamic ambivalence of the text. The compensatory relationships between the determinacy and indeterminacy of the structure of the expression and of the content of the text and the formal relationships between their abstract representations, consisting respectively in the markup system applied to the expression and in the database management system that describes its content, are hence defined by the alternative formulations of their respective rules of structural organization.

The primary distinction between expression and content is itself the result of an analysis of the unity which constitutes the text in its entirety. The

structural instability of the representation of the text is the immediate outcome of this distinction. Expression and content constitute two *subsystems* of the whole textual system; but once considered as separate units, they constitute two new distinct and related wholes. By analysing the expression, we determine its structure in relation to its integral whole. However, several ways of analysing its content correspond to any given and self-identical structure of the expression. By equating the identity of the text with the partial unit subsisting in mere expression, the other partial unit subsisting in its content remains indeterminate. A symmetrical and similar phenomenon comes about by equating the identity of the text with the other partial unit subsisting in its content. The phenomena of instability and indeterminacy of the structure of the text come about when we reduce the integral identity of the text to the identity of one of its two partial subunits.

7. Formal representation of the structure of the text

The forms of the *formal representation* of the structure of the two partial subunits can be constituted respectively by the markup of the expression and by the model of the content. The structural relationship between expression and content can thus be represented by the relationship between the markup system applied to the expression and by the database management system that provides a model of its content. The structural indeterminacy of the expression shows itself in the set of rules of the markup system and in its degree of freedom. The structural indeterminacy of the content shows itself in the set of operational rules of the database management system and in its degrees of freedom, that is to say in the intrinsic power of the formalism applied to the structural elements of the model.

A representation of the structural relationships among the elements that constitute the content of the text, that is to say an expression of the results of the operations applied to such elements, may be obtained through different *rules of linearization*. In this way, it is possible to obtain several descriptions, or representations, of the internal relations of the whole of the content. Likewise, we could assign different *logical forms* to a single linear representation of the content of the text by means of different markup schemes. It is thus possible to obtain several models of the internal relations of the whole of the expression.

That which expresses the internal relations of the whole of the expression can in turn be considered part of the expression, or something separate and external to it. Likewise, the rules governing the operations applicable to the elements that constitute the content of the text may either be described and defined outside the model, or may be represented directly within the model by specifying all the elements of the domain and co-domain that they respectively

correlate. This *ambivalence* in the forms of representation of the rules of structural articulation within the two subunits that constitute the text makes it possible to convert interpretative variants into textual variants, and vice versa.

The formal representation of the internal structure of the two textual subunits also allows the forms of compensation between the structural determinacy and indeterminacy of the text to become explicit, and affords the possibility of analysing the processes of dynamic instability in the textual structure. Such an explicit representation can be obtained thanks to the very specific nature of the *digital representation* itself. The semiotic properties of the primary meaningful constituents of the expression and the content of the text do not constitute elementary properties that cannot be further analysed, but they can be determined as a function of semiotic properties of even more elementary units of the data structures that represent them, for example single codified characters. A digital representation makes it possible to apply automatic and rigorous procedures to the processing and analysis of the minimal units which make up the more complex data structures that represent the semiotic properties of the primary meaningful constituents of the text (Samuels and McGann 1999: 35). A digital representation enables us to treat the causes of dynamic instability in a text not as fixed and invariable data, but as a controllable and analysable result of more elementary processes which determine the structure of the text and produce textual phenomena of higher complexity.

8. Formulation of the law of compensation

The formal representation of the structure of the expression and of the structure of the content of the text may be constituted respectively by the markup of the expression and by the database of the content. Being tied as it is to the expression of the text, markup may be considered as a way of representing explicitly in linear form complex internal relationships that are represented explicitly in non-linear form by a database that describes the content of the text. How is it possible to use these formal representations of the structural subunits of the text to express their *law of compensation*?

