# Proposals for a Discourse Analysis Practice Integrated into Digital Humanities: Theoretical Issues, Practical Applications, and Methodological Consequences

**Julien Longhi**

Agora Lab (EA7392), Institute of Digital Humanities (FED4284), CY Cergy Paris Université,
F95000 Cergy, France; julien.longhi@u-cergy.fr

**Abstract:** In this article, I put forward a linguistic analysis model for analyzing meaning which is based on a methodology that falls within the wider framework of the digital humanities and is equipped with digital tools that meet the theoretical requirements stated. First, I propose a conception of the digital humanities which favors a close relationship between digital technology and the humanities. This general framework will justify the use of a number of models embodied in a dynamic conception of language. This dynamism will then be reflected in the choice of metrics and textual analysis tools (developed in the field of textometry, especially the *Iramuteq* software). The semantic functioning of linguistic units will be described by using these tools within the identified methodological framework and will help to better understand the processes of variations, whether temporal or generic, within vast discursive corpora. I propose a way of analyzing corpora with specific tools, confronting the humanities with computing/numerical technology.

**Keywords:** discourse analysis; corpus linguistics; digital humanities; semantics; textometry

## 1. Introduction

The aim of this article was to put forward a linguistic analysis model for analyzing meaning which is based on a methodology that falls within the wider framework of the digital humanities and is equipped with digital tools that meet the theoretical requirements stated. This article will thus first propose a conception of the digital humanities (DH) which favors a close relationship between digital technology and the humanities. This general framework will justify the use of a number of models embodied in a dynamic conception of language. This dynamism will then be reflected in the choice of metrics and textual analysis tools. Using these tools within the identified methodological framework will help describe the semantic functioning of linguistic units and better understand the processes of variations, whether temporal or generic, within discursive corpora (for this paper, the corpus will be composed of political discourse extracted from the digital social network Twitter).

The subject DH has been heavily mobilized in recent scholarly output and has become a predominant element in the various projects that structure the academic landscape (Longhi 2017). While the digital humanities "are more than a passing fad, as malicious tongues would have it," and are "a real fundamental movement called on to redefine all of the research fields in the human and social sciences" (Dacos and Mounier 2014, p. 6), they do raise many questions about the definition of their scope in relation to what is meant by both the terms "digital" and "humanities." The description and phrase "digital humanities" has a discursive potential which is due not only to the polysemy of the terms "digital" and "humanities" but is also connected to the status of this description (description? categorization?). There are of course many works questioning this very definition.

The English version of Wikipedia defines the digital humanities as follows:

DH is an area of scholarly activity at the intersection of computing or digital technologies and the disciplines of the humanities. It includes the systematic use of digital resources in the humanities, as well as the analysis of their application. DH can be defined as new ways of doing scholarship that involve collaborative, transdisciplinary, and computationally engaged research, teaching, and publishing. It brings digital tools and methods to the study of the humanities with the recognition that the printed word is no longer the main medium for knowledge production and distribution. (en.wikipedia.org/wiki/Digital_humanities)

The equivalent French Wikipedia page on "*humanités numériques*" or "numerical humanities," which is the French translation for "digital humanities," speaks more specifically of the intersection between "computing and the arts, literature, humanities, and social sciences" ("*arts, lettres, sciences humaines et sciences sociales*") (fr.wikipedia.org/wiki/Humanit%C3%A9s_num%C3%A9riques). The examples from Wikipedia are given because for a first step of the analysis, they give good insight into / overview of the conceptual difference between the "digital humanities" and the "humanités numériques". The clarification in the French definition points to the particularity of francophone and especially French digital/numerical humanities in that they are linked to disciplinary demarcations that stem from academic and institutional definitions. They are further defined as "an area of research, teaching, and engineering" ("*un domaine de recherche, d'enseignement et d'ingénierie*") characterized by a twofold relationship between the digital and the humanities: digital tools for the human and social sciences (HSS) and digital content for HSS studies.

The *Digital Humanities Manifesto* (tcp.hypotheses.org/318) drawn up by "digital humanities (*humanités numériques*) players and observers who came together at the THATCamp [unconference] in Paris on 18 and 19 May 2010" has put forward the following definition:

I. Definition

1.  Society's digital turn changes and calls into question the conditions of knowledge production and distribution.
2.  For us, the digital humanities concern the totality of the social sciences and humanities. The digital humanities are not *tabula rasa*. On the contrary, they rely on all the paradigms, *savoir-faire*, and knowledge specific to these disciplines, while mobilizing the tools and unique perspectives enabled by digital technology.
3.  The digital humanities designate a "transdiscipline," embodying all the methods, systems, and heuristic perspectives linked to the digital within the fields of the humanities and social sciences.

Here, the human sciences (in the French meaning of the phrase, as discussed below) take center stage more explicitly. They are described as a cross between disciplines or a trans-discipline, taking into account the way in which different HSS paradigms are impacted by digital technology. The latter seems to be considered both from the perspective of its computing dimension (tools, development, etc.) and practices. Problems arise as soon as we try to characterize, map, or practice the digital humanities. A particular difficulty lies, I believe, in the relationship to computing/digital technology since the thematization of the digital humanities seems to pass primarily through the prism of the humanities.

The English term "digital" usually translates into French as *numérique* or "numerical" and, much less often, *digital*. However, "digital humanities" are not necessarily identical with "numerical humanities." Thus, bearing in mind that the word "digital" stems from and can refer to the word "finger" (Latin *digitus*), Olivier Le Deuff (2015) notes that digital humanities "remind us of the importance of fingers and thus of possible ways of manipulating and pointing." He states that digital humanities are an integral part of a much longer history

which long preceded the emergence of numerical technology and computing. It is the history of scholarly practices, but also and especially the history of tools for classifying and organizing thought and knowledge. Just as it is worth remembering that the history of

hypertexts started before numerical technology and their well-known concrete manifestations within web browsers, digital humanities are part of a tradition of organized reading that goes outside the linearity of texts and documents.

