# Did Darwin really talk about Evolution?[1]

**Jean Guy Meunier, Maxime Ste Marie, Jean-Francois Chartier**

**Laboratoire d'analyse cognitive de l'information, (LANCI)**

**Université du Québec a Montréal**

***Published in Agazzi E ,Minazzi, F. Eds . ( 2011) Evolution and Religion. Mimesis 2011 p.127 -146***

Before his Beagle trip (1831-1836), Darwin believed with most of his contemporaries that each species had been created independently with features or characteristics that are naturally suited to the conditions for his living. This allowed harmony in nature and adaptation. However, following his voyage and having debated intellectually with many contemporary theories that aimed at describing not only their characteristics but also the changes observed in each species, he began a slow march towards his own theory, which was exposed in his seminal work of 1859, the Origin of Species. This slow pilgrimage toward his own theory of specific modification in a species went through many intellectual stages. Each new step or transition being rigorously justified by Darwin. And he went through inner arguments to justify this transition. Many scholarly works on this transformation have been writing ( Gould 1977, Bowler, 1987, Osprovat,1995, Limoges 19xx,Richard 1992 Bowler).

And in this intellectual transformation for describing and explaining species' change, the concept of **evolution** was used. But, as noted by Thomas Henry Huxley in his 1878 article on *evolution* published in the *Encyclopaedia Britannica,* this usage of the word "evolution" corresponded at that time to a specific meaning. Indeed, "evolution" already had two distinct biological uses. Initially, it was used to refer to the various embryological theories but it had regain importance with the preformatists.And it was only later, that it was used to characterize the general belief that species have descended from one another over time" (Richards, 1992). And H.Spencer played an important role in the development of this more general meaning of *evolution*.

However, the transition towards a more contemporary usage of *evolution* occurred only many years after the publication of the first edition of the *Origins of Species*, the dual use of "evolution" was still in effect at the time Darwin first edition.

Such dual biological meaning wasn't specific to the word "evolution", however. Von Baer and other Germans would alternatively refer to the Entwickelung of embryo and species, just as English speakers would use "development", or sometimes

[1] Conference given at the Academy International of Philosophy of Science. Florence 2009 On Darwin Evolution and Religion. To be published.

“transformation” and “transmutation”. The more important linguistic phenomenon to recognize, though, is that these several terms were used for both embryological and species change. Thus, while Lyell, Grant, Green, Owen, Chambers, and Carpenter spoke of “evolution”, as well as of “development” and “transformation”, what emerges as significant is that they used the same term to describe the two processes (Richards, 1993: 168).

Still, even if the term “evolution” was used at that time with the embryological meaning, in the popular view, Darwin is seen as the founder if not the father of modern theory of evolution. If this is so, Darwin must have talked constantly about *evolution* in his writings. This term and the concept to which it refers must be present everywhere in his theory.

Unfortunately, and-it may be a surprise for many- Darwin did not often use the term evolution itself in his writing. In fact, in the first four editions, that is, from 1859 to 1866, there is only one occurrence related to the term “evolution”: it is the last word of the conclusion of the work, that is : “evolved”.

In the fifth edition, the same term evolved appears a twice the first occurrence appearing in the fourteenth chapter (p. 573) and the second at the same last spot as in the earlier editions (p. 579). The word and its derivatives, though still scarce, are somehow just slightly more frequent and diverse in the sixth and last edition : evolve (Chapter VII: p. 191), evolved (Chapter VII: pp.191, 202 (2); Chapter VIII: pp. 425,429), evolution (Chapter VIII: pp. 201(2), 202; Chapter VIII, p. 215; Chapter X, p. 282; Chapter XV, pp. 424, 424, 424). In fact it appears only 14 times!

While the most influential and renowned biologists of the time all seemed to endorse a monolexical reference to embryological and specific development, little is known about Darwin's own stance in this matter. These statistics then raise the following question: Did his use of the word *evolution* or other synonyms refer to both embryological and species’ development? Or did he use it only to refer to species’ development? Hence our research title : did Darwin really talk about evolution? If not, what did he speak of?

