



The Origins of Humanities Computing and the Digital Humanities Turn¹

Dino Buzzetti, University of Bologna

Abstract: At its beginnings Humanities Computing was characterized by a primary interest in methodological issues and their epistemological background. Subsequently, Humanities Computing practice has been prevalently driven by technological developments and the main concern has shifted from content processing to the representation in digital form of documentary sources. The Digital Humanities turn has brought more to the fore artistic and literary practice in direct digital form, as opposed to a supposedly commonplace application of computational methods to scholarly research. As an example of a way back to the original motivations of applied computation in the humanities, a formal model of the interpretive process is here proposed, whose implementation may be contrived through the application of data processing procedures typical of the so called artificial adaptive systems.

1. Introduction

A retrospective overview of the first stages of development of the newly emerged forms of reflection and methodological practices that, in Italian, have been properly named *informatica umanistica* (humanities computing) can also foster a better understanding of the new current trends. As a matter of fact, the limitations of the technological tools available at the time conferred more space to the ideation of what could have been achieved through the application of computational methods. The essential nature of the available technology focused the attention on the vast range of future opportunities enabled by the implementation of computational procedures on the model of the universal Turing machine. The origins of humanities computing are therefore characterized by a marked attention to the methodological and theoretical implications of research projects based on the application of computational procedures. The ensuing technological developments produced a rather paradoxical drawback. By polarizing the attention of scholars on the functionalities of application programs occasionally imposing themselves as dominant technologies, they induced a conceptual dependence on the available technology, to the detriment of a well-grounded choice of appropriate methods and alternative solutions. It may therefore be worthwhile to review the successive phases of

¹ This article was published in Italian: *Alle origini dell'Informatica Umanistica: Humanities Computing e/o Digital Humanities*, in *Il museo virtuale dell'informatica archeologica*, a cura di Paola Moscati e Tito Orlandi. Atti della "Segnatura" (13 dicembre 2017). «Rendiconti dell'Accademia Nazionale dei Lincei», Classe di Scienze morali, storiche e filologiche, S. ix, 30.1-2 (2019), 71-103.

applied computation characterized by new technological developments in order to evaluate their impact on the research practices of humanities computing and to reconsider its current orientations in the light of the theoretical discussions of its opening phase.

2. The era of the mainframes

The initial phase of humanities computing was characterized by employment of large computers available in computer centers, or public or private institutions that used them for administrative purposes. A “special thanks” for the intertextual analysis of the correspondences between two fundamental legal texts of the Jewish tradition, conducted by Sergio Noja in 1968, was addressed “to the Management of the S. Paolo Banking Institute of Turin,” who made “the electronic computer available,” which in turn made the publication of the essay presenting the results (Noja 1968, 582). The most famous example of this type of facility is that of the 56 printed volumes of the *Index Thomisticus* prepared by father Roberto Busa (1974-1980) thanks to the support of Thomas J. Watson, founder of IBM, and completed, after thirty years of work, only in the 70s.

It has surely been noted that both computational projects conducted by the *electronico IBM automato* resulted in printed publications. At first glance, in an era characterized by the ever-increasing pervasiveness of the digital, this seems surprising. However, the paradox is only apparent and, if well considered, leads us to conclusions of a different sort. This circumstance leads us to reflect on the working conditions that the mainframe technology then allowed. The memories, consisting of punched cards and only later of magnetic tapes, did not allow any form of visualization of the data, and the output resulting from their processing was returned in print format from the output units of the computers specifically used. Therefore, the purpose of computation could not consist in reproduction and visualization of digital data, or object sources of investigation, but only in the elaboration and analysis of their informational content. In this situation the purpose of the research was primarily directed to computation, that is to the application of computational procedures to objects, in our case, of humanistic research. Hence the name of humanities computing or, in Italian *informatica umanistica*, for the research practices of this first period.

The possibility of tackling research problems in the humanities with computational methods involved, at this stage and in the absence of technological mediation consisting of instruments already available and ready for use, a reflection directly addressed to the fundamentals of computation. IBM itself is likely to have supported Father Busa’s project for the opportunity it offered to extend the application of computation, up to that point almost exclusively aimed at processing numeric data, into the realm of processing textual information. Therefore, researchers did not simply have to make use of computational tools already set up, and choose those most suitable for the purposes of the research, but they had to contribute their own design, focusing attention on the specific aims of their own research.

3. A definition of humanities computing

Attention to “the fundamentals of computing theory and science, which today absolutely nobody in Humanities Computing mentions” (Orlandi 2016, 80-81), together with the constant reference to the objectives of the research, led in the first phase of development of humanities computing to a

definition expressed in the title of Nicholas Wirth's widely read book, *Algorithms + Data Structures = Programs* (Wirth 1976). In this regard, a volume that appeared at the conclusion of an investigation on the "impact of new technologies in the humanities" in Europe included a chapter, coordinated by Tito Orlandi, dedicated to study of formal methods, in which we find this specification – proposed by Manfred Thaller – of humanities computing as it was characterized in previous years and understood fundamentally as "applied computer science" (de Smedt 1999):

we will attempt to define the core in terms of the traditional combination of data structures and algorithms, applied to the requirements of a discipline:

- The methods needed to represent the information within a specific domain of knowledge in such a way that this information can be processed by computational systems result in the data structures required by a specific discipline.
- The methods needed to formulate the research questions and specific procedures of a given domain of knowledge in such a way as to benefit from the application of computational processing result in the algorithms applicable to a given discipline.

Crucial in this definition was the awareness that computation applied to the humanities requires both representation (*data structures*) and processing (*algorithms*) of the information contained in the objects of study, a requirement often overlooked in the subsequent phases of development influenced by advances in technology. This essentially *theoretical* characterization of humanistic information technology placed formalization in the foreground as a necessary and unavoidable prerequisite of research.

From this point of view, the "real beginnings" of humanities computing can be directly traced, together with other initial "experiments" to the seminal works of Jean-Claude Gardin (Orlandi 2016, 79) on "formalization of aspects of archaeological research connected with the processes of representation and classification of data" (Moscati 2013, 7). As Paola Moscati rightly notes (2013, 10), Gardin stated that "the interest of method [...] rather comes from its logical implications, and from the consequences it seeks to provoke in the general economy of the archaeological research" (Gardin 1960, 5); and later, in 1971, in a letter to René Ginouvès, Gardin reiterated that "[...] the comparative merits of such or such machine model or punch cards since 1955 worried us less than the methods of formalization (*mise en forme*) of the data and of the archeological reasoning, in the perspective of a 'mechanization' conceived without referring to any of these cards or these machines in particular" (JCG 205, 12 January 1971). What mattered more than the technology, therefore, was the formal organization of data. From this point of view Jean-Claude Gardin, "really is at the source of Humanities Computing" (Orlandi 2016, 82).

4. Representation vs. Data processing

Subsequently, the impressive technological development that occurred within a few years with the introduction of personal computers, graphic interfaces and then the implementation of the World Wide Web, profoundly transformed the research practices in the domain of the humanities computing, and greatly influenced the relationship itself between the representation and the elaboration of the content of the data under examination. The new and more advanced opportunities for practical use of computers, made possible by the progressive advancement of technology, have paradoxically

caused, if not quite a setback, at least one obvious delay in the theoretical elaboration necessary for the planning of applications specifically designed for specific research purposes. Once more Manfred Thaller, in the aforementioned volume, stated that, for an adequate training in humanities computing, “the study of computational methods themselves” was essential to the development of “new methods” for “the explanation,” framed “according to formal principles,” of the phenomena studied in the various humanities disciplines (de Smedt 1999). But in the new technological context, the center of the discussion was progressively moving from the investigation of the formalization of research methods and of the applicability of computation, to the evaluation of the possibility of using the new technological tools that gradually became available. The theoretical discussion about the design of applications specifically developed for specific projects of research and their intrinsic methodological implications thus fell into the background.

To give just one example, in the years in which exclusively textual terminals were still in use, the graphic layout of a document was decided by the author himself during the composition of the text with the insertion of the markup, or print instructions, in declarative or directly procedural form. Subsequently, with the introduction of graphical interfaces, writing programs working in WYSIWYG mode (What You See Is What You Get) automatically inserted the markup, removing from the author the direct control of the layout, which could be carried out exclusively in the ways provided by the program functionalities. The alleviation of effort in the process of composition was obtained at the cost of renouncing the direct design of the graphic characteristics of the document. This example, after all banal and certainly not relevant in terms of research, is nonetheless useful to point out certain consequences, often unnoticed, implicit in the development of technology, and to highlight some actual reasons for the progressive renunciation of the design of computer applications usable for research purposes, in favor of the passive and uncritical use of new instruments thrust upon us.

Also the very rapid spread of the Web has had profound consequences in the evolution of humanities computing. On closer inspection, the specific role of the computer in the practical use of the Web is quite limited, since it does nothing but guarantee remote access to data or documents stored at a distance, to be viewed on the computer screen. But the elaboration of the information content of the resources displayed remains entrusted to the reader’s ability to understand and, from this point of view, nothing changes. About the so-called “liberational effect of electronic technology on texts” Marilyn Deegan and Kathryn Sutherland have acutely observed that “the narrative of redemption from print” – anticipated by McLuhan and repeated emphatically by its followers of the 90s – did not foresee that

further, unimagined, developments in electronic technology, like the Google search engine, the brainchild of two college students, would lovingly extend the culture of the book through instant delivery of high-resolution images of the pages of thousands of rare and previously hard-of-access volumes. (Deegan and Sutherland 2009, 10).