This is one reason why, for the rest of this article, I will employ the term "numerical" in its French sense and speak of "numerical humanities," "numerical technology," or simply "the numerical." This choice clarifies my position about numerical humanities: the numerical humanities are a field conducive to advancing the history of ideas and the history of science, since they require a meta-disciplinary approach which reveals the challenges of each field (the humanities and computing) and the particularities of specific disciplines (literature, history, linguistics, etc.; data mining, image processing, knowledge management, etc.). I follow Jean-Guy Meunier (2019a) who noted "a tension, a paradox, a contradiction between numerical and humanities" and raised the following questions: How should this relationship be conceptualized? What forms of knowledge are NH? Can they be seen as a type of scientific knowledge? A detour into disciplinary issues (inter/poly/multi/trans-disciplinarity) as proposed by Morin (1994) will help better situate things in this section:

> In fact, it is inter-, poly-, and trans-disciplinarity complexes that have operated and played a fertile role in the history of science; we must remember the key concepts involved, namely, cooperation and, better still, articulation, common object and, better still, common project. Finally, it is not only the idea of inter- and trans-disciplinarity that is important. We must "environmentalize" disciplines, that is, take into account everything that is contextual, including cultural and social conditions, that is, see in which environment they are born, raise issues, ossify, and metamorphose. We also need meta-disciplinarity, where the term "meta" means to go beyond and preserve. We cannot break what has been created by disciplines, we cannot pull down every fence—this problem pertains to any discipline, science, and life; a discipline must be both open and closed. In conclusion, what would be the use of all our fragments of knowledge if not to be compared and contrasted in order to form a configuration that meets our expectations, our needs, and our cognitive questions.

My approach to the numerical humanities is therefore fundamentally multidisciplinary, and requires a very strong understanding, especially between the humanities and the computer sciences, thanks to a meta-disciplinary perspective.

## 2. Materials and Methods

In order to "rise above" this questioning and adopt a reflexive position, we can query the "cyclical dynamics" proposed by Jean-Guy Meunier which presents the reciprocal relations between conceptual, formal mathematical, formal computational, and physical computer models, as we can see with Figure 1:
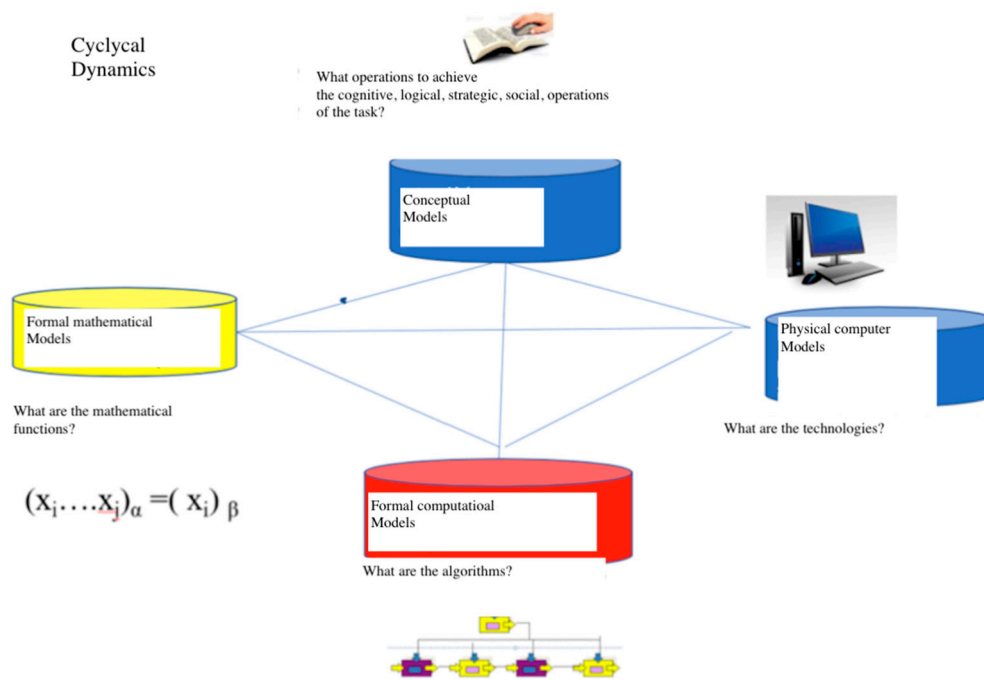
Cyclycal
Dynamics

What operations to achieve
the cognitive, logical, strategic, social, operations
of the task?

Conceptual
Models

Formal mathematical
Models

Physical computer
Models

What are the mathematical
functions?

What are the technologies?

$$(x_i \ldots x_j)_\alpha = (x_i)_\beta$$

Formal computatioal
Models

What are the algorithms?

**Figure 1.** Model dynamics (Meunier 2019a).

The articulation of these models proposed by Meunier (and which will be explained in the different points of this section) makes it possible to reshuffle the cards of the separation between the humanities and computing since while the "numerical" can be embodied in the technologies used and mobilized, the other models transcend the separations established between the humanities and computing and ultimately, it is all about being able to develop coherent models between conceptualization, mathematical formalization, and algorithm design. In the next section, I will therefore show how the theoretical model of discourse analysis that I propose (Longhi 2015, 2018), in connection with a dynamic semantic theorization, can be embodied in formal models that can be subject to a computational and physical transposition.

### 2.1. Models and Tools: Clarifications

This detour through the theoretical semantic model that underlies my discursive approach is needed since the analytical architecture I will mobilize must, both in terms of methodology and tools, be anchored in a certain conception of discourse. Meunier (2019b, p. 23) has spelled out the necessary link between four types of models: conceptual, formal, computational, and computer models. The objective here is to show the coherence of my analysis approach by looking at these four models. While my conceptual model is the theory of semantic forms applied to numerical corpora and implemented using textometric resources, it is important to highlight the way in which concepts are mobilized in the formal and computational models and how they can be dealt with by using a program pertaining to the computer model. I will therefore go in detail over the different models and their application to my research.