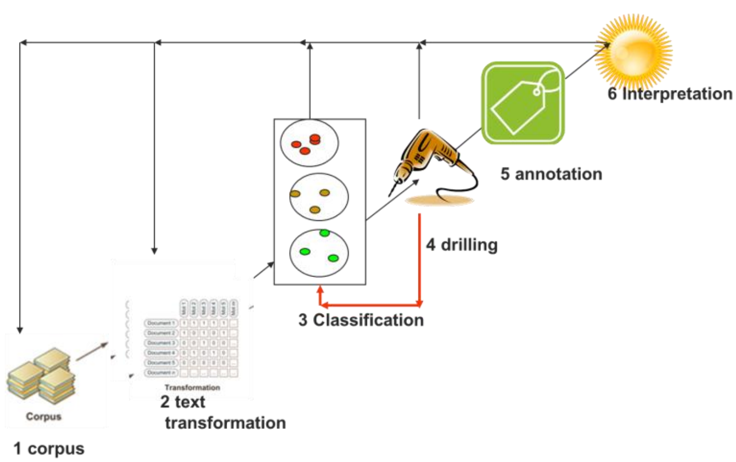
This lexical scarcity of the word "evolution" and its derivate in the OoS somehow complicates the analysis of Darwin’s theory of evolution as it exposed in this book and developed in his later writings. This problem can however be bypassed by focusing not on the use of the term “Evolution, but on the underlying concept. In other words, even if Darwin scarcely used the word *evolution* in the *Origin of Species*, this does mean that the concept of *evolution* itself is not present in the book. But to affirm this is to recognize a crucial distinction between words and concepts, distinction which is in itself it a deep theoretical problem, at the heart of many philosophical, logical and linguistic debates. And the position one takes on this distinction depends naturally on the theory of concept one defends. In this research, we take the classical conceptualist position: a concept cannot be reduced to its lexical expression in any language. For us, there exist many lexicalisations of a same concept. This means then that an inquiry into a concept is an analysis which is not one on the meaning of a term- for this a semantic endeavour - but an inquiry into the predicative and inferential structure that underlies the lexical expression itself. Such a conceptual analysis is at the heart of all conceptual analyses at work in many disciplinary practices, be it history, philosophy, literature, etc.

**The Conceptual analysis in the Origin of Species**.

In view of this, we aim to show that, in Darwin's writings, more specifically the sixth edition of the *Origins of Species*, underlying the scarcely used word of Evolution is a complex conceptual structure by which Darwin understands the dynamics of the change of life forms. To do this we, shall present here a strategy to discover how an author -here Darwin- unwraps this conceptual structure. More so we will show how this can be done by means of a computer methodology on which we have been working on for many years: CACAT.. a computer-assisted conceptual analysis of text.

## Computer assisted conceptual analysis

This computer methodology aims at offering professional expert text analysts "algorithmic" assistance for the exploration of the conceptual structure expressed in the terms and sentences in a text. This methodology unwraps into an algorithmic processing chain which is composed of 6 phases that we can represent in the following diagram. We present them here in schematic manner.




.

We shall give here but slight details of each phase of this methodology and insist more on its application to the sixth edition of Darwin's *Origin of Species*.

## Phase I: the corpus: Preprocessing Units and domains

In this first phase, text corpora are chosen and prepared for analysis. This task is not as menial as it may at first appear. In our research on OoS, two text samples are chosen. The first one is the whole text itself, and the second is a limited subtext constituted of the concordance of the term EVOLUTION (or its derivate) in the same text. This dual strategy will allow for comparisons between results.

Here are a few examples of the subtext produced by the concordance (set of sentences of paragraphs) of the word Evolution (or its derivates):

> *And even if one was so, what chance was there of the perpetuation of such a variation?" But the case is not here put fairly. It is admitted by most* ***evolutionists*** *that mammals are descended from a marsupial form; and if so, the mammary glands will have been at first developed within the marsupial sack.*
>
> *4At the present day almost all naturalists admit evolution under some form. Mr. Mivart believes that species change through "an internal force or tendency," about which it is not pretended that anything is known.*

Each segment here is but a sample of the whole segment. Be it on the whole corpus or on the concordance, basic lexical units and textual units, that is : words and segments or paragraphs, are chosen. On the whole text, this gives 974 paragraphs and words. On the concordance, this gives 13 segments and x words.