Therefore, it is possible to assert that one “Google *Book* Search” – as they emphasize – “is not providing electronic text, it is providing books” (147). In other words, this common tool uses electronic technology to archive and instantly return “simulations of print and manuscript documents” (27). The elaboration of the informational content of the object of study is substantially neglected and the interest is turned to the digital reproduction of the source. That being said, we certainly do not

want to underestimate the importance of availability of multimedia resources. Digital images can, for example, be technically elaborated to improve the readability of deteriorated manuscripts, but the precedence given to the mere visualization of the sources profoundly modified the functionality of the fundamental link between representation and information processing. Attention came to be mainly addressed to the representation of information transmitted by objects of study, to the detriment of the elaboration of their contents for purposes of analysis and interpretation.

The lack of attention paid to the elaboration of information content of textual data can also be seen in the strategic choices of Text Encoding Initiative (TEI), with which in the 90s the community of scholars of humanities established “a standard for the representation of texts in digital form” (TEI 2016). The purpose of the TEI was indeed to publish *Guidelines for Electronic Text Encoding and Interchange*, in order to “define and document a markup language for representing the structural, renditional, and conceptual features of texts,” above all “in the humanities and social sciences” (TEI 2015). With the introduction of document production systems, among computer scientists, a so-called *document community* was formed, which took care of the automation, the visualization and printing processes of the documents, distinct from the community denominated in a similar way *data processing* or *database community*, dedicated instead to the design of archives for structured data. Now, while for the document community, in the interchange of data between different systems, it was essential to maintain the invariance of the representation of the documents, for the data processing community, it was instead fundamental to ensure the invariance of data processing operations. As a result, while the document community “chose to standardize the representation of data,” to guarantee its interchangeability, the database community “chose to standardize the semantics of data,” by developing “data models that described the logical properties of data, independently of how it was stored,” and regardless of the particular format of their representation. To tell the truth, even the “data semantics was not irrelevant to the document community, but the definition of semantics did seem to be a difficult problem” and the attempts undertaken proved too easily exposed to criticism (Raymond et al. 1996, 27). So, for quite similar reasons,

attempts to define semantics in the scholarly community, most notably the Text Encoding Initiative, similarly met with resistance. Thus, the route proposed by SGML was a reasonable one: promote the notion of application and machine independence, and provide a base on which semantics could eventually be developed, but avoid actually specifying a semantics. (28).

The technology of document management systems thus affected the choices for the digital representation of the text and led to the adoption of Standard Generalized Markup Language (SGML) as a standard language for the codification of textual data. As a language of simple representation and not data processing – because it lacks a semantics of its own – SGML represented a clear limit for the processing of textual data for the analysis of contents and the interpretation of texts. In the field of humanities computing, the prevalence assigned to the representation as opposed to the processing of information radically changed, in this phase, the prevailing orientations of the research.

5. Semantic Web and Digital Humanities

In the last and most recent period, the development of technology has had contrasting effects on the research practices of humanities computing. The Semantic Web project has brought back to the

fore the fundamental demand for the elaboration of the information accessible online. Languages have been developed for the representation of the content of Web resources, such as the Resource Description Framework (RDF) and for the construction of formal ontologies. Through the use of these languages, the so-called DIKW hierarchy (Data, Information, Knowledge, Wisdom), already in use in information science, could be treated more formally, allowing meaning to be assigned to data conceived as pure symbols that have not been interpreted and to represent the *information* conveyed by them through descriptions of the linked content between them. The networks and graphs of semantic relations thus obtained (linked data) made it possible to define connections rigorously and to organize certain fields of *knowledge* according to logically defined structural relationships and as such to allow the application of real procedures of formal inference. All this made it possible to bring attention back – through the tools provided by Semantic Web technologies – to the problem of processing the content of digital resources accessible online.

At the same time, and against this trend, the use of the expression *digital humanities* to define the customary field of humanities computing has imposed itself. The deliberate adoption of this name seems to be due to the preference expressed by the publisher Blackwell Publishing for a catchy title for its *Companion* for the introduction to the discipline (Schreibman et al. 2004). However, this has favored the tendency to comprehend under this definition all the phenomena in which the digital medium is used to disseminate contents related to the humanities. Even a simple e-book, or all applications for mobile devices designed for the access to multimedia concessions, thus seems to enter the field of interest of humanities computing.

The transition from *humanities computing* to *digital humanities* also comes explicitly theorized as a positive evolution of humanities computing. In fact, literary and artistic practice itself is more and more taking place in a directly digital form. In a recent interview with the online magazine *Il lavoro culturale* (Cultural Work) Jeffrey Schnapp, founder and director of metaLAB of Harvard University, claims to fully share “the point of view according to which a definition of digital humanities that reduced it to the application of a series of IT tools for the study of cultural heritage would be a relatively trivial operation,” and argues that

already in the 1990s, when the formula, Digital Humanities, was established in the United States, and we stopped talking about Computational Humanities or Humanistic Computing, we wanted to emphasize two aspects: the emergence of the Network as a public space and the personalization of the computer... The expression Digital Humanities marked precisely this moment of transition, in which the distinction between the world of digital technologies and culture in society no longer existed. This is a moment of unification in which there has certainly been a rethink on what research in the field of human sciences can be.

Consequently, *humanities computing* should give way to a “new experimental model of the human sciences,” to a new social practice of “Knowledge Design,” as opposed to nineteenth-century practice of philology (Capezzuto 2017).

In the face of all this, there is no lack of authoritative positional statements that in different forms instead recommend the opportunity for a return to origins. So John Unsworth, already in the title of an essay in which he draws coherent attention to the results achieved in the various phases of

humanities computing development, urges us to go beyond the simple digital representation of the primary research sources (Unsworth 2004):

we are, I think, on the verge of what seems to me the third major phase in humanities computing, which has moved from tools in the 50s, 60s, and 70s, to primary sources in the 80s and 90s, and now seems to be moving back to tools, in a sequence that suggests an oscillation that may repeat itself in the future. But whether or not the pattern ultimately repeats, I think we are arriving at a moment when the form of the attention that we pay to primary source materials is shifting from digitizing to analyzing, from artifacts to aggregates, and from representation to abstraction.

The exhortation to proceed beyond the simple digital “representation” of the studied documents, passing to the “analysis” of the content and to the “abstraction” necessary for the formal specification of computational procedures, is here quite evident, while the reference to “tools” should be understood as being functional with respect to the formalization of the methods adopted, and not as simple technological devices prepared in advance, regardless of the specific procedures applied, just placed “in the hands of researchers” for the computer-assisted practice of the usual activity of examining and annotating documents (Leon s.d.).

6. The “logicism” of Jean-Claude Gardin

Is it then possible to envisage the forms of this desired return to origins in our new context? The indications are not lacking and many inspiring principles can be drawn precisely from the illuminating anticipations of Jean-Claude Gardin. First of all, in the whole of his theoretical reflection, his rigorous methodological perspective and the reference to its necessary epistemological foundation take a central role. In a text published for the proceedings of a seminar held on 17 January 1994 at the University of Bologna, in which he presents the “research program” in which he was engaged “for thirty years” (Gardin and Borghetti 1995, 17), Gardin states that he is interested more than in “extending the field of application of computer science in the human sciences than in the progress and consolidation, with or without the calculator, of methodologies and their epistemological status.» (70). On the other hand, without giving priority to theoretical reflection on the possibility of applying computational methods to the humanities, one would inevitably run “the risk of confusion between the means and the ends of research (70),” and humanities computing would lose those “features of an autonomous intellectual project, with its own tools and goals” that actually characterize it (33).

Hence the proposal by Gardin of the “logician method” (30 ff) and “of the inevitable reference to epistemology” (19) that the application of this method necessarily involves. In the “analysis of archaeologists’ and historians’ texts” (17), or of the human sciences in general, considered “in their entirety” and as “constructs” (1980) or “scientific constructions” (Gardin and Borghetti 1995, 19), Gardin interest lies not so much, with Wittgenstein, in “erecting a building,” but rather “in having the foundations of possible buildings transparently before me” (18; see Wittgenstein 1998, 9). It is therefore necessary to face a problem of method and “practical epistemology” (Gardin and Borghetti 1995, 19), that is, of a type of epistemological reflection which he considers as “an activity whose purpose is to clarify the basic conceptual constructs of the human sciences, as they arise in practice, through the combined study of the symbolic systems that provide the materials, and the chains of operations that govern its architecture (Gardin *et al.* 1987, 29).”

The logicist method presented by Gardin consists therefore in the “study of the mechanisms and foundations of scientific argumentation” (Gardin and Borghetti 1995, 19) and in applying its principles to scientific «constructions» of the humanities, defined as follows:

I define ‘constructions’ the texts elaborated on the model of scientific works, with the following three components: (a) a set of observation facts or data ascertained on any type of foundation; (b) hypotheses or conclusions based on these data and which constitute the end of the construction, its reason for being; lastly, (c), the argument produced to link these two components: data to conclusions or, conversely, hypotheses to the facts, with modalities that can be of different nature: natural reasoning or common sense, mathematics, formal logic, computer science, or any type of conjunction of these instruments that is considered as distinctive of our intellectual procedures in the human sciences (18).