### 2.1.1. The Conceptual Model

The conceptual model expresses the objects, operations, and methods that are epistemically relevant to humanities research by means of natural language (concepts, statements, arguments, discourses, etc.). This model is limited: its mode of expression (natural language) is burdened with ambiguities, bias, and impressions. Nevertheless, it remains essential in the research process because it is the only language immediately accessible to us. In the present case, the conceptual model of the *Theory of Semantic Forms* draws on the concepts of patterns, profiles, and themes which, in addition to

being polysemic, are used in several theoretical frameworks. It is therefore necessary to be able to transcribe these concepts in a way that is faithful to their definition and to render them functional for a tool-based analysis.

### 2.1.2. The Formal Model

Meunier explains that "in a 'numerical humanities' project the role of the formal model is to translate certain elements of the conceptual model into some formal language (mathematical, geometrical, logical, grammatical, etc.)." This "translation" is a work in progress in my research and cannot be detailed in this article. But to clarify the direction taken by my research, my reflection is particularly related to the work conducted by Yves-Marie Visetti (2003, 2004) who, especially in relation to the links between the conceptual and formal models within this theoretical framework, states that his approach, "guided by the question of forms in semantics and not directly by that of the continuum[,] consisted of a critical return to the historical Gestalt schools and, at the same time, to phenomenological philosophy by following a line going from Husserl to Merleau-Ponty via Gurwitsch" (2004). In this approach, "the importance of mathematizing the theories and techniques within cognitive disciplines is still largely unknown, especially in AI," and it is a question "of allowing a real schematization of descriptive theoretical concepts by means of both formal and intuitive mathematical structures" (2003). In this spirit, topological, geometrical, and dynamic methods "must be promoted in the construction of models in the same way as symbolic or numerical methods are" (2003). It therefore seems to me that formalizations based on topological, geometrical, and dynamic models can lead to relevant avenues for "translating" concepts into formal language. The computational and the computer models I will now present can also provide a way to access the formal issues of my proposal because they help make them understandable: indeed, the fact of thinking about how to implement this theory with metrics and tools requires to make "concrete" these propositions, and to confront the textual data.

### 2.1.3. The Computational Model

"For its part, the computational model is directly linked to formal models. Its role is to translate the computable statements of a formal model into the statements of a computational language, that is, into algorithms or programs." (Meunier 2019b). Within my numerical humanities research, the transcription of the dynamic and topological models referred to earlier has been echoed in, on the one hand, similarity analyses and, on the other hand, a certain use of descending hierarchical classifications (and their graphic representations). This echo, as indicated in the previous section, is in the process of being precisely characterized through more in-depth thinking on the formal models, the mathematical conceptions underlying them, and the challenges of the transition between the conceptual and formal models. I will, however, indicate the characteristics of the content of the computational model envisaged (analysis examples will be given in Section 3):

-   With regard to similarity analyses, Loubère (2016) explains that this model stems from "graph theory (Flament 1962, 1981; Vergès and Boumédine 2001) and presents the structure of a corpus by schematizing these relationships, thus making it possible to highlight the links between forms in text segments." More precisely, Marchand and Ratinaud (2012) explain that "after segmenting, recognizing, and lemmatizing forms, followed by ECU partitioning, the matrix of the overall corpus can be represented in various ways (linear or circular trees; form size proportional to frequency or statistical link, etc.). The tree of lexical links in the corpus is represented here (co-occurrence calculation and Fruchterman-Reingold algorithm)." The co-occurrence calculation and Fruchterman-Reingold algorithm[1] are conceived here as a means to take into account the

---

[1]    github.com/gephi/gephi/wiki/Fruchterman-Reingold: The Fruchterman-Reingold Algorithm is a force-directed layout algorithm. The idea of a force-directed layout algorithm is to consider a force between any two nodes. In this algorithm, the nodes are represented by steel rings and the edges are springs between them. The attractive force is analogous to the spring

"profiles" since they reflect both the syntactic proximity and frequency of associations, and the force of the relationship between units.

- With regard to descending hierarchical classifications, I have followed Loubère (2016) by choosing a Reinert-type classification proposed by the Iramuteq software: "this classification, implemented for the first time in the Alceste® software (Reinert 1983), makes it possible to highlight lexical worlds. These discourse structures assume that a statement is a stance that is dependent on the subject but also on its activity and context." At a methodological level, as described by Loubère, the vocabulary of the corpus is "used to build a double-entry table listing the presence/absence of the full forms selected in/from the segments; a series of bi-partitions are [then] performed on this table based on a factorial analysis of correspondences." These classifications are very useful for understanding the themes of a corpus through the lexical worlds that compose them.

- In terms of the visualization and representation of the results, one can use the factorial analysis of correspondences (FAC): this is a statistical method "that can be applied to contingency tables such as tables resulting from counting different types of vocabulary (table rows) in different parts (table columns) of a corpus of texts" (Salem n.d., Lexico3 software tutorial). We start by "calculating a distance (known as the $\chi^2$ distance) between each pair of texts making up the corpus. These distances are then broken down into a hierarchical succession of factorial axes. ... This method helps obtain synthetic representations of both the distances calculated between texts and those that can be calculated between the textual units that make them up." It is nevertheless important to note that while "the main advantage of FAC lies in its ability to extract from vast data tables that are difficult to grasp simple structures that can approximately reflect large underlying oppositions within a corpus of texts," this is only an "approximation" and the results of previous functions (calculations, tables of figures) must be precisely considered. These visualizations are an issue in the context of research in numerical humanities, which aim, in particular, to make complex results comprehensible through visualizations, which are based on metrics and rigorous calculations.

### 2.1.4. The Computer Model

"Finally, the computer model translates the algorithms created in the computational model into electronic-type mechanical forms. It provides a hardware architecture that allows the calculations or computation in the computational models to be actually performed." (Meunier 2019b)

Here, we therefore have the technological realization of a theoretical model as we have formally envisaged it. This corresponds to the definition of the digital humanities given in the first section of the article: by approaching the object of study from a multidisciplinary perspective and its realization from a meta-disciplinary one, we can use functions, metrics, and tools consistent with the theoretical anchoring. In the present case, within my own practice, the Iramuteq software allows to bring together these different algorithms and functionalities. The Lexico software (3 then 5) is also frequently used, particularly for the functions of repeated segments and FCA which can help reflect the temporal dynamics of a corpus.