### Phase II: Text transformation

In this phase, the two corpora chosen are transformed into vectors and then combined into a matrix. And on this matrix mathematical classification strategies are applied. So for each paragraph of the text, or each segment of the concordance, a vector is created. The properties of each vector are the values (presence or absence or other indexes) given to each word in them. For instance the segments 2 and 4 b become vectors 2 and 4.

| | Admitted | naturalist | tendency | marsupial | force | developped | Form | sacks |
|---|---|---|---|---|---|---|---|---|
| Vector s2 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| Vector s4 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 |

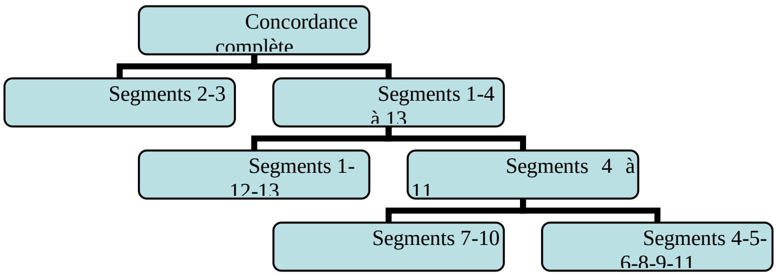
By repeating this procedure for all segments, a matrix of the whole corpus can be formed.

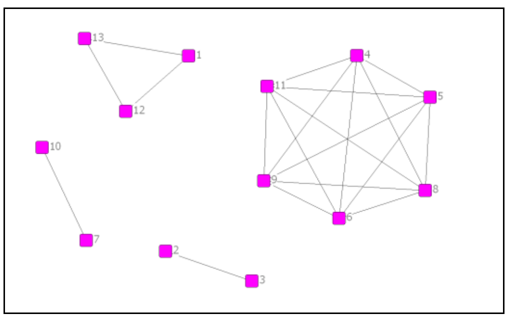
For the whole text the matrix is one of 974 by 9442.For the concordance the matrix is one13 x 3256.

### Phase III:Text Classification

Follows the third phase:. On this matrix three classification algorithms were used : the EM, K-means, and the Adaptive Resonance (ART) classification algorithms (Carpenter & Grossberg 2003). Here, for simplicity reasons, we illustrate the classification results obtained on the concordance. Similarly results but with a wider range of classes has been found on the whole text. The results on the concordance were the following :

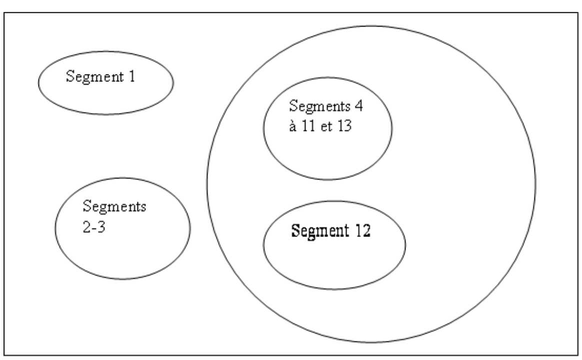
**The EM algorithm**

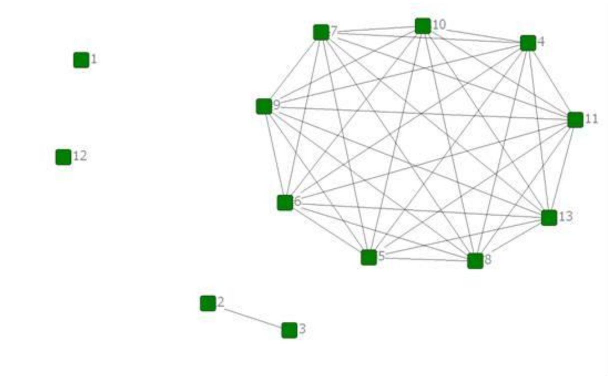



The above tree can be represented by a CLIC graph which shows that segment 2 and 3 are related and form a particular class but which is separated from all the others.