The application of the logicist method involves the use of “schematizations,” which in turn are defined – following “the logician J.B. Grize” (30) – as “the exercise that aims to isolate the operations called ‘natural logic’, currently practiced in the argument of ordinary language” (31; cf. Grize 1974, 204); the schematizations “are therefore nothing more than exercises in transferring into logical or, rather, semiological form, specialized texts in a particular discipline or field of research” (31).

The assignment of a logical form to the discursive arguments obtained through the schematizations “shows that every construction can be defined through the combination of two elements” (34), the “initial” propositions that describe the “facts” (35) and the “rewriting operations,” or the discursive passages, “whose sequence constitutes the reasoning” that leads to the “conclusions” (34), that is the propositions called “terminal” (35). The rewriting operations constitute real “logical operations that are in reality particularly diversified” (30) and depend on the peculiar principles of inference of the different “modes of reasoning” (93) that in the “discursive practices” (37) of the human sciences can take the most varied forms: “inductions, implications, abductions, inferences, deductions, etc.” (30) Now, “the two elements” of the schematization previously cited are found again unchanged in the structure of the knowledge base in the field of artificial intelligence” and “the organizations thus defined constitute” practically “the specific subject of expert systems or ‘knowledge-based systems.’” (35) The result was “the possibility of conceiving the schematizations as one source of knowledge for the elaboration of expert systems or, conversely, expert systems as a possible development of schematizations” (36); and “the computational paradigm” could become “the main tool” of the logicist analysis, that is, of that “rewriting modality which consists in expressing interpretative constructions in the form of chains of propositions that link observed data” to theoretical statements such as “in a calculation procedure” (Gardin 1993, 12).

The adoption of the computational tool, therefore, originates from a precise methodological choice and is based on the “homology” between the “architecture” of expert systems and that of schematizations (Gardin and Borghetti 1995, 36). Now, the more relevant aspect related to the switching to expert systems “concerns the ‘added value’” which is expected to be obtained “on the epistemological level” in which our scientific arguments are situated. But if, on the one hand, “the mandatory conversion of rewriting operations in reasoning ‘rules’ ” offers the possibility of “applying these rules in an experimental way, through simulations aimed at proving their validity”; on the other hand, “nothing allows us to affirm that our discursive practices can be assimilated to the true and

proper, and rigorously formal, rules of reasoning” (37). The computational methodological option thus requires an equally well-founded epistemological justification.

7. Epistemological reflection and expert systems

In response to the alleged “scientist infection” of which he would be a victim (Gardin 1993, 15) and to other criticisms of his logicist approach, Gardin presents a thorough picture of the contrasting positions in an attempt to arrive at an adequate characterization of the method of human sciences, without, however, refraining from considering “the limits and possibilities of logicism” (19). Among the different positions taken into consideration we can distinguish, on the one hand, those that presuppose an exclusive “dualism” between the expositional methods of the human sciences and those of the natural sciences (Gardin and Borghetti 1995, 18-19) and, on the other, the “middle positions” that insist on the intermediate nature of the human and social sciences “should one characterize them only by negations (‘neither this nor that’), or should one opt for an intermingling of categories (‘a little of this, a little of that’)” (Gardin 1999, 125). In this debate, the position of logicism “seems to be confused with that of the human sciences themselves, in that ‘entre-deux’ (Passeron 1991) where they intend to legitimize their location today,” unless they put in question the very definition of “this ‘third way’ of knowledge which, according to the some, would not be that of science or literature; according to the others, it would be neither that of the symbolic constructions, separated from the logic and ‘natural’ languages, nor entirely that of argumentation current in everyday life.” (Gardin 1993, 19).

Rather than following this debate in detail, it is important to note here that Gardin declares that he “feels uncomfortable in these intermediate spaces where the rules of the discursive game remain obscure” (1999, 125), even admitting in the end, that the “*substance de l’entredeux, elle, m’échappe*” (1991a, 32). Indeed, it is perhaps more important to observe that, today, both the epistemological reflection and the most up-to-date computational procedures actually seem to converge in offering a way out that addresses the issue Gardin leaves unresolved. In retrospect, its difficulty seems to depend upon the state of research in the field of expert systems at that precise moment and upon the availability of inference engines, which at the time were still too tied to the classical model of hypothetical-deductive reasoning, typical of the natural sciences. Now, however, a possibility of solution is in sight, in full compliance with the homology recognized by Gardin between the “logical form” (1995, 31) of the scientific constructions of the human sciences and the procedures of formalized inference of expert systems; and all this without overthrowing, however, from another point of view, the relationship of priority between the adoption of the logicist method and the “computer applications” that accompany it, but which – it should be reiterated – “do not constitute its main objective, nor its inevitable extension” (1993, 12). Therefore, it seems appropriate to pay attention to the possibility of a more detailed analysis of the interpretative and inferential practices of the texts expressed in natural language both in general and, for what concerns us more directly, in the field of human sciences.

8. Adaptive systems and methodological issues

In this regard, in an essay published on *Archeologia e Calcolatori* we find an affirmation that sounds almost surprising to those who usually rely on the classical deterministic paradigm of computation, but that is, however, of particular importance for our purposes since it is precisely based on the

“homology” already highlighted by Gardin between the “architecture” of the expert systems and that of the schematizations of scientific constructions expressed in natural language (Gardin and Borghetti 1995, 36). This essay is dedicated to the epistemological foundations of “adaptive systems,” whose theory developed thanks to contributions from “several fields” of research such as biology, cognitive sciences, and artificial intelligence (cf. Holland 1962, 297). Precisely at the beginning Massimo Buscema, the author of this essay, writes expressly, “I shall use an analogy to explain the difference” or, to say it in a better way, the relationship, “between artificial science and natural language; the computer is to the artificial sciences as writing is to natural language” (2014, 53). In other words,

in the artificial sciences, the computer is what writing represents for natural language: the artificial sciences consist of formal algebra for the generation of artificial models (structures and processes), in the same way in which natural languages are made up of semantics, syntax and pragmatics for the generation of texts. (2011, 17)

It follows that in the “artificial adaptive systems,” that are “part of the vast world of natural computation” – which constitutes in turn “a subset of the artificial sciences” – the functioning of the texts composed in ordinary language is assimilated, in an apparently unexpected way, to the algorithmic operation of the computer. What is illustrated here, in fact, is a homology between the forms of computation and the analysis of cultural phenomena, which amounts to a renewed proposal of the homology already theorized by Jean-Claude Gardin between the schematizations of scientific constructions in the humanities and the architecture of expert systems. In this sense, we can also read the definition of artificial science proposed by Buscema: “artificial sciences are those sciences for which an understanding of natural and/or cultural processes is achieved by the recreation of those processes through automatic models” (2013, 17). One could then almost say that Gardin’s logicism also fits, like the adaptive systems studied by Buscema, in the field of the so-called “natural computing,” which is, however, described as “the computational version of the process of extracting ideas from nature to develop computational systems” (de Castro 2006, 3). In both cases, the homology between discursive procedures and automatic systems works in reverse: while in natural computation the rules of the system adapt to the processes from which they are derived, in Gardin’s logicist analysis, the discursive schematizations are necessarily adapted to the formal rules of the expert system in use.

Among the expert systems examined by Gardin and the adaptive systems studied by Buscema there is therefore a crucial difference. What is characteristic of the adaptive systems is the presence of “rules that determine the conditions of possibility of other rules”; by their nature, these rules – formed by ‘constraints (links)’ which give the artificial models of natural processes the capacity to dynamically generate other rules – “are similar to the Kantian transcendental rules” and constitute the regulatory overarching principles on which the adaptive functioning of the system depends. In this way,

natural computation does not try to recreate natural and/or cultural processes by analyzing the rules which make them function, and thus formalizing them [statistically] into an artificial model. On the contrary, natural computation tries to recreate natural and/or cultural processes by constructing artificial models able to create local rules dynamically and therefore capable of change in accordance with the process itself. (Buscema 2013, 20)

Based on these considerations, the idea of building an adaptive model of this type – functional with respect to the analysis of texts expressed in natural language, in order to overcome the difficulty

encountered by Gardin in assigning a well-defined logical form to that intermediate form of argument that seems plausibly proper to all humanistic disciplines – looks quite legitimate, without thereby legitimizing non-clarified discursive rules and the residual mingling of models.