This choice is motivated by my dual conceptual interest in the links between forms and profiles on the one hand, and forms and themes on the other. It should be noted that these algorithms do not directly address the dynamic/topological models mentioned in relation to the formal model. The proposed articulation lies in the fact that the work I am doing on discourse analysis is based on variation and comparison/differentiality: thus, it is the variational application of Iramuteq (or Lexico) algorithms that can help reflect corpus dynamics (comparing states of discourse with "snapshots" of different stages, taken as a whole that varies at different discursive moments). The illustrations in the third part of this article will make these considerations clearer. As an example, textual time

---

force and the repulsive force is analogous to the electrical force. The basic idea is to minimize the energy of the system by moving the nodes and changing the forces between them. For more details refer to the Force-Directed algorithm.

series can be used to chronologically grasp corpus dynamics. As Salem (n.d.) explains in a Lexico 3 tutorial (although these functions can be generalized to other textometry tools), textual time series "are corpora built by bringing together similar texts produced by the same textual source over a period of time"; taking into account the chronological dimension of such corpora makes it possible to "highlight variations that occur over time in the use of the vocabulary and important moments in its evolution." My work, therefore, falls squarely within the analysis of the dynamics of meaning specific to a corpus, in this case in relation to its temporal dimension.

According to the theoretical model presented, the computer model that I believe provides the best "hardware architecture" is a computer (which allows the software to be installed) or a platform which would integrate the algorithms described into its operation (this was the case with the #Idéo2017 platform: http://ideo2017.ensea.fr/plateforme/). It remains to be clarified how the described algorithms can reflect the concepts of the theory of semantic forms.

### 2.2. From Theory to Tools: Profiling and Similarity Analysis, Themes and Lexical Classification

For Cadiot and Visetti (2001), profiling dynamics "refer in part to new usages already recorded in a lexicon and, in a much more generic form, in a grammar. But they also occur within new themes taking them up in their own grid which is possibly extrinsic to either given patterns in a language or already established lexical profiling standards" (130). Thus, the Analysis of Similarities, which makes calculations based on "a co-occurrence index (how many times the elements will appear at the same time)" in order to give a visual result "in which the size of the words is proportional to their frequency and the size of the edges is proportional to their force," allows us to reflect these profiling operations, their new usages, their stability, etc. This makes it possible to visually show the frequency of words in relation to specific associations. This functionality represents, in a way, the "profiling" of units, that is, it reflects their stabilization within a corpus through frequent associations that "profile" the uses of forms in a given domain or practice.

As regards the themes, Cadiot and Visetti (2001) explain that a theme "translates the stabilization and actualization within and through a "referential" or even "conceptual" domain." Thus, the lexical classification implemented in Iramuteq, which makes it possible to highlight the themes specific to a corpus and to group together "lexical worlds," can correspond to this theme-characterization goal (it seeks to "reflect the internal order of a discourse, to highlight its lexical worlds").[2]

### 2.3. Material

In order to illustrate these choices and the bringing into coherence of the different models, I will now offer an example-based illustration. I focus on manually corrected automatic transcripts of "morning" interviews: a dataset currently made up of 3166 interviews corresponding to 561 political figures interviewed between 10 June 2016 and 4 December 2017 (representing approximately 10 million words). Echoing other work done together with André Salem on the corpus of the French newspaper *Père Duchène* (Longhi and Salem 2018), the analysis here focused on the term "*ennemi*" ("enemy").

The *Père Duchène* corpus can be downloaded on the website of the Lexico software (http://www.lexi-co.com/L3-5.Textes.html); the "morning interviews" corpus is not yet available online, but will be published during the TALAD ANR project r https://web.u-cergy.fr/anr-talad/), on an infrastructure including a repository of language data (probably Ortolang: https://www.ortolang.fr/). These questions and methodological choices will here be applied, in an illustrative way, to the corpus of the TALAD ANR project. To provide some context for this choice, the TALAD project aims to show how Natural Language Procesing (NLP) (*Traitement Automatique des Langues [TAL]*) allows Discourse Analysis (*Analyse du Discours [AD]*) to go further in its explorations, test its theoretical apparatus, and reinforce its methodological tools. The aim is to adapt NLP techniques in order to provide DA

---

2　datahist.hypotheses.org/11.

with more complex sets of descriptors relative to different levels of discursive organization. In return, DA will provide a range of complex phenomena to be studied, which will represent just as many challenges to be addressed by the latest advances in NLP.

## 3. Results

In order to examine this term (here treated as the lemma /ennemi/), I survey all lemmas (see Appendix A) in order to then be able to compile the concordance of all text segments containing this lemma. Then I can extract a subcorpus that brought together all these text segments, presented in Figure 2:



**** *theme_J_Chirac *theme_Patron_du_parti *theme_Extreme_droite *theme_General_de_Gaule *theme_Gaullisme *theme_Synthese *theme_L_Wauquiez *theme_Constructifs *theme_Veut_rassembler_et_exclut *theme_Plus_de_discours_a_droite *theme_Que_la_France_se_remette_a_marcher *theme_E_Macron *theme_P_Seguin *theme_N_Sarkozy *theme_JP_Raffarin *theme_Europe

c est toujours un espoir d aller comme disait la gauche jamais d **ennemi** à gauche à droite on dit

**** *theme_J_Chirac *theme_Patron_du_parti *theme_Extreme_droite *theme_General_de_Gaule *theme_Gaullisme *theme_Synthese *theme_L_Wauquiez *theme_Constructifs *theme_Veut_rassembler_et_exclut *theme_Plus_de_discours_a_droite *theme_Que_la_France_se_remette_a_marcher *theme_E_Macron *theme_P_Seguin *theme_N_Sarkozy *theme_JP_Raffarin *theme_Europe

jamais d **ennemi** droite hé bien on retrouve les mêmes choses mais encore une fois à l époque il y avait un patron qui fixait et ce patron il était parfois contesté mais il était admis par tout le monde