The law of compensation has been presented by Jerome McGann in the following form (McGann 1999: 81; McGann 2001: 175):

A = A <=> A =/= A22

This law expresses the paradoxical principle that any text *A* is not identical to itself. McGann derives this law from the relationship between the identity of the whole and the distinction produced by its primary partition, which George Spencer-Brown has expressed formally by means of the introduction of the notion of 'form of distinction' (Spencer-Brown 1969: 1) and its corresponding

laws. Applied to the text, Spencer-Brown's specification of the primary partition may be provided by the distinction of its primary subunits, expression and content. The 'indication' of the expression or the 'indication' of the content (*ibid.*) presuppose their distinction, produced by the primary partition operating upon the whole of the text. The 'indication' of the expression makes it a subunit of the text identical to itself and determines its structure. The determination and the identity of the expression with itself is expressed formally by the law of idempotence of the expression with respect to its representation — Spencer-Brown's first axiom, or 'law of calling' (*ibid.*), and 'form of condensation' (*ibid.*: 5). The 'indication' of the content makes it a subunit of the text identical to itself, and determines its structure. In the same way, the determination and the identity of the content with itself is expressed formally by the law of idempotence of the content with respect to its representation. But how can the *identity* of the text with itself, expressed as the idempotence of its partial subunits with respect to their representation, both depend on and at the same time be cancelled by the primary *distinction* that defines them?

It can be shown that the law of compensation presupposes and implies an endomorphism (*f*) between the structural constituents of the text (Buzzetti-1999: 156)

$$(A = A \iff A =/= A) \iff A \overset{f}{\longrightarrow} A$$

This endomorphism can set up a correspondence of elements of the content to elements of the expression, or conversely, of elements of the expression to elements of the content. The compensation between expression and content is represented by the *inversion* of the domain and co-domain of the endomorphism. The endomorphism between the structural elements of the text can be expressed by markup. The logical status of markup is ambivalent, and the *ambiguity of markup* constitutes the explicit expression of the inversion of the endomorphism. Each of the two partial subunits, expression and content, produced by the primary partition of the text can, as occurs, be assumed as the domain or the co-domain of the endomorphism. The inversion is produced by substituting distinct and equivalent forms of markup for one another, and the process of holistic compensation ensuing from the inner dynamics of the text shows itself explicitly in the oscillation between the equivalence and the functional distinction of the different forms of markup. I shall try to describe this phenomenon of compensation in greater detail.

Markup may be considered either as a *metalinguistic* representation of the structure of the expression, or as a direct, self-reflexive form of representation of the structure of the expression, consisting of second-order *object-language*

assertions (Buzzetti 1999: 152–3). In the latter case, markup assigns a second-order logical form to the structure of the expression, while in the former it assigns it a first-order logical form. The common reference to the integral unity of the text taken as a whole determines the equivalence of the two different interpretations of markup and brings about its *diacritical ambiguity*. In its turn, the *equivalence* of the two different interpretations of markup with respect to the integral unity of the text allows us to transfer the indeterminacy of one of its partial subunits to the other. The structure of the content can be referred to the form of the expression by means of out-of-line markup, that is a form of metalinguistic markup independent of the position of the tags in the sequence of codified characters (Raymond *et al.* 1992: 4). This enables us to assign to the expression of the text also non-linear and overlapping hierarchical structures. Distinct *interpretative variants* can thus be assigned to the structure of the text. It is, however, possible to transform interpretative variants into *textual variants* by substituting other equivalent forms of markup for this kind of metalinguistic markup, namely by substituting self-reflexive second-order object-language expressions for metalinguistic expressions. Conversely, by substituting equivalent forms of metalinguistic markup for self-reflexive second-order statements that describe the structure of the expression of the text within its own object-language, it is possible to transform textual variants into interpretative variants.