Even from a more general epistemological point of view, “this ‘third way’ of knowledge,” still viewed with suspicion by Gardin (1993, 19), has some plausible justifications. It is true that even the field of humanistic research can be subdivided in “two sectors,” one more strictly “governed by logic” and the other “governed by what you might call intuition”. It is also true that it is “difficult to subject intuition to scrutiny” for validity (Orlandi 2016, 79), an objection that Gardin frequently aims at positions that are prone to justify “the plurality and accumulation of interpretations” (1999, 119), without defining a precise criterion for validation. However, “a phenomenologically inclined cognitive scientist,” reflecting on the origins of cognition, might reason as follows:

We reflect on a world that is not made, but found, and yet it is also our structure that enables us to reflect upon this world. Thus, in reflection we find ourselves in a circle: we are in a world that seems to be there before reflection begins, but that world is not separate from us. For the French philosopher Maurice Merleau-Ponty, the recognition of this circle opened up a space between self and world, between the inner and the outer. This space was not a gulf or divide; it embraced the distinction between self and world, and yet provided the continuity between them. Its openness revealed a middle way, an *entre-deux*. (Varela *et al.* 1991, 3)

The recognition of this *entre-deux*, of this intermediate path between the self and the world, brings into play the fundamental relationship between the subject and the object of knowledge. Gardin also considers the problem posed by the “incisive formula,” frequently cited in epistemological debates, of the *retour en force du sujet* (1991b, 99); however – without going into the discussion of the complex relationship between model or subjective representation of phenomena and objective reality, or between observer and observed – he tends to treat the “subject” from a predominantly objective point of view and to deal above all with the “objective evaluation of the role of the subject in human sciences” (98). Suffice it to mention, however, that in the face of this, even in the natural sciences and especially in physics, the problem has been addressed in a direct way: “when a theory is highly successful and becomes firmly established, the model tends to become identified with ‘reality’ itself, and the model nature of the theory becomes obscured,” writes the theoretical physicist Hugh Everett, who thus goes on:

once we have granted that any physical theory is essentially only a model for the world of experience, we must renounce all hope of finding anything like “the correct theory.” There is nothing which prevents any number of quite distinct models from being in correspondence with experience (i.e., all “correct”). (Everett 1973, 134)

Also in physics, therefore, the ‘multi-interpretation,’ considered so problematic by Gardin, does not cause scandal and, except for the criterion of empirical conformity, the problem of choice doesn’t arise anymore. Then the question shifts rather to the formal reconstruction of the interpretative process in the discursive practices of the ‘third way’ mainly followed in the human sciences, whose characterizing element seems to be constituted precisely by a form of self-referentiality, which includes in itself the role of the observer. Thus, one understands the relevance of the processes of redefining their own rules which are typical of automatic adaptive systems. Formal analysis of self-referential

procedures of internal transformation imposes itself as the primary task of a research that can be extended – thanks to the analogy previously reported – to the interpretative practices of the texts expressed in natural language.

9. Ordinary language: formal model and natural computation

Which formal model can therefore be proposed for the representation and the formal analysis of the texts in natural language, which constitute the main product of scientific constructions in the humanities? Inspiration can come only from an analysis of language and from the perception of the enormous distance that separates the rigid “formalist’s motto” – characteristic of one of the most raw formulations of Good Old-Fashioned Artificial Intelligence (GOFAI),

If you take care of the syntax, the semantics will take care of itself (Haugeland 1985, 106)

and the compelling image of the connection between the text and its meaning offered by Samuel Beckett:

There are many ways in which the thing I am trying in vain to say may be tried in vain to be said.
(1965, 123)

This illuminating sentence dissolves with immediate naturalness the extreme trivialization of the relationship between syntax and semantics of the previous maxim. In the conception of good old-fashioned artificial intelligence, a formalization of the syntax should lead to an alleged one-to-one correspondence between the syntactic structure and the semantic structure of the text, an assumption which persists also philosophies of language of analytical orientation. As Davidson maintains, “to give the logical form of a sentence is, then, for me, to describe it in terms that bring it within the scope of a semantic theory” (Davidson 1970, 145).

The illusory postulation of this cherished one-to-one relation between syntax and semantics is completely debunked by the iconic representation of the fundamental indeterminacy of the relationship between the many ways of saying the same thing and the many ways to understand what is said by the same sentence: an identical content can admit different forms of expression, while an identical expression each time can be assigned different meanings. Here we encounter opposite conceptions of the relationship between the “expression” and the “content” of the text (see Hjelmslev 1961, 47-60). Following Saussure, who speaks of a plane of ideas (*plan... des idées*) and a plane of sounds (*celui... des sons*), (49) Hjelmslev states that an adequate description of the functioning of language “must analyze content and expression separately,” and that each of the two analyses may identify a certain number of entities “which are not necessarily susceptible of one-to-one matching with entities in the opposite plane” (46). On the one hand, the logicians, perhaps too conditioned by the symbolic character of the formal languages, are led to suppose that a syntactic system has “essentially the same structure as a semiotic” system and to consider it “normative for the concept of a semiotics.” On the other hand, for linguists it is the language which must be “considered as normative” for the functioning of a syntactic system (110). Accordingly,

the task of the linguistic theoretician is not merely that of describing the actually present expression system, but of calculating what expression systems in general are possible as expression for a given content system, and vice versa. (105)

In fact, “the two planes,” the syntactic one and the semantic one, “cannot be shown to have the same structure throughout,” with a “one-to-one relation” between the functioning of the one and the functioning of the other (112). Therefore, while a logician like Carnap proposes “a sign-theory where, in principle, any semiotic is considered as a mere expression system without regard for the content,” from a linguistic point of view the “formal” description “is not limited to the expression-form, but sees its object in the interplay between the expression-form and a content-form” (110-111).

However, it would be misleading to think that the radical difference between the two different conceptions of the relationship between the expression and the content of the text puts into question the possibility of establishing a functional homology between the discursive practices of ordinary language and the most advanced systems of artificial intelligence and natural computation. In Jean-Claude Gardin’s view, the formalization of the scientific production of the humanistic disciplines consists essentially in the formalization of the *discours savant* (1974, 57): in fact, the possibility of formalizing the textual phenomena in no way requires, as a necessary condition, a one-to-one correspondence between the syntactic structure and the semantic structure of the text. Rather, it is necessary to reflect on other characteristic aspects of the text and in particular on its diacritical or self-referential forms of expression.

Also, in this regard, however, the approach of the logicians and that of the linguists diverge. As Hjelmslev observes, “the logistic theory of signs finds its starting point in the metamathematics of Hilbert,” which considers the system of mathematical symbols only as “a system of expression-figurae with complete disregard of their content,” and which treats its “transformation rules,” or rules of rewriting as Gardin would say, “without considering possible interpretations.” The same method was then “carried over by the Polish logicians into their ‘metalogue’” and eventually “brought to its conclusion by Carnap” (1961, 110). In particular, Hjelmslev, who had defined language in general as “a semiotic into which all other semiotics may be translated,” (109) argues that

this is the advantage of everyday language, and its mystery. And this is why the Polish logician Tarski (who reached the same conclusion independently of the present author) rightly says that everyday languages are characterized in contrast to other languages by their ‘universalism.’ (1970a, 104-105)

For Tarski, on the other hand, rather than constituting an advantage,

it is presumably just this universality of everyday language which is the primary source of all semantical antinomies, like the antinomies of the liar or of heterological words. (1956, 164)

In fact, “one does not realize that the language *about* which we speak must not at all coincide with the language *in* which we speak” and if the semantics is elaborated in that same language, the analysis of antinomies shows that “the language which contains its own semantics and within which the logical rules commonly accepted apply must inevitably be inconsistent” (1936, 2). So while for Hjelmslev “owing to the universalism of everyday language, an everyday language can be used as metalanguage to describe itself as object language,” (1970a, 132) for Tarski “in contrast to natural languages, the formalized languages do not have universality.” (1956, 167). In fact, formal languages are developed as pure symbolic systems regardless of the content and for this reason, “when we investigate the language of a formalized deductive science, we must always distinguish clearly between the language *about* which we speak and the language *in* which we speak,” (167), between the “metalanguage” and the “language under investigation” (172). However, proceeding in this way, the normative relationship

between semiotic structure and logical structure is reversed and what is deemed to be an obstacle due to the self-referentiality of natural language is avoided through the sharp separation of the “metalanguage,” the language “to describe,” from the “object language,” the “language described” (Hjelmslev 1970a, 132). But it is precisely the search for the forms of expression of metalinguistic import within the natural language that can make our investigation proceed to the construction of a formal model of its self-referential semiotic system. To use Hjelmslev’s linguistic terminology natural language can actually be described as a “semiotics” that includes its “metasemiotics,” which is more specifically described as “a semiotics whose content plane is a semiotic” (1961, 114), also expressed in the natural language itself.

10. The markup: diacritical function and self-referential cycle

A useful starting point for this research can be found precisely by considering the current model for the digital representation of the text. It is well known that the text understood from a computational point of view as a data type of, that is, exclusively as “information coded as characters or sequences of characters” (Day 1984, 1), fails to represent all the information contained in the text understood in its current meaning. To solve this problem one resorts to the markup, whose standard form, accepted by the community of scholars of the humanities, consists in embedding, in the ordered sequence of the set of characters, marks or tags that define the properties of its partial segments or subsets. Now, if the markup represents textual information, the legitimate question arises about the status assumed by the markup in relation to the text. Thus, as Allen Renear puts it, one can enquire “about just what markup really is,” and in particular, “when it is *about* a text and when it is *part* of a text,” or in other words whether it belongs to the object language or the metalanguage of the text, without however excluding that “it may sometimes be both” (2000, 419).