**** *theme_Proces_Merah *theme_Justice_laxiste *theme_Bracelet_electronique *theme_D_Trump *theme_Attentat_New_York *theme_Peine_de_mort *theme_M_Le_Pen *theme_L_Wauquiez *theme_Liberation_534_elus_FN_de_2014_ne_siegent_plus *theme_Une_France_village_d_Asterix *theme_Viol_a_Calais *theme_Refondation_du_FN *theme_Alliance_LR_FN *theme_J_L_Melenchon *theme_A_Berge *theme_Sens_Commun *theme_Europe *theme_Opposition_a_LREM *theme_E_Macron *theme_Sondage_IFOP

écoutez donald trump réagit à la mesure de ces attentats qui touchent le peuple américain c_est_à_dire qu il a pris conscience de qui était l **ennemi** d où il venait au nom de quoi il combattait

**Figure 2.** All the text segments with *ennemi*.

And I use a "building corpus" function to extract a subcorpus presented in Figure 3:



**** *theme_Francois_Hollande *theme_Candidature_presidentielle *theme_Primaires_a_Gauche *theme_TAFTA *theme_CETA *theme_Russie *theme_Poutine *theme_Alep *theme_Syrie *theme_Al_Nosra *theme_Daech *theme_Mossoul *theme_Notre_Dame_des_Landes *theme_Manuel_Valls

qui s intéressent superficiellement aux choses me coller une étiquette de cette nature c est douloureux mais ça ne fera pas changer d avis monsieur bourdin les russes ne sont pas nos **ennemis** mais nos partenaires et je ne veux pas la guerre avec les russes

**** *theme_Syrie *theme_Poutine_Bacha_El_Assad *theme_Daesh *theme_Europe *theme_Pays_est_europeens *theme_Grand_congres_des_consciences_europeennes *theme_Libye *theme_Donald_Trump *theme_Defense_des_interets_francais_et_promotion_de_la_paix

qui est notre **ennemi** numéro un aujourd hui en syrie l état islamique ou bachar al assad daech et c est daech que nous bombardons ce n est

| Cancel | Save | Construire un sous-corpus |

**Figure 3.** Extraction of the subcorpus.

The concordancer displays the context in which the word is used in the corpus. Iramuteq offers relatively broad contexts, which makes it possible, by grouping all the concordancer results, to provide a subcorpus centered on the searched term, used in its context and context of use. This give a specific subcorpus centered on the segments containing the lemma /ennemi/ (Table 1):

**Table 1.** Lemmas of the subcorpus[3].

| Form | Freq. | Types |
|---|---|---|
| ennemi | 135 | common name |
| finance | 19 | common name |
| France | 19 | proper name |
| droite | 17 | common name |
| daesh | 16 | proper name |
| guerre | 13 | common name |
| état | 12 | common name |
| aller | 11 | verb |
| national | 10 | adjective |
| penser | 10 | verb |
| république | 10 | common name |

Of course, this procedure leads to the loss of some of the interviews' wider contexts but, nevertheless, since the left and right contexts are of a suitable size, this stage allows to focus the attention on /ennemi/ itself. In order to describe its profiling, I can use the similarity analysis defined earlier, giving the following result presented in Figure 4:
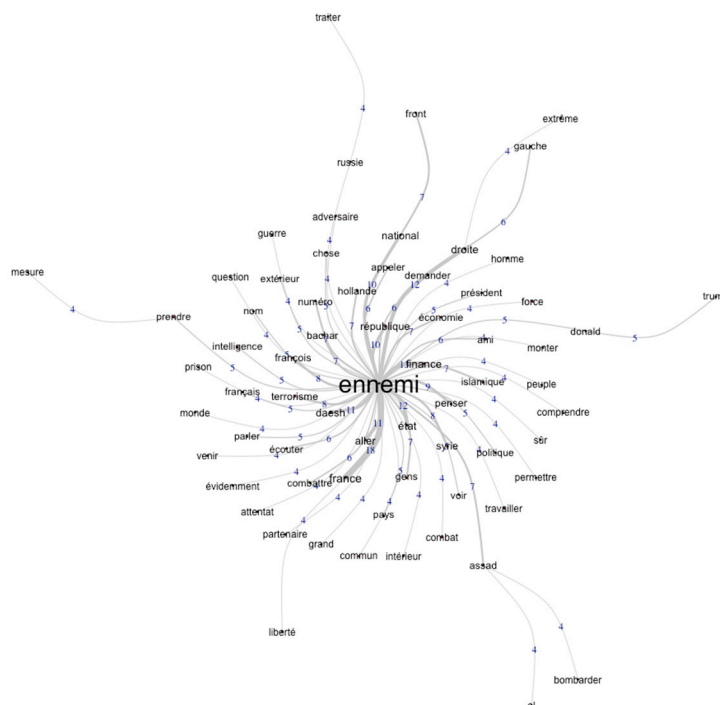


**Figure 4.** Similarity analysis.

This result makes it possible to observe the polysemy of a term, and to be able to "objectify" the instability of meaning by describing it in a quantified manner and based on indices of co-occurrence. There are frequent co-occurrences with the lemmas /france/, /daesh/ ("Daesh" or "ISIS"), /république/, /bachar/ ("Bashar"), and /islamique/. Here, the goal is not an exhaustive analysis of the co-occurrences around this lemma, but rather, to describe the way in which this similarity analysis can give access to "new usages already recorded in a lexicon" or occurrences "within new themes." Indeed, I find different ways in which /ennemi/ or "enemy" is characterized in this corpus of morning political shows: an enemy of values (French Republic), a political enemy (France, Bashar), a religious enemy (Islamic), or a political-religious enemy (Daesh or ISIS). These results are very close to the way in which the "Theory of Lexical Concepts and Cognitive Models (LCCM Theory)" (Evans) describes polysemy. For example, Evans (2018, about prepositions such as *in* and *on*) "argue[s] that the 'state' lexical concept associated with *in* selects for co-occurring open-class lexical concepts which access conceptual structure concerning emotional or psychological 'force' such as being 'in love', 'in pain' and so on. In contrast, the semantic arguments that co-occur with *on* relate to content that has to do with time-restricted activities, as well as actions that involve being currently active," which "suggests that each of the prepositions is associated with a distinct 'state' lexical concept." Here, my tooled-based approach allows us to show meaning constructions and linguistic representations. This also echoes the conclusions of textometric approaches (Mayaffre 2007) which use co-occurrence from a hermeneutical perspective.