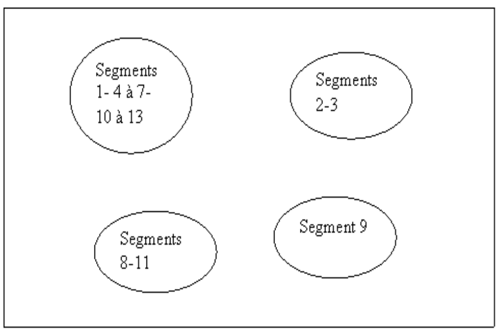
**The K means**

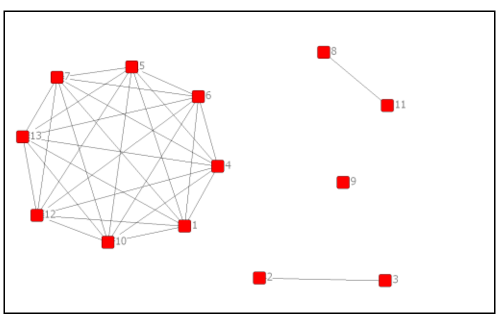




Here, a Euler representation shows that segment 2 and 3 are a class apart from the others but has distinguished segments 1 and 12, also reflected in the clic representation

**The Adaptative Resonance Theory (ART1)**

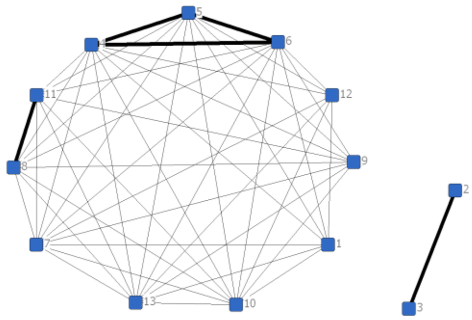


Here too, the Euler representation shows that segments 2 and 3 form a distinctive class, but this algorithm has also distinguished segment 9 as well as segments 10 and 13, distinctions also reflected in the clic representation.

**Consensus**

In light of these classification results and their similarities, the following “consensus synthesis” can be obtained:

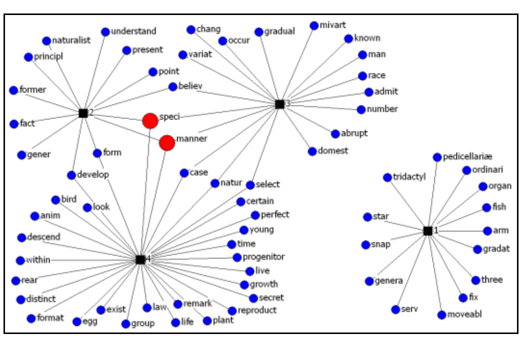



This representation shows that segments 2 and 3 form a distinctive class, apart from the others, that the two groups made of segments 3 and 11 and segments 5, 4, 6 which are internal classes to the greater class that included the rest of the other segments1-7-9-10-12-13.

How can we understand these classifications? From a mathematical point of view, it means that, in each class, some vectors are more similar to each other; in linguistic terms, it means that there is a high probability that these segments have some similar semantic content, that is thes" talk about the same thing because containing more similar words" But to know more about this common semantic content, a deeper analysis or "digging" is necessary. This is the object of the forth and fifth phase.

## Phase IV Conceptual Drilling

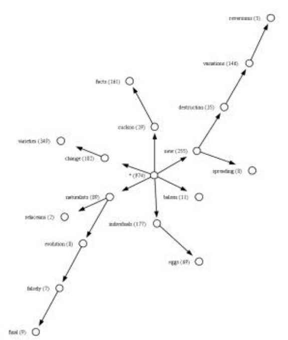
In this phase, a filtering process that "drills" into each class of sentences is applied. The aim of this drilling is to discover particular, important terms that are metaphorically said to be semantical "nuggets", that is `: one looks for terms that, because of certain mathematical property (high TF,IDF, high frequency, etc.), open up the sub-semantic fields of the main class. As an illustration. here is the result of the drilling into the segments of the segment 1, 2, 3, 4, in which we discover semantical nuggets MANNER and SPECIES.




From a semantic point of view, this means that the segments of class (2,3) share with segment 4 a common semantic subfield, one that pertains to both SPECIES and MANNER.

If this digging is done recursively on the whole text and for each nugget term, this produces the following "semantic" graph which can be conceived as a sort of semantic gallery to which EVOLUTION gives access.

From the center node, one is directed toward a first layer of nodes, and then a second one, and so forth. As shown in the graph, some semantic paths quickly come to a dead end, as in the case of the path leading to INDIVIDUAL and EGG. Other ones allow for deeper digging and new semantic paths.