Trying to arrive at a satisfactory answer, we can examine the case of punctuation. Alluding to the importance of this topic for the interpretation of the text, the title of a book dedicated to punctuation, *Eats, Shoots and Leaves* (Truss 2003), presents an interesting example: written with the comma the title means “eats, shoots and leaves” and can allusively describe the rude behavior of a young man invited to dinner by a friend; written without comma it means “eats buds and foliage” and may describe the eating habits of a panda. Now that comma, which completely changes the meaning of the sentence or of single words like shoots and leaves, can be considered, like any other another diacritical sign of the text, both as an element of the text, in that it is part of the writing system, and as an indication or a metalinguistic rule, as it prescribes the way in which the text must be interpreted.

It therefore can be affirmed that “punctuation is not simply part of our writing system,” but that “it is a type of document markup” (Coombs et al. 1987, 935). Along the same lines, the condition of the markup in general can be assimilated to that of a diacritical sign which, as such, has a double function: when it is used “to describe a document’s structure” (Raymond et al. 1992, 1) it carries out a metalinguistic function, but since it is expressed with “assigned tokens” which denote “specific positions in a text” (4) it constitutes itself the structure. The markup is therefore “simultaneously embedded and separable” from the text, is “part of the text, yet distinguishable from it” (3). Therefore, because it “*denotes* structure” in the text and at the same time it “*is* structure” itself (Buzzetti 2002, 80), the markup plays with respect to the text – in addition to “a properly diacritical function” – also “a self-reflexive function” and “can be considered, respectively, as an extension of the expression that

explains its structure” and the implicit rules that determine it, “as a form of external reference to its functional and structural aspects” (Buzzetti 2000). In short, “markup is at once *representation* and *representation of a representation*” (Buzzetti 2002, 81).

Because of its ambivalent nature, every form of diacritical expression generates a cyclic process (markup loop) within the textual dynamic:

we may say that an act of composition is a sense-constituting operation that brings about the formulation of a text. The resulting expression can be considered as the self-identical value of a sense-enacting operation. By fixing it, we allow for the indetermination of its content. To define the content, we assume the expression as a rule for an interpreting operation. An act of interpretation brings about a content, and we can assume it as its self-identical value. A defined content provides a model for the expression of the text and can be viewed as a rule for its restructuring. A newly added structure mark can in turn be seen as a reformulation of the expression, and so on, in a permanent cycle of compensating actions between determination and indetermination of the expression and the content of the text. (Buzzetti and McGann 2006, 68)

All this can also be appropriately expressed with a diagram (Fig. 1). It is worthwhile to pause and consider in more detail some of the formal aspects both of the cycle and of the diagram that represents it.

The diagram refers in particular to the markers that complete the digital representation of the text and that can be inserted inside it, or be made up of external elements connected through pointers to certain positions in the linear sequence of the characters. Since there is no direct correspondence between the elements of the syntactic structure and the elements of the semantic structure, the internal (embedded) markup – as it is part of the sequence of characters and it forms itself its structure – diacritically describes syntactic and expressive properties of the text. The external markup (stand-off), on the other hand, not being bound to the linear structure of the expression of the text, can freely express aspects that are not necessarily linear in the structure of its content. In the multidimensional diagram of the self-referential cycle of the text, there is therefore a correspondence between the dimension of the expression and that of the internal markup, as well as that of the content and external markup.

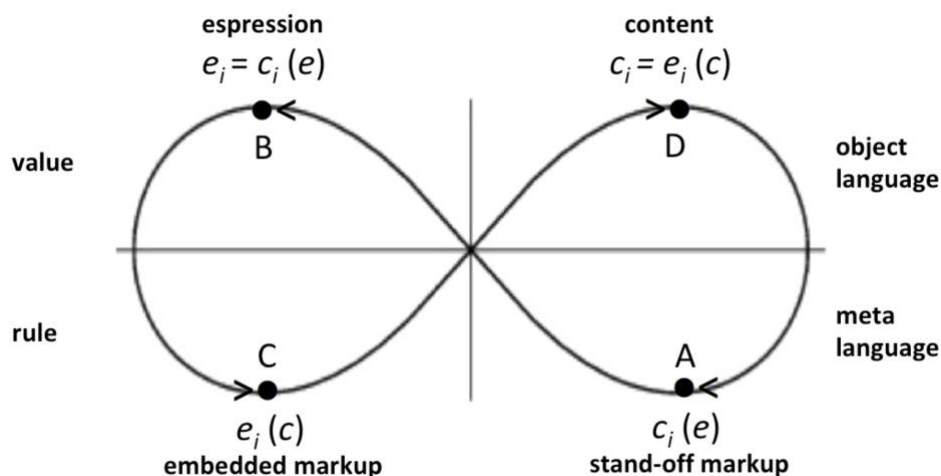


Figure 1. The markup loop (Buzzetti and McGann 2006, 68).

The dual linguistic and metalinguistic function of the markup owing to its diacritical nature means that the same marker is both a self-identical element of the expression of the text, and a rule that determines the structure of the content and defines its specific elements, which in turn behave in the same way with respect to the expression. Therefore, the structural diacritical elements of expression and content can be considered both as the result of a restructuring operation and as the operations themselves that determine, respectively, the organization of the structure, both of the expression and of the content of the text. Formally, therefore, they can be understood as values of a function, or as the functions themselves which formally represent the rules for structuring the text.

The relationship between the formal representation of the value of the function and the formal representation of the function or the rule itself, deserves to be carefully considered from the logical point of view, in order not to run into serious confusion between the linguistic and the metalinguistic levels present in natural language. In his careful analysis of ordinary language use, Gilbert Ryle appropriately cautions against easy “category-mistakes” (1949, 17) which one incurs if one does not pay attention to the “logical type or category,” (16) of commonly used expressions. As for what concerns us, Ryle observes that “a ‘variable’ or ‘open’ hypothetical statement” (120) – that is, a propositional function that contains variables, and all propositions of this type that express law-statements or a rule – “belong to a different and more sophisticated level of discourse from that, or those, to which belong the statements of the facts that satisfy them” (121). These propositions therefore constitute real rules of inference, for

a law is used as, so to speak, an inference-ticket (a season ticket) which licenses its possessors it to move from asserting factual statements to asserting other factual statements (121).

So the rules of inference, and with them the diacritical expressions we are dealing with, can be considered as statements of a higher order that belong to the logical type of the “inference-licenses,” studied by Stephen Toulmin in *The Uses of Argument* (2003, 91), who by his own admission “owes

much” to Ryle’s ideas, which he also “applied to the physical sciences” in his own *Philosophy of Science* (2003, 239).

In his review of this work, Ernest Nagel (1954) observes that, thanks to the so-called deduction theorem, the principle, now “canonical in modern logical theory” (405), that “a rule of inference can in general be replaced by a premise,” holds in the case of our inference-licenses, and that “in the case of material rules of inference,” consisting of true non-tautological propositions like the ones we are dealing with, “this can apparently always be done.” Nagel tells us too that this “maneuver” can also “be introduced in reverse” (406). This means, according to standard logic, that the same sentence can act both as a first-order asserted premise in the object language, and as a rule of inference in the metalanguage. One should note that while in logic the object language and the metalanguage are necessarily kept separate – and are made up of statements respectively in the “material” and in the “formal mode of speech,” to use the terminology introduced by Carnap (1934), or by statements *de re* and statements *de voce*, to use a terminology drawn from the use of medieval logical Latin (Henry 1984), a technicized, but still natural language – in the case of ordinary language, which contains its own metalanguage, inference rules are expressed by statements. Such object-language higher-order *de re* statements are however inferentially equivalent to first-order *de voce* statements expressed in an external metalanguage, separated from the object language. Consequently, in natural language, self-referential object-language diacritical expressions take on a double function: considered as first-order statements, they are used as structural markers both of the expression and the content of the text; whereas considered as second-order statements, they constitute rules of inference which are used as functions of the expression to determine the structure of the content or, conversely, rules of inference used as functions of the content to determine the structure of the expression.

11. Generalization of the model

Still on a formal level, we can observe that the structure of the markup cycle, represented above – which can however be generalized for all forms of diacritical expression – corresponds exactly to the “conversational cycle,” which according to Frederick Parker-Rhodes represents the actual “speech process” between the speaker and the listener (1978, 16) or, dealing with texts, between the writing and the reading of a text (Fig. 2).

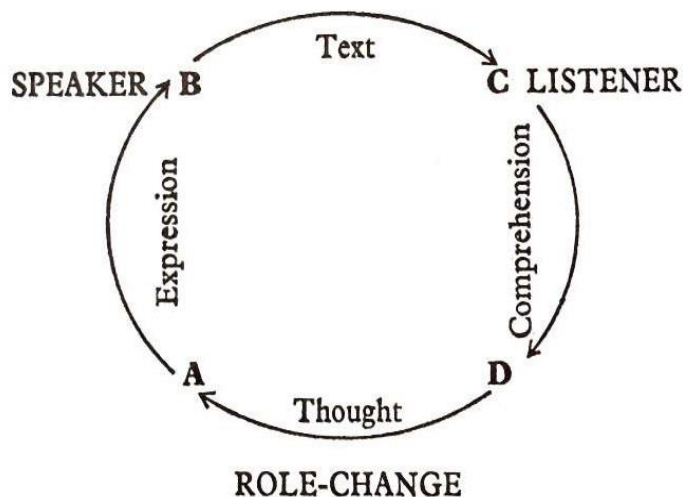


Figure 2. The conversational cycle (Parker-Rhodes 1978, 16).