These profiles open the way to the themes within which this lemma occurs and this is illustrated in the thematic classification that could be produced on the same subcorpus, presented in Figure 5:

---

3    Translation of the words: "enemy," "finance," "France," "[the] right," "Daesh," (ISIS), "war," "state," "go," "national," "think," "republic".
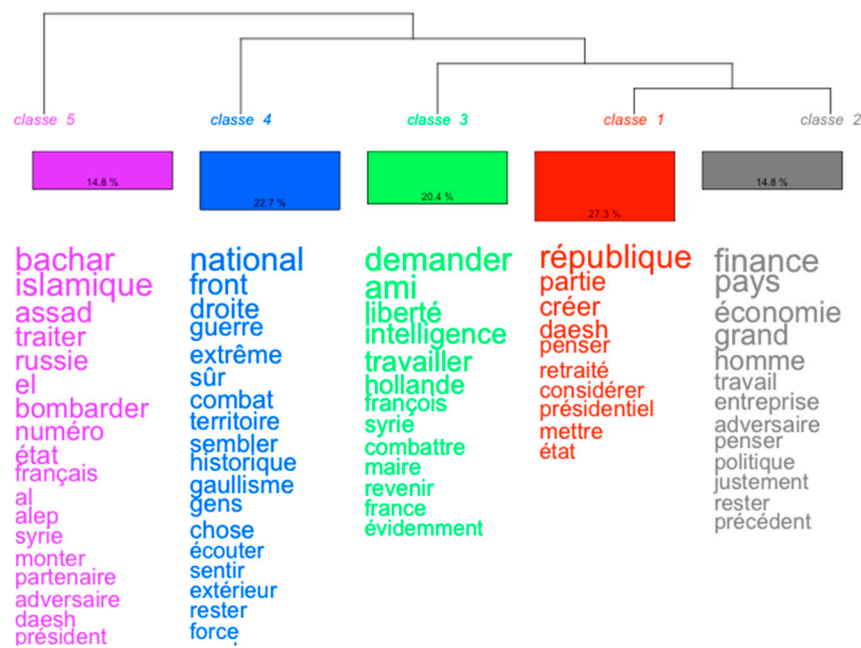
**Figure 5.** Reinert-type classification proposed by the Iramuteq software.

Five main classes are identified which can also be represented using FAC proposed in Figure 6:
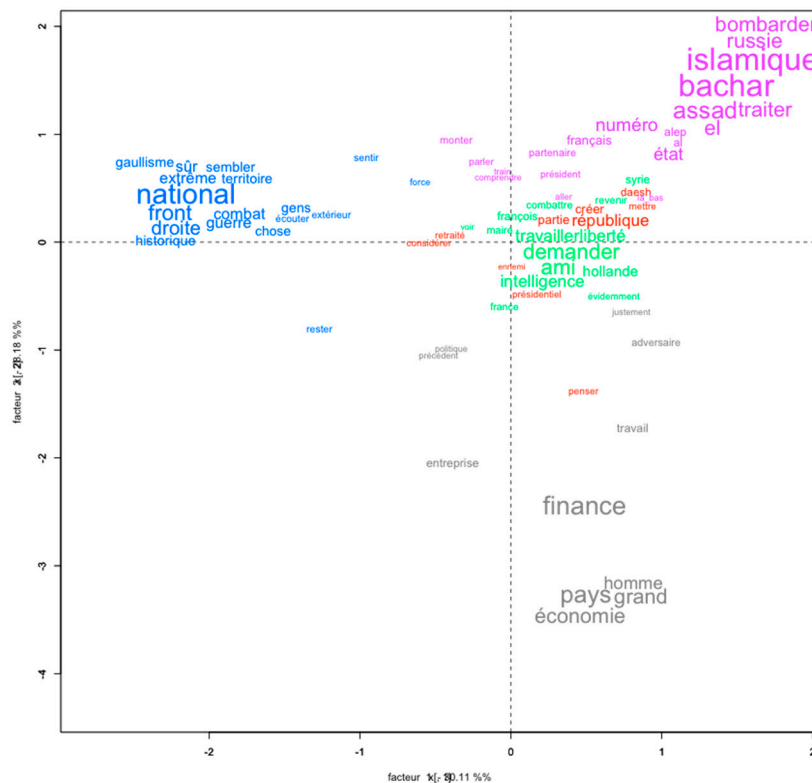


**Figure 6.** Factorial analysis of correspondences (FAC) of the thematics.

One characterization of "enemy" is its anchoring in the far right and the French National Front; another is related to finance and economics; another concerns the Islamic State and Bashar al-Assad; one relate to the French Republic and Daesh (ISIS); and a final one contains the lemmas /ami/ ("friend") and /liberté/ ("freedom") and which deserve to be contextualized.

I can then return to the corpus with the text segments typical of the different classes and understand the profiling and thematization of /ennemi/. Here are some examples of class 3 which will help better understand it (the highlighted terms are typical of this class):

1. *c_est_à_dire que des gens qui **reviennent** de **syrie** qui sont allés égorger qui sont nos ennemis **reviennent** en **france** en **liberté** je suis là aussi un des rares à **demander** leur arrestation immédiate auprès du tribunal pour **intelligence** avec l ennemi* ("That is to say that people who **return** from **Syria**, who went to slaughter people and are our enemies, **return** to **France** and keep their **freedom**. I am again one of the few **asking** for their immediate arrest and prosecution for **colluding** with the enemy.")

2. *donc vous **demandez** vous aussi aux **maires** de **france** de ne pas accorder aux ennemis islamistes de la **liberté** la moindre parcelle de **liberté** mais **évidemment*** ("So you too are **asking** the **mayors** of **France** not to grant the Islamist enemies of **freedom** the slightest ounce of **freedom**? **Obviously**.")