Other experiments varying the mathematical constraints on the discovery of these semantical nuggets produces comparable results (Ste Marie et al.2010)

-**Phase V Annotation**

All these classification and digging strategies build up a complex network of terms that are semantically significant. For an expert reader, these associated terms are indices of the conceptual architecture build in the text. But from these conceptual indices, a whole portrait of the architecture has to be built in order to discover the fine details of the underlying conceptual structure. This is done in part through the annotation phase. Here, interpretative informations(comments, paraphrase, citations etc.) are added to each segment retained. These are then aggregated into a global comment that summarizes the semantic content of each segment. This annotation process can be done automatically[2] or manually. In this project, the annotations produced were all manual ones. Here are some examples of these annotations:

1- The class of segments 2&3 is the main and constant class. And it contains a general thesis on evolution **: Evolution is different in individual and species : it relies on minute variations of forms or structures**.

[2] Some experimentations were made in this matter with the help of platform EXCOM's (see Desclés et all 200)

2-The class of segments 10 & 7 affirms that **Evolution happens in an extremely gradual manner**

3-The class of segments4 5, 11,9, 6, 8 **: Evolution is not abrupt, it happens through natural selection, by accident and requires intercrossing. Some characters in species are less probable than others.**

These types of commentaries are repeated for all segments**.**

## PhaseVI: Interpretation

Given these different classification, term drilling, and annotation strategies, a question naturally rises: What are the results of this procedure for understanding Darwin's conceptual structure of "evolution"? For the moment and as revealed by the annotative summaries done, the conceptual structure of evolution in Darwin's *Origin of Species* unwraps in two main directions:

The first direction can be stated as a thesis: **Evolution is different in individuals and species : it relies on minute form or structure variations.**

To talk in mathematical terms of lattice, this thesis is the supremum of the conceptual architecture. It is it's "clef de voute". And it dominates both the whole text and in the concordance:

In the *whole text* for instance, the individual vector opens directly on the original moment of evolution of an embryo : the egg (69 occurences) but, as shown in the graphs above it stops there. That is : not much more it developed in the lign of the indivual embryo. This means that, even if, Darwin consecrated a chapter to embryology he did not, in the OoS associate the concept of *evolution* to embryological question. The lexical field involved does not call upon or related to the lexical fields normally used, for instance, by the preformatists or the recapitulative theories for explaining the ambryo' evolution.

But we cannot infer here that, for Darwin, the model of the individual was not important in a theory of evolution. All we may say here, it that it seemed to be used only as a point of departure.

In the *concordance*, the same direction is to be found: Apart the technical details about the arms of star-fisches or the urchins, the segment 2 and 3 point mostly to lexical field of the MANNER in which the SPECIES develops. These segments do not open up into the lexicon of the development of the *individual*. But more importantly, the analysis reveals that the lexicon mainly pertains to the development of *the Species and genera";* And this development is expressed in terms of a "perfect series of gradations".

In summary, as we can see, Darwin' conceptual architecture of *evolution* does not unwrap into the semantical field related to discourses on the *embryo*. At first sight, even if this conclusion may seem simple it is of the outmost importance : For such a conceptual architecture was an intellectual tendency that was only starting to emerged in Darwin's time. Indeed, in Darwin's scientific milieu, many thought that the various concepts of the models of individual development, mainly the ones used for the explaining the development of the *embryo*, could serve as a model for the development of the *species*.

According to our research, Darwin's own stance on the question does not go in this direction: the lexical field related to it is simply absent, as if pushed aside. One reason we could advance to explain this divergence is that, if Darwin accepted to dwell in the conceptual architecture proper for explaining the development of the individual, he would surely become entangled into the various explanatory concepts proper to the explanation of the individual development. For instance, this would imply to address questions and problems such as those of finding a primary cause, a final cause, a progressive development, etc. And implicitly, this would ultimately mean that if there is a genitor for the embryo, there should be a " creator" for the species. All these issues would have been very problematic for Darwin, for they are not part of his inferential and propositional architecture.

For a more complete understanding of the underlying conceptual structure of this first conceptual thesis, deeper digging into the question of the evolution of the *species* is needed. This is part of our ongoing research. But as we may already see, our present results might give us some indications for further research:

a) in the concordance, we find that the model of species' evolution will call upon "natural selection" and this evolution has to be "gradual, not abrupt". And anybody familiar enough with this question will notice that the concept of "progressiveness" is absent. This concept was a building block of the conceptual structure of the model of the evolution of the embryo, and it was implicit in it's use that the embryo was "progressively" oriented towards a final well formed adult. In other words, when applied to the embryo, one must infer that it was guided by some final "cause" or design. Embryos do not evolve in a haphazard way. They " progress " toward " the adult forms. This dimension is absent from Darwin's model of the evolution of species.