In this cycle, the “expression” (A) is an operation performed by the speaker “which takes a ‘thought’ as input (which we must think as formalized in some manner)” and produces a “text” (B). One should note, incidentally, that here by expression we mean an operation, which is a function of the content, and not its result, a fact that proves the ambivalence of the diacritical mark on which it operates. In turn, the “comprehension” (C), or interpretation, is an operation performed by the listener, who receives the text as an “input containing all the information imparted to it by the speaker” and that produces “again a thought” (D) as its “output” (17). It is clear, regardless of the use of a different terminology, that the structure of this cycle corresponds exactly to that of the previously examined markup cycle (Fig. 1).

However, an important observation by Parker-Rhodes should not be overlooked. It explicitly refers to the indeterminacy of the interpretation process: the “thought that the speaker had intended to convey,” once received and interpreted in the mind of the listener, “could produce the elaboration of a new thought” as a “result” (17). In this case the diagram could take the form of an open spiral, which is more suitable to represent the case of several possible interpretations (Fig. 3).

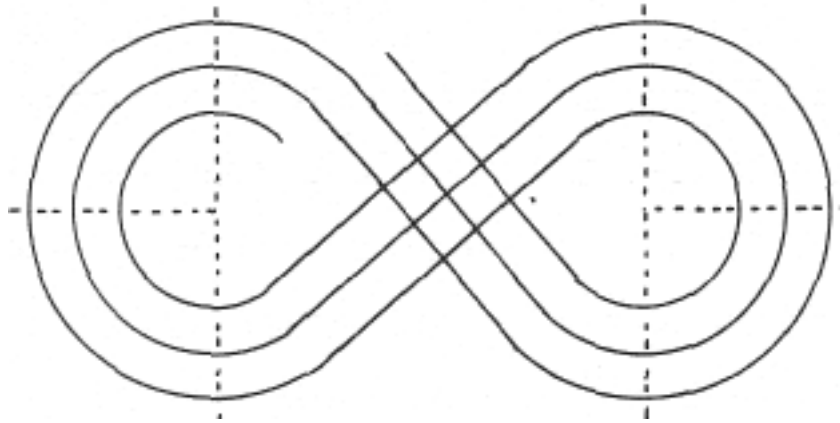


Figure 3. The helicoidal cycle (Gardin 1980, 45).

Such a cycle could end at some point, returning to the starting point, or proceed indefinitely, depending on the context in which a given expression of the text is received. Jean-Claude Gardin also recognizes as “self-evident” the “cyclical nature” of the process of scientific construction. However, similarly to Parker-Rhodes, he believes that the cycle is not necessarily closed and therefore can be best represented by a “helicoidal curve,” more suited to retrace “the successive steps of its formation,” which are produced through a series of choices, depending not only on the data and their organization, but also and above all on the “logico-semantic rules of interpretation” and the different “interpretative models” that are equally possible (1980, 145).

12. Epistemological foundations

The analysis of the cyclical nature of discursive practices brings us back again to the question of its epistemological foundation. As Gardin observes, the process of scientific construction can be considered both “from within,” and “from without,” or in other terms, subjectively from the author’s point of view, and objectively from the point of view of those who examine it, as an alternative to other constructions, in order to express a judgment of “validation” (145). This allows us to better evaluate the intermediate nature of the humanistic disciplines’ methodology, which many locate in the *entre-deux* between the predominantly objective nature of the methods of the natural sciences and the predominantly subjective nature of literary or discursive production in general. In other words, one has to decide whether this *entre-deux* divides or joins the two points of view, establishing what relationship exists between the subject and the object, or else the subjectivity and the objectivity absolutely considered.

The cyclic and self-referential nature of the discursive process, which in the ordinary language form of expression jointly includes both the representation of its own object, and the representation of the way in which the subject represents it, inclines towards an answer that excludes the absolute separation between the subject and the object, or in other words between the observer and the observed. This is the position embraced, for example, by the theorists of autopoiesis (Varela et al. 1991), that draw inspiration from the epistemological discourse of Maurice Merleau-Ponty and his notion of “chiasm.” In one of his most iconic descriptions, Merleau-Ponty presents the chiasm as

an exchange between me and the world, between the phenomenal body and the “objective” body, between the perceiving and the perceived: what begins as a thing ends as consciousness of the thing, what begins as a “state of consciousness” ends as a thing exchange between me and the world, between the phenomenological body and the ‘objective body’ between the perceiver and the perceived: what begins as a thing ends as consciousness of the thing, what begins as a “state of consciousness” ends up as a thing. (1968, 215)

In his essay on *La structure du comportement*, in order to clarify the connection between the subject and the object, Merleau-Ponty, once again, cites (1942, 11) the physiologist Viktor von Weizsacker, who describes that relation in these terms: “the properties of the subject and the intentions of the subject (...) not only mix with each other, but also constitute a new whole” (1927, 45). This means that the subject and the object must not be conceived as separate, but as constantly connected in a continuous process of “overlapping or encroachment (*empiètement*)” (Merleau-Ponty 1968, 123), as if one would over and over again take the place of the other. The chiasmic interlacement thus consists of a relationship of “activity and passivity coupled,” (261) a representing and being represented of the subject and the object both in language and perception. Thus, the understanding of the “chiasm,” as described by Merleau-Ponty, leads to the conclusion that language, understood as natural language, “is *the same*” thing that simultaneously represents and is represented, but not the same “in the sense of real identity,” but rather “the same in the structural sense,” that is, in the sense of a unique and self-identical semiosis, which also includes the semiosis that represents it (261).

The same relationship between the subject that represents and the object being represented, when conceived as ‘the same thing,’ that is, as the ‘new whole’ that they constitute, is found in the notion of the subject proper to the cybernetics “of the second order,” the cybernetics of the “observing systems,” in which “the observer enters the system by stipulating *his own* purpose,” as opposed to the cybernetics of the “observed systems,” or “first-order” cybernetics, in which “the observer enters the system by stipulating the *system’s* purpose” (von Foerster 2003, 285-286). Thus, in this context, one can find this enlightening definition of the subject: “I am the observed relation between myself and observing myself.” (257). Here the subject is defined as one and the same thing, a new whole, constituted by the representation of the relationship between the self observing itself and the self observed by itself.

Hence the idea that the conception of systemic self-referentiality – that takes place, for example, both in natural language and in its formal model – could constitute a new fundamental scientific paradigm. A new paradigm of this kind necessarily leads one to believe that the nature of the human sciences can be considered as an intermediary one, only as long as the natural sciences and the literary or the discursive disciplines in general are conceived of as absolutely separate and incompatible with each other. However, the recognition of the unavoidable relationship between the observer and the “systems observed,” or the principle of the autonomous organization of the “observing systems,” today extends manifestly beyond the field of the disciplines characterized by the interpretive method to the field of the physical and biological sciences. Thus, the paradigm of self-referentiality seems to open a new perspective of convergence between the methods of the natural sciences and the methods of the human sciences, whose median nature would then be based more on the nature of the object of the investigation than on the specific nature of the method whereby knowledge is constructed.

13. Subjectivity and objectivity: formalization and implementation

At this point, our diagram of the self-referential cycle of the discursive process can be reconsidered, taking into account the reflexive character of the relationship between the subject and the object. Language, in as much as it is seen as expression, is subjective, because it is the representation of the form of our act of representing; however, in as much as it is seen as content, language is objective, because it is the representation of the form of what it represents. In turn, a form of diacritical expression of the text, subjective in itself, can be considered both from an objective point of view, as an element of the expression identical to itself, and from a subjective point of view, as a function that determines a structural element of the content (Fig. 4). The same can be said of an element of the content: objective in itself, which can be considered both from an objective point of view as an element identical to itself, and from a subjective point of view as a function that determines a structural element of the expression.

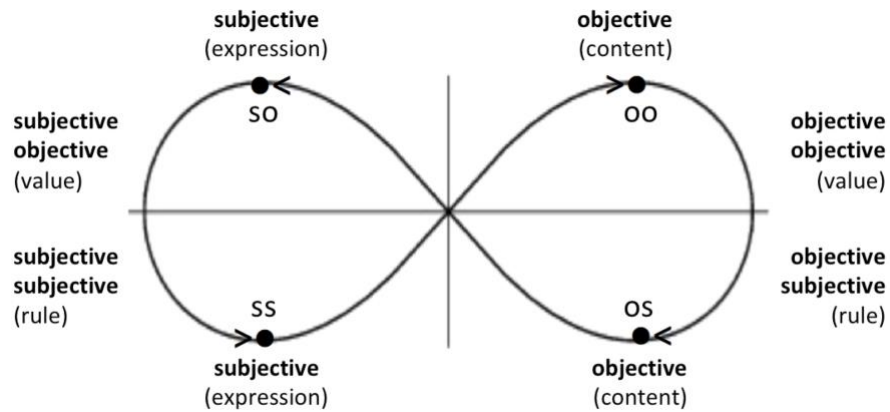


Figure 4. Subjectivity and objectivity in the speech process (Parker-Rhodes 1978, 16).