3. *alors pour vous qui sont les rebelles d alep ce sont des **amis** ou des ennemis de la **france** on aurait dû les aider ou pas on les a aidés malheureusement* ("Who are then, for you, the Aleppo rebels? Are they **friends** or enemies of **France**? Should we have helped them or not? Unfortunately, we did help them.")

Class 4 included the following examples:

4. *et que pour moi comme tous ceux qui étaient attachés au **gaullisme historique** ou comme ceux qui avaient une histoire centrée ou comme les humanistes de **droite** le **front national** l **extrême droite** était et **reste** un ennemi* ("For me as for all those who subscribed to **historical Gaullism** or those with a centered background or the humanists of the **right**, the **National Front**, the **far right** was and **remains** an enemy.")

5. *moi mon ennemi c est le **front national** d_abord parce_que à paris vous ne le **sentez** pas mais en province ça monte et vous savez pourquoi ça monte parce_que les **gens** se **sentent** complètement abandonnés* ("My enemy is the **National Front**, first because in Paris you don't **feel** it but in the provinces it's on the rise. And do you know why it's on the rise? Because **people feel** completely abandoned.")

6. *notre ennemi à nous c est bien l **extrême droite** avant toute chose et bien sûr la **droite** représentée par françois fillon c est bien l **extrême droite** qui a le projet le plus dangereux pour la france et c est bien la **droite** qui a le projet le plus inégalitaire* ("Our biggest enemy is first and foremost the **far right** and of course the **right** represented by François Fillon is very much the **far right**, which has the most dangerous plans for France, and it is very much the **right** whose ideas are most riddled with inequality.")

The semantic dynamics in the sense of meaning trajectories as illustrated in the similarity analysis, supported by the thematic analysis, and detailed in the corpus examples, can be understood now.

By adopting the principle of variation favored, I can compare the analysis of this contemporary political corpus with the results of another analysis (Longhi and Salem 2018) on the *Père Duchesne* corpus (made up of issues edited by Jacques-René Hébert between 1793 and 1794; see also Salem 1988). By studying repeated segments linked to /ennemi/, the analysis shows that the earlier phrases *plus cruels ennemis, plus mortels ennemis, ennemis du dehors* ("cruelest enemies," "deadliest enemies," "outside enemies"—i.e., foreign powers, expatriates) were later followed by *les ennemis du dedans et du dehors* ("enemies inside and outside"—outside enemies were not the only danger) and then by *ennemis de l'intérieur* ("inner enemies") which completed the notion of *ennemis du dedans* ("enemies inside"). Gradually, *nos ennemis* ("our enemies") became *vos ennemis* ("your enemies") and then *les ennemis* ("the enemies"). Towards the end the enemies—now preferentially designated in the plural—were no longer qualified by their location or by their relationship to the message recipients ("our/your enemies") but by assumed common values that they were meant to oppose: *ennemis du peuple* ("enemies of the people"), *ennemis de la république* ("enemies of the republic"), *ennemis de la révolution* ("enemies of the revolution"), *ennemis de la liberté* ("enemies of freedom"), *ennemis de l'égalité* ("enemies of equality").

We can therefore see convergences between these corpus analyses which could be distinguished by their temporal anchoring as well as type: an inner/outer distinction; the question of values; and

taking into account a viewpoint when designating the enemy. Thus, these studies provide us with information on both the meaning and context of specific corpora, but also on the ways of interpreting meaning and the stabilization processes occurring within themes.

## 4. Discussion

By questioning the various accepted meanings of "digital" and "humanities" and by putting into perspective various ways of conceptualizing or practicing the digital humanities, this article will seek to show that

- on the one hand, the humanities cannot keep using computing simply as a reservoir of tools without knowing how they are actually designed (the "black box"), or why and how they are relevant to their research—otherwise they will lose their own distinctive mark within them;
- on the other hand, computer scientists cannot keep blindly applying tools which work properly elsewhere and are declared applicable in the humanities—otherwise results in the humanities will lose in quality. Indeed, the interpretation of textual data is subject to semiotic constraints (discursive practices, discourse genre, intertextuality, etc.) and it is necessary to be able to characterize, before any computer processing, the corpora. As we have seen, it is also fundamental to maintain coherence between the types of models called upon, in particular so that the results obtained are really answers to the scientific questions asked.

We therefore need to think of the digital humanities as a co-construction of objects, knowledge, and tools as we consider, in a reasoned and mutual way, issues and expectations pertaining to the humanities and computing.

### 4.1. Humanities, Human Sciences, and Human and Social Sciences

In his article, "Sciences Humaines" in *Encyclopædia Universalis*, Edmond Ortigues (1979) explains the institutionalization of the French human sciences:

> A decree issued on 23 July 1958 (published in the Journal Officiel [official gazette] on 27 July 1958) turned faculties of arts [*lettres*] into faculties of arts and human sciences [*lettres et sciences humaines*] with the aim of encouraging some of the social sciences (psychology and sociology) to be taught in the proximity of arts and humanities subjects [*humanités littéraires*]. The phrase "human sciences" in this academic sense–which has come into widespread use–is a typically French idiomatic expression. The English language uses it in fairly loose contexts and speaks more commonly of "social sciences".

However, the "human sciences" (*sciences humaines*) or the "arts and human sciences" (*lettres et sciences humaines*) or even the "arts and human and social sciences" (*lettres, sciences humaines et sociales*) are not the same as the humanities (*humanités*), even though the former could be seen as an institutional transposition of the extent of the latter. Indeed, as Ortigues reminds us in relation to the humanities, "the human mind manifests itself in its works. Historically, the arts and humanities [*humanités littéraires*] which study the *works* of the mind came before the birth of the social sciences [*sciences sociales*] which seek to study more directly human *activities* (by means of observation and hypothesis)": "the Latin word *humanitas*, when translating the Greek word *paideia*, means "culture," "education," "civilization"; the medieval concept of tradition "was of an ecclesiastical or legal and theological nature: the transformation of the concept of ecclesiastical tradition into that of a humanist or cultural tradition went hand in hand with the development of philological and historical criticism. This observation is interesting because it allows us to see the humanities as an integral part of the cultural sciences (*sciences de la culture*) as defined by François Rastier (2004) and gives us a methodological insight because, according to him, "cultures can only be described differentially, like the cultural objects—above all languages and texts—that make them up."