In the whole text, we also find many conceptual "galleries" that open up into epistemological questions pertaining to evolution of the species as such, for instance the " naturalist" approach, its falsity, etc. But more interestingly, we find the questions of "variations","destruction", "reversions" "varieties". These lexical galleries lead us into the second dimension of Darwin's model of evolution: the dynamics of change in life forms.

The second direction pertains to the methodological question regarding the modelling of the changes of life forms. It becomes the following thesis : **Evolution is a theory of change in species lifes forms.**

This dimension is directly related to the preceding thesis that is: evolution is different in the individual and in the species. Indeed, if Darwin cannot call upon the model of evolution of the individual, more specifically of the embryo, for the explanation of changes in species, how can these be explained and accounted for? What could be then the model for specific development?

If we look at the results of the classification and drilling both of the concordance and of the whole work, we find conceptual terms that point at properties of these changes, for instance *development*, *perpetuation*, *variation minute variations*, *transition*, *coordination gradation*, *series*, *perpetuation*, *modification*, *adaptation*, *retrogression states of nature*, *form*, *groups of forms*, *transitions*, and so forth.

All these terms are not typical of a theory of the development of an individual. They are in fact the typical vocabulary for describing informally a specific type of system that has forms, states, groups of forms and changes through time. In other words, and this is our main point: Darwin's theory of specific development is entangled into the informal modelling of a dynamical system.

We could summarize our position by saying, that for Darwin, a correct theory of evolution is one that must:

a) firstly, *describe* individual forms or group of individual forms as "states of nature" embedded in changes, variation, gradation, etc., dynamics which they go through in a certain lapse of time;

b) secondly, it must *explain* the causes of these changes;

c) thirdly these changes must be modeled in some formal or informal manner. This is precisely the aim of a dynamical system model. What is specific here is that this dynamical system is applied to living forms, not purely material entities.

In this perspective, the concept of *evolution* corresponds to the classical essential characters of a dynamical system. An evolutionary process will contain a set of forms and a time span. And it contains the changes of the forms undergone in a time span, be they discrete (as found Lyndenmayers systems) or continuous, as found in linear or non linear dynamical systems.

This understanding of the concept of evolution was already implicitly present in its first usage by the Dutch entomologist Jan Swammerdam (1637-80), slowly "evolved" through time and theories into Darwin's conceptual architecture, and later became explicit in the works of Spencer. And it is this sense it still has today.

In this perspective, a theory of evolution in Darwin's conceptual structure focuses on the set of transformations that organic life forms go through, be it at the level of the individual or the species, that is at the ontogenic level or the phylogenic level. In such a paradigm the individuals forms are but st*ates* in a dynamical process. And their change are their transformations.

In such a model, the" principle" or "laws" of such transformations are understood by Darwin as the *causes* of evolution. And this means then that these causes are internal to the system. Such a position has a immediate consequence, these causes are not intentional causes grounded in some agent be it God or humans (as it the case in artificial breeding).

The question that now has to be answered is how Darwin understood these principles or laws of transformation. We believe that he did not possess formal concepts clear enough to model fomally these dynamic changes. And these principles or laws were not easy to identify, even though models of such principles or laws of changes in life forms have already been developed in Darwin's time.

For instance, some saw the development as changes in a sequence of forms, as the recurrence of some pre-existing forms into a following one. This was the model defended by preformationist or recapitualionist models : transformations are a kind of morphism between states.

Others saw the changes between forms as a kind of heteromorphism, where the changes focus not on the continuity of the constituting forms but on the modification of their organization. Development is hence just a variation of the organization of forms. This is the essence of a transmutationist approach to evolution.

Some would declare that there was an implicit design driving the development. Some (Lamark) maintained that the transformations were oriented towards better interaction of the form with its environment: this is the source of adaptation. Some thought (Spencer) that the organisation is self-contained, gradual, progressive, and complexifying.