The distinction between something subjective and something objective is therefore a recursive distinction that could continue indefinitely (Fig. 5):

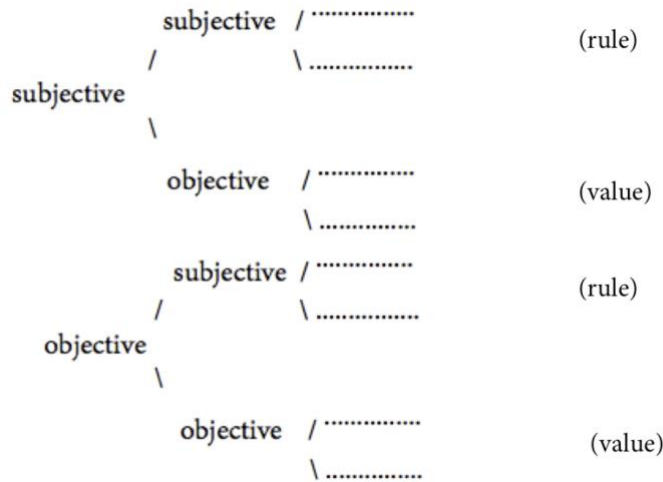


Figure 5. Recursiveness of the subjective/objective distinction.

But this does not happen, for the very reason that language is self-referential, as one can clearly infer from the diagram shown in Fig. 6:

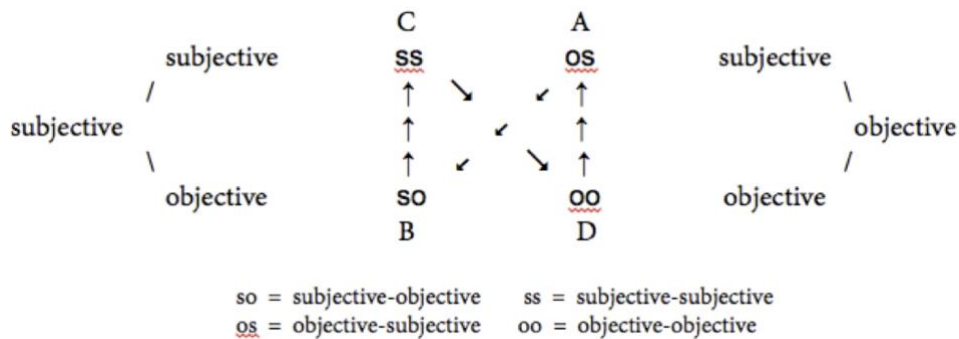


Figure 6. Chiastic self-referentiality of the subjective/objective distinction.

It can be reasonably assumed that this scheme represents a possible formal model of the ‘chiasm,’ that is, of the relationship between the subject and the object that involves a continuous process of mutual “encroachment, infringement (*empiètement, enjambement*)” (Merleau-Ponty 1964, 175), or reciprocal displacement, dismissal, and override. An image of continual oscillation between what is subjective and what is objective, as an uninterrupted process, is aptly allusive of the self-referential mobility of the text and the dynamic nature of the ambivalence of the diacritical structural elements of both the expression and the content of the text.

This is an aspect which cannot be described only metaphorically, but which finds formal expression also in rigorous mathematical terms. As David Hestenes writes about the mathematician who introduced the algebras that bear his name,

Clifford may have been the first person to find significance in the fact that two different interpretations of number can be distinguished, the *quantitative* and the *operational*. On the first interpretation, number is a measure of “how much” or “how many” of something. On the second, number describes a relation between different quantities. (1999, 60)

In other words, seen from the latter point of view, a number describes the operation that connects two different quantities. The same concept of ambivalence between value and function can be found in the “calculus of indications” introduced by the English mathematician George Spencer Brown (1969, 11): he admits that there can be a “partial identity of operand and operator,” since an operand “is merely a conjectured presence or absence of an operator” (88).

Granted that a rigorous formalization of such a model is possible, it can be surmised that a computational implementation could be obtained by developing a suitable adaptive system, endowed with functional capabilities as those previously illustrated. If the adaptive artificial systems that we described are built on the basis of a recognized analogy with the operation of natural language, that is, foreseeing the presence of rules capable of modifying other rules of the system, the same analogy allows us to suppose that a formal model of the discursive processes of natural language could be implemented precisely using an adaptive computational system of the same type.

In fact, the ambivalence of precise mathematical objects strictly defined (operation and operand, function and value), can constitute the formal expression of the relationship between subject and object that we have described by recalling the epistemological notion of the ‘chiasm.’ Secondly, it is precisely the indeterminate character of the relationship between syntax and semantics in the natural language that gives origin to the self-referential cycle of “rules” of the second order “that establish the conditions of possibility of other rules” of the system (Buscema 2013, 20). Thus, in this way, the road is open for the possibility of implementing a computational model of the discursive processes proper to scientific constructions in the humanities, consisting in an automatic system of an adaptive type.

14. Conclusions

This concludes our long, extended argument aimed at supporting the opportunity for a return to the origins of humanities computing to avoid the risk, of which Jean-Claude Gardin has made us aware, of exchanging means for research purposes. The period of the origins, or the so-called humanities computing, was distinguished by an attitude aimed primarily at reflecting on the methods and their epistemological foundations, as a preliminary condition to the choice of computational means suitable for the solution of the research problems of a specific disciplinary field. Subsequently, subordination to the rapid technological development of the 1990s, favoring the importance of the digital medium in artistic and literary production, has actually reversed this relationship. The priority given to practices of cultural production directly in digital forms and to research activities assisted by the computer, although still conducted in all-traditional forms, has produced a veritable mutation of the humanities computing practice of the original period and led to the advent of the so-called digital humanities. Thus, interest has abated in what Jerome McGann considers, in this new digital environment, the urgent and very current philological “imperative” of the “preservation of cultural memory” (2012), in agreement with the famous definition of August Boeckh, *die Erkenntnis des Erkannten*.

In the second part of the essay, I therefore tried to present, through an example, a form of restoration of the original humanities computing attitude to deal with theoretical and methodological issues in regard to the interpretation of texts. Building on the thoughts of Jean-Claude Gardin on the analysis of discursive practices in the human and social sciences – in particular, on the homology between the structure of knowledge base expert systems and the structure of the schematizations of data and argumentations in humanities scientific texts – I have noted a significant convergence, or rather a substantial homology, between the analysis of the self-referential phenomena of natural language and the establishment of data processing rules in automatic adaptive systems. Thus, this correspondence has allowed me to outline a formal model for the analysis of the interpretative practices of texts in ordinary language congruent with data processing procedures proper of adaptive systems. A possibly successful implementation of this model would undoubtedly confirm the fecundity, for humanities computing, of a re-proposal of the priority of the theoretical and methodological reflection that particularly characterized the period of its origins.

translated by Massimo Lollini

Works Cited

- S. BECKETT, *Proust and Three Dialogues with Georges Duthuit*, London 1965.
- R. BUSA S.J. (ed.), *Index Thomisticus: Sancti Thomae Aquinatis operum omnium indices et concordantiae in quibus verborum omnium et singulorum formae et lemmata cum suis frequentius et contextibus variis modis referuntur quaeque, auspice Paulo 6 Summo Pontifice, consociata plurimum opera atque electronico IBM automata usus digessit Robertus Busa*, 56 voll., Stuttgart-Bad Cannstatt 1974-1980.
- M. BUSCEMA, *Artificial Adaptive Systems: Philosophy, Mathematics and Applications*, in M. Buscema, M. Ruggieri (eds.), *Advanced Networks, Algorithms and Modeling for Earthquake Prediction*, Aalborg 2011.
- M. BUSCEMA, *The General Philosophy of Artificial Adaptive Systems*, in M. Buscema, W.J. Tastle (eds.), *Intelligent Data Mining in Law Enforcement Analytics: New Neural Networks Applied to Real Problems*, Dordrecht 2013.
- M. BUSCEMA, *The General Philosophy of Artificial Adaptive Systems (AAS)*, in M. Ramazzotti (ed.), ARCHEOSEMA. *Artificial Adaptive Systems for the Analysis of Complex Phenomena*. Collected Papers in Honour of David Leonard Clarke, “Archeologia e Calcolatori,” Supplemento 6 (2014), pp. 53-84. http://www.archcalc.cnr.it/indice/Suppl_6/04_Buscema.pdf [11/04/2018].
- Buzzetti, Dino. “Digital Representation and the Text Model.” *New Literary History*, vol. 33, no. 1, 2002, pp. 61–88. doi:10.1353/nlh.2002.0003.
- D. BUZZETTI, *Ambiguità, diacritica e Markup: Note sull’edizione critica digitale*, in S. Albonico (ed.), Soluzioni informatiche e telematiche per la filologia, Atti del Seminario di studi (Pavia, 30-31 marzo 2000), Pavia 2000. http://studi.umanistici.uni-pv.it/diplamm/pubtel/Atti2000/dino_buzzetti.htm [11/04/2018].