When applied to the integral unity of the text, the two different interpretations of markup are in fact equivalent and make the expression of the structural articulation of the text and its internal relations basically ambiguous. When applied, on the other hand, to each of the two partial subunits of the text, the two different interpretations of markup remain *functionally distinct* and produce the reciprocal determination and indetermination of expression and content. This can easily be seen. If we represent the *logical form* assigned by the markup to the expression of the text in *first-order* logic, such a form can exhibit only the structure of the expression as being completely independent of the structure of the content. The formal system for the treatment of a logical form of this kind may be *complete*, but it cannot be categorical. If it is complete, we can find a model that satisfies all its demonstrable assertions, but not all the models that satisfy the demonstrable assertions of the system are isomorphic. Hence, we have here a case of compensation. If we determine the form of the expression merely in relation to itself, the structure of the content is left indeterminate. On the contrary, if we represent the logical form assigned by the markup to the expression of the text in *second-order* logic, the formal system for the treatment of a logical form of this kind may be *categorical*, but it cannot be complete. This means that the logical form of the expression in this case also exhibits the structure of the content, because all the models that satisfy the demonstrable

assertions of the system are isomorphic, even if it cannot be proved on principle that all the demonstrable assertions of the system can be verified. Once again, the structural determination of the form of the content implies the indeterminacy of its expression. The law of compensation between the reciprocal determinacy and indeterminacy of expression and content may therefore be formally represented by means of the relationship between their respective representations, that is to say by means of the relationship between the logical form assigned to the text by markup and the model assigned to it by a database. The relationship between the *logical form* of the *expression* of the text (markup) and the *model* of its *content* (database) can in turn be considered an example of Spencer-Brown's second axiom, or 'law of crossing' (Spencer-Brown 1969: 2), and of the 'form of cancellation' (*ibid.*: 5). For, the reference of the structural articulation of one of the two textual subunits to the structural totality of the other cancels its identity with itself and brings about its indeterminacy. In conclusion, then, the identity of the text with itself is posited by the primary partition between expression and content, and is cancelled by the crossing from one subunit to the other, which revokes the separate identity of each determinate partial unit and reintegrates the indeterminate totality of the text. The text can be considered and described, in brief, only as a holistic unit.

In the light of the preceding considerations, it is clear why forms of strongly embedded markup depend on the structure of the expression alone (in the case of texts presented in SGML format, on the context-free grammars defined by their respective DTDs). Any attempt to project this structure onto the form of the content breaks the law of compensation. A context-free grammar related to expression alone is incorrectly used as a constraint language for a database management system related to the content of the text (Sperberg-McQueen and Huitfeldt 1999: 30).

9. Towards an analysis of the structural representations of text

As has been seen, in order to understand the phenomena of the digital representation of a text, it is essential to consider those of its constituent units that are of an even more elementary character either than the ultimate elements of the logical form assigned to its expression, or the abstract structural entities constituting the model of its content. The *expressive categories* and the *ontological categories* that constitute respectively the basis of the structural articulation of the expression and the content are in their turn the result of the aggregation of *more elementary constituents*. The values of the variables of the programming languages that can be applied respectively to the structure of the expression and the structure of the content are likewise the result of the composition of

codified data of a more elementary nature. Markup may be applied not only to the logical meaningful units of the expression of the text but also to the single codified characters that make them up, and the same is true for the description of the data structures to which the operational formalism of the database management system representing its content can be applied.

It is at this most elementary level of representation that the procedures for the automatic processing of textual information must be applied, and it is consequently at this basic level that the designing principles of systems for the representation and edition in digital form, or for the automatic analysis of texts, should be tested. Textual and literary criticism of the digital text must also take this path.

Bibliography

Buzzetti, D., 'Image Processing and the Study of Manuscript Textual Traditions', *Historical Methods*, 28:3 (1995), 145–54.

—, 'La faculté des arts dans les universités de l'Europe méridionale: Quelques problèmes de recherche', in Weijers, O. and Holtz, L. (eds), *L'enseignement des disciplines à la Faculté des arts (Paris et Oxford, XIIIe-XVe siècles)*, Actes du colloque international, Studia Artistarum, no. 4 (Turnhout: Brepols, 1997), 457–66.