It was Spencer's hope to offer ultimately a set of general concepts and propositions to explain, at least in a general manner, the dynamics of the evolutions of all forms of life. For him, evolution refers to the dynamic process of complex but gradual development of life forms. And it is this theoretical perspective that in fact underlies the modern general meaning of evolution.

Whatever these thesis are, each one generates its own set of concepts and propositions for exploring, expanding, and justifying these views. The set may not be well defined or tidy but still forms a conceptual architecture of its own.

Darwin had his own answers to this question. But he was confronted to a great theoretical challenge. Even if his underlying conceptual architecture implicated a dynamical system, he would have to deal with the specificity of the dynamics of organism life in species. And at that time, the only variety of other known dynamic models were inspired by thermodynamic. And this could not be applied to the dynamical change of species. The evolution of a species cannot be modelled like the evolution of a gas! One may identify laws of the evolution of a gas in a closed container because all the molecules that are found in the initial states are conserved in their change process and are still present at succeeding and final states. More so, the control parameters are given by an exterior agent (heat, for instance).

But for Darwin these types of transformation could not be the ones present in the evolution of species, for:

a) the population at the initial state is never the same at the final state of the change process. But more importantly, he refused that there could be an external agent causing these processes.

b) And at that time, the mendelian, genetic, dissipative, the Lyndenmaer, fractal, morphogenetic, generative, chaotic, dynamic, and emergentist formal models did not exist. The only means he could call upon was "natural selection" : a concept that is deeply metaphoric and as Ospovat (19xx) showed was lately integrated in Darwin's conception. And it has it own conceptual problems. At a general meaning, it is "nature" that "selects" and this selection produces varieties in the species. But how is this done, is a problem in itself.

## Conclusion

The aim of the research presented here is both technical and epistemological. We aimed at showing that computer technology can allow for and assist the expert reader in the rigorous exploration of difficult conceptual structures present in the writing of an author. We have illustrated this on the concept of *evolution* in Darwin's *Origin of Species*. The approach shows us that this concept seems to have a specific and complex architecture in this book. First, the model of evolution for the species is different from the model of evolution of the individual. And this model conceives evolution as a dynamical system were the control parameters are internal and take the form of natural selection. In so doing, Darwin evade the possible traps entangled in the models of embryological evolution.

On a more technological basis, the present study might also give "a more rigorous idea of what reading and analyzing a "text" with the help of a computer involves" (Meunier, Forest and Biskri, 2005: 959). While such computer-assisted techniques "can diminish the burden for many researchers in the field of social sciences and humanities" (Meunier et al, 2005: 976), they haven't yet made their mark in the humanities and social sciences. The completion of this research and the diffusion of its results precisely aims to help these scientific communities recognize the full importance of computer-assisted text applications and technology.

*References.*

Bowler P.J. (1984). *Evolution: the History of an Idea*. University of California Press.

Bowler P.J. (1975). The Changing Meaning of 'Evolution'. *Journal of the History of Ideas*, 36: 95-114.

Darwin Charles. (2002-2009). *The Complete Works of Charles Darwin Online*. http://darwin-online.org.uk

Forest D. and Meunier J.G. (2000). La classification mathématique des textes: un outil d'assistance à la lecture et à l'analyse de textes philosophiques. In: Rajman M. and Chappelier J.-C., editors. Proc. of JADT 2000 (5$^{es}$ Journées internationales d'Analyse statistique de Données Textuelles, Lausanne). pp. 325–329.

Limoges,C (1970)*La selection naturelle*. Presses Universitaires de France, 1970

Meunier J.G., Forest D. and Biskri, I. (2005), Classification and Categorization in Computer-Assisted Reading and Text Analysis. In Cohen, H. and Lefebvre, C. editors, *Handbook of Categorization in Cognitive Science*. Elsevier.

Ospovat, D.(1995) *The development of Darwin'Theory*, Barns and Nobles,1995

Richards R.J. (1992). *The Meaning of Evolution: the Morphological Construction and Ideological Reconstruction of Darwin's Theory*. University of Chicago Press.

Ste Marie, M., J.G. Meunier, J.C. Chartier,(2010) *Reading Darwin between the Line*. Actes JADT 2010. Rome,

Gould J.S.(1977) *Ever since Darwin*,New York, Norton

Carpenter, G.A. & Grossberg, S. (2003), Adaptive Resonance Theory, In Michael A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks*, Second Edition (pp. 87-90). Cambridge, MA: MIT Press