- D. BUZZETTI, J. MCGANN, *Critical Editing in a Digital Horizon*, in L. Burnard, K. O'Brien O'Keefe, J. Unsworth (eds.), *Electronic Textual Editing*, New York 2006, pp. 51-71.
- S. CAPEZZUTO, *Il design della conoscenza: Intervista a Jeffrey Schnapp*, "Il lavoro culturale" (2017). <http://www.lavoroculturale.org/intervista-a-jeffrey-schnapp/> [11.04.2018].
- R. CARNAP, *Logische Syntax der Sprache*, Wien 1934.65
- J.H. COOMBS, A.H. RENEAR, S.J. DEROSE, *Markup Systems and the Future of Scholarly Text Processing*, "Communications of the ACM" 30: 11 (1987), pp. 933-947.
- D. DAVIDON, *Action and reaction*, "Inquiry" 13:1-4 (1970), pp. 140-148.
- A.C. DAY, *Text Processing*, Cambridge 1984.
- L.N. DE CASTRO, *Fundamentals of natural computing: An overview*, "Physics of Life Reviews" 4:1 (2007), pp. 1-36.
- K. DE SMEDT et al. (eds.), *Computing in Humanities Education: A European Perspective*, University of Bergen 1999. <http://www.hd.uib.no/AcoHum/book/> [11/04/2018].
- M. DEEGAN, K. SUTHERLANDT, *Transferred Illusions: Digital technology and the forms of print*, Farnham-Burlington 2009.
- H. EVERETT, III, *The Theory of the Universal Wave Function*, in B.S. de Witt, N. Graham (eds.), *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton 1973, pp. 3-140.
- H. VON FOERSTER, *Understanding Understanding: Essays on Cybernetics and Cognition*, New York NY 2003.
- J.-C. GARDIN, *Les applications de la mecanographie dans la documentation archeologique*, "Bulletin des Bibliothèques de France" 5:1-3 (1960), pp. 5-16.
- J.-C. GARDIN, *Les analyses de discours*, Neuchatel 1974.
- J.-C. GARDIN, *Archaeological Constructs: An aspect of theoretical archaeology*, Cambridge 1980.
- J.-C. GARDIN, *Le calcul et la raison: Essais sur la formalisation du discours savant*, Paris 1991. (a).
- J.-C. GARDIN, *Le role du sujet dans les sciences de l'homme: Essais d'évaluation objective*, "Revue européenne des sciences sociales" 29:89 (1991), pp. 91-102. (b)
- J.-C. GARDIN, *Points de vue logicistes sur les methodologies en sciences sociales*, "Sociologie et sociétés" 25:2 (1993), pp. 11-22.
- J.-C. GARDIN, *Archeologie, formalisation et sciences sociales*, "Sociologie et sociétés, 31:1 (1999), pp. 119-127.
- J.-C. GARDIN, M.N. BORGHETTI, *L'architettura dei testi storiografici: Un'ipotesi*, a cura di I. Mattozzi, Bologna 1995.
- J.-C. GARDIN, M.-S. LAGRANGE, J.-M. MARTIN, J. MOHO, J. NATALI-SMIT, *La logique du plausible: Essais d'épistémologie pratique en sciences humaines*, 2e éd. Revue et augmentée, Paris 1987.
- [JCG] Fondo Équipe Archéologie de l'Asie Centrale et Jean-Claude Gardin, Archivi della Maison Archéologie & Ethnologie René-Ginouvès, Nanterre.

- J.-B. GRIZE, *Logique mathématique, logique naturelle et modèles*, in *Formalisierung in den Geisteswissenschaften / Sciences humaines et formalisation*, “Jahresbericht der Schweizerischen Geisteswissenschaftlichen Gesellschaft” (1974), pp. 201-207.
- J. HAUGELAND, *Artificial Intelligence: The very idea*, Cambridge MA 1985.
- D.P. HENRY, *That Most Subtle Question (Quaestio Subtilissima): The metaphysical bearing of medieval and contemporary linguistic disciplines*, Manchester 1984.
- D. HESTENES, *New Foundations for Classical Mechanics* (Second Edition), New York 2002.
- L. HJELMSLEV, *Prolegomena to a Theory of Language*, Madison WI 1961.
- L. HJELMSLEV, *I fondamenti della teoria del linguaggio* [1961], introduction and translation by G.C. Lepschy, Torino 1968.
- L. HJELMSLEV, *Language: An introduction*, Madison WI 1970a.
- L. HJELMSLEV, *Il linguaggio* [1970], a cura di G.C. Lepschy, Transl. A. Debenedetti Woolf, Torino 1970b.
- J.H. HOLLAND, Outline for a Logical Theory of Adaptive Systems, «Journal of the ACM» 9:3 (1962), pp. 297-314.
- S. LEON, *Digital Public History*, s.d. [http://www.6floors.org/dossier I personal-statement/digital-public-history/](http://www.6floors.org/dossier-I-personal-statement/digital-public-history/) [11/04/2018].
- J.J. MCGANN, *Memory Now*, posted on “4Humanities,” 19 August 2012. <https://4humanities.org/2012/08/jerome-j-mcgann-memory-now-2/> [27/04/2018].
- M. MERLEAU-PONTY, *La structure du comportement* (1942), 6^e ed., Paris 1967.
- M. MERLEAU-PONTY, *Le visible et l'invisible*, Paris 1964.
- Merleau-Ponty, Maurice. *The Visible and the Invisible: Followed by Working Notes* (1964). Northwestern University Press, 1968.
- P. MOSCATI, *Jean-Claude Gardin (Parigi 1925-2013): Dalla meccanografia all'informatica archeologica*, “Archeologia e Calcolatori” 24 (2013), pp. 7-24. [http://www.archcalc.cnr.it/ indice/PDF24/01_Moscati.pdf](http://www.archcalc.cnr.it/indice/PDF24/01_Moscati.pdf) [27/04/2018].
- E. NAGEL, *Review of The Philosophy of Science by S. Toulmin*, “Mind” 63:251 (1954), pp. 403-412.
- S. NOJA, *Saggio di un confronto a mezzo di un elaboratore elettronico tra lo “Šulhan ‘arûk” di Karo e quello del rabbino di Lubavitch*, “Atti della Accademia delle Scienze di Torino” 102 (1967-68), pp. 555-582.
- T. ORLANDI, Interview, in J. NYHAN, A. FLINN, *Computation and the Humanities: Towards an oral history of Digital Humanities*, Cham 2016.
- A.F. PARKER-RHODES, *Inferential Semantics*, Hassocks 1978.
- J.-C. PASSERON, *Le raisonnement sociologique: l'espace non-popperien du raisonnement naturel*, Paris 1991.
- D.R. RAYMOND, F.W. TOMPA, D. WOOD, *Markup reconsidered, paper presented at the First International Workshop on Principles of Document Processing*, Washington DC, 22-23 October 1992, pp.

- 1-25. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.80.9369&rep=rep1&type=pdf> [09/04/2018].
- D.R. RAYMOND, F.W. TOMPA, D. WOOD, *From Data Representation to Data Model: Meta-semantic issues in the evolution of SGML*, “Computer Standards and Interfaces” 18 (1996), pp. 25-36.
- A. RENEAR, *The Descriptive/Procedural Distinction is Flawed*, “Markup Languages” 2:4 (2000), pp. 411-420.
- G. RYLE, *The Concept of Mind*, London 1949.
- S. SCHREIBMAN, R.G. SIEMENS, J. UNSWORTH (eds.), *A Companion to Digital Humanities*, Malden MA 2004.
- G. SPENCER-BROWN, *Laws of Form*, London 1969.
- Tarski, Alfred. “Grundlegung der wissenschaftlichen Semantik.” In *Actes du Congrès international de philosophie scientifique*, Sorbonne, Paris 1935. vol. III: *Language et pseudo-problèmes*. Hermann, Paris 1936, pp. 1-8. <https://gallica.bnf.fr/ark:/12148/bpt6k383668/f5.image.r=.langFR>
- Tarski, Alfred. “The Concept of Truth in Formalized Languages.” In *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. Translated by J.H. Woodger. Clarendon Press, 1956, pp. 152-278.
- TEXT ENCODING INITIATIVE *TEI: Text Encoding Initiative*, 2016. <http://www.tei-c.org/> [09/04/2018].
- TEXT ENCODING INITIATIVE, *Guidelines for Electronic Text Encoding and Interchange*, 2015. <http://www.tei-c.org/Guidelines/> [09/04/2018].
- S. E. TOULMIN, *The Philosophy of Science: An introduction*, London 1953.
- Toulmin, Stephen. *The Uses of Argument* (1st ed. 1958). Updated ed., Cambridge University Press, 2003.
- L. TRUSS, *Eats, Shoots and Leaves: The zero tolerance approach to punctuation*, London 2003.
- J. UNSWORTH, *Forms of Attention: Digital Humanities beyond representation*, paper delivered at “The Face of Text: Computer-Assisted Text Analysis in the Humanities,” The third conference of the Canadian Symposium on Text Analysis (CaSTA), McMaster University, 19-21 November 2004. <http://www.people.virginia.edu/~jmu2m/FOA/> [10/04/2018].
- F. VARELA, E. THOMPSON, E. ROSCH, *The Embodied Mind*, Cambridge MA 1991.
- V. von WEIZSÄCKER, *Reflexgesetze*, in A. Bethe et al. (hrsg.), *Handbuch der normalen und pathologischen Physiologie*, Bd. 10, Berlin 1927.
- N. WIRTH, *Algorithms + Data Structures = Programs*, Englewood Cliffs N.J. 1976.
- L. WITTGENSTEIN, *Culture and Value: A selection from the posthumous remains*, edited by G. H. von Wright in collaboration with H. Nyman, revised edition of the text by A. Pichler, translated by P. Winch, Oxford 1998.