The centrality of texts within the scope of the cultural sciences and the humanities, in interconnection with philology and the textual sciences (*science des textes*), and more broadly, a semiotics that would take into consideration all sign systems, leads us to adopt a reflexive approach to characterizing the digital humanities themselves: describing them as "transdisciplinary" and "an intersection" drawing on HSS introduces a perspective which is likely to restrict both the very nature of this field and the practices developing within it. Thus, the digital humanities need to be conceived in a way that is less initially focused on their place within the field of the humanities from a disciplinary viewpoint and more in line with a natively digital approach (based simultaneously on objects, issues, resources, models, etc.). This, in fact, means understanding, at least reflexively, but ideally, as soon as DH are conceptualized and therefore, as soon as research emerges, that digital technology changes the view that the humanities have of their own field, acting not as an added dimension, an "added value," or a "bias" depending on the way in which its contribution can be described, but as a prism that recursively remodels the field of the humanities. This does not mean dismissing everything that the humanities produced before DH, but it does mean considering that the existence of DH changes the ecology of the humanities.

These leads to two distinctions being drawn:

- Regarding the nature of the field, the HSS prism (or AHSS if we add the arts) leads us to conceive the digital either as an area or a means, not an approach angle, paying attention to what underlies this area or means (I will come back to this later);
- Regarding practices, we are often faced with the use of generic tools that can be applied to the "humanities"; however, the computer sciences often develop specific projects on similar objects but from a perspective other than that of DH.

Rather than being seen as inter-/trans-/multidisciplinary, DH can be seen as a new research field or paradigm with its own academic understanding which is distinct from that of the humanities or computing. As for the tools themselves, these have to be specifically designed to meet DH needs and integrate the humanities' view of digital technology. This observation on DH must necessarily be accompanied, just as importantly as for the humanities, by a definition of what is meant by "digital" and the wider use of computing. Computing must work hand in hand with the humanities, rather than just the humanities use computing for its digital needs.

### 4.2. Computing, Numerical, Digital: The Data in Question

Milad Doueihi (2015) gives us the following description of the passage from computing to the numerical: [ . . . ]

From computing (which has obviously not disappeared completely) to the numerical we go from one type of technicality, which is often exaggerated and cultivated for itself but requires a certain degree of technical skill, to more common uses requiring other skills which are put to use by a new online sociability built on texts and driven by "shares." Nowadays, it is in relation to this popular numerical practice that the work of the numerical humanities also needs to be conceived. (711)

Indeed, we know, for example, that "the messages posted on Twitter, although each limited to 140 characters, amount to 7 TB per day which is the equivalent of half the collection of the French National Library" (Ganascia 2015) and that more and more tools, projects, and research focus on the analysis, extraction, and representation of numerical social data. An increasing number of software programs which are sometimes quite intuitive and functional, well documented, and enriched with tutorials give access to an advanced numerical analysis practice. I hail, for example, the development of the Iramuteq software (which I use extensively)—an R interface which gives access to statistical calculations and classifications without having to go into the configuration of functions or algorithms. However, in order to have a clear view of the results produced and the quality of the data (depending

on the research objectives), it is important to know the specificities of the various statistical, probabilistic functions and consider the impact of the data representation, visualization, and exploration systems.

And conversely, the computer sciences approach certain tasks or objectives pertaining to the numerical humanities (NH) as research objects but only rarely do they play an integral part in NH projects except in the form of the "engineering" side of such projects. There can be a "utilitarian" bias in the humanities towards the numerical which explains, for example, why one of the chapters in Mounier's recent work (Mounier 2018) is entitled "What the computer brings to the humanities" (*Ce que l'ordinateur apporte aux humanités*) (45). Obviously, it is not the computer as such that contributes to the humanities, rather, the different programs, resources, tools, etc., that are developed. However, with the growing ease of use of computing tools, the design of the tools tends to be sidelined in favor of their simplified use. But behind these tools, there lies fundamental research which mobilizes research objects from a computing viewpoint. This "forgetting" of fundamental research produced by computer scientists has been picked up on by the DAHLIA Working Group (WG) (dahlia.egc.asso.fr) whose "aim is to bring together players (researchers and institutions) who, in the context of the numerical humanities and cultural heritage, are interested in managing but also analyzing data for the purpose of producing knowledge." It is therefore important that such a community specializing in "data management" and often interacting with data processing specialists be in permanent contact with humanities researchers in order not only to create turnkey inductive tools (such as machine learning) but also to develop, within the very problematization of its objects and methods, ways to manage/mine semiotic data considered within their own signifier environment.

This group—supported by the *Association EGC* (*Extraction et Gestion des Connaissances* or Knowledge Mining and Management)—is founded on the observation that "many researchers in the data analysis and management community are developing new models, algorithms, and software which help to effectively process complex data" and "these innovative tools are often developed by working jointly with researchers from disciplines other than computing, in particular the human and social sciences (HSS): information and communication sciences (ICS), sociology, history, geography, etc." However, "despite this enthusiasm, it must be noted that the NH trend has not been given much prominence in the French world of computing even though it is very much alive in some HSS disciplines such as ICS." It must be recognized that the computing work pertaining to the humanities needs to rise to the challenge of considering signifier data and not just signifiers (words, images, constructions); the algorithms and tools developed need to ensure that the signified is preserved in their understanding of the data so that meaning can be conceived as an achievable research object.

This WG brings together "computer scientists working on this type of questions in close collaboration with partners from other disciplines stemming from HSS." The change in perspective is interesting since heritage, as a pilot field of study, is considered a "subfield of the human and social sciences" made up of "oeuvres, bibliographical documents, and analyses/studies carried out on the oeuvres;" "in this context, science and technology can help solve cultural heritage problems in the management and analysis of data which give rise to research questions."

Thus, from the viewpoint of computing researchers, computing cannot be confused with numerical technology since, according to Dacos and Mounier (2