—, 'Rappresentazione digitale e modello del testo', in *Il ruolo del modello nella scienza e nel sapere*, Roma, 27–28 Ottobre, 1998. Contributi del Centro Linceo Interdisciplinare 'Beniamino Segre', no. 100 (Rome: Accademia Nazionale dei Lincei, 1999), 127–61.

—, 'Il testo "fluido": Sull'uso dell'informatica nella critica e nell'analisi testuale', in Floridi, L. (ed.), *Filosofia & informatica*, atti del primo incontro italiano sulle applicazioni informatiche e multimediali nelle discipline filosofiche, Convegno Nazionale della Società Filosofica Italiana, Roma, 23–24 novembre 1995 (Torino: Paravia, 1996a), 85–93, <http://helmer.aksis.uib.no/allc-ach96.html>. Accessed, July 2004.

—, 'Digital Editions: "Variant Readings and Interpretations"', in *ALLC-ACH'96: Conference Abstracts*, 1996 ALLC and ACH Joint International Conference, Bergen, 25–29 June 1996 (Bergen: University of Bergen, 1996b), 254–6.

Buzzetti, D. and Rehbein, M., 'Textual Fluidity and Digital Editions', in Dobreva, M. (ed.), *Text Variety in the Witnesses of Medieval Texts*, Proceedings of the International Workshop, Sofia, 21–23 September 1997 (Sofia: Institute of Mathematics and Informatics of the Bulgarian Academy of Sciences, 1998), 14–39.

Buzzetti, D. and Tabarroni, A., 'Informatica e critica del testo: Il caso di una tradizione "fluida"', *Schede umanistiche 1* ns (2) (1991), 185–93.

Buzzetti, D., Pari, P., and Tabarroni, A., 'Libri e maestri a Bologna nel XIV ecolo: Un'edizione come database, *Schede umanistiche 2* ns (2) (1992), 163–9.

Goldschlager, L. and Lister, A., *Computer Science: A Modern Introduction* (London: Prentice Hall, 1982).

Joloboff, V., 'Document Representations: Concepts and Models', in André, J., Furuta, R., and Quint, V. (eds.), *Structured Documents* (Cambridge: Cambridge University Press, 1989), 75–105.

McGann, J., 'What Is Text? Position Statement', in *ACH-ALLC'99 Conference Proceedings* (Charlottesville VA: University of Virginia, 1999), 80–81.

—, *Radiant Textuality: Literature after the World Wide Web* (New York: Palgrave, 2001).

Raymond, D. R., Tompa, F. W., and Wood, D., 'Markup Reconsidered', paper presented at the First International Workshop on Principles of Document Processing, Washington DC, 22–23 October (1992), [1–20], <http://www.cs.ust.hk/~dwood/sgml-xml>. Accessed, July 2004.

—, 'From Data Representation to Data Model: Meta-Semantic Issues in the Evolution of SGML', *Computer Standards and Interfaces 10* (1995), [1–22], <http://www.cs.ust.hk/~dwood/sgml-xml>. Accessed, July 2004.

Samuels, L. and McGann, J., 'Deformance and Interpretation', *New Literary History 30* (1999), 25–56.

Segre, C., with Kemeny, T., *Introduction to the Analysis of the Literary Text*, trans J. Meddemmen (Bloomington and Indianapolis: Indiana University Press, [1985] 1988).

Spencer-Brown, G., *Laws of Form* (London: Allen and Unwin, 1969).

Sperberg-McQueen, C. M. and Huitfeldt, C., 'Concurrent Document Hierarchies in MECS and SGML', *Literary and Linguistic Computing 14* (1999), 29–42.

Thaller, M., Κλειω: *A Database System* (St Katharinen: Max-Planck-Institut für Geschichte i.K.b. Scripta Mercaturae Verlag, 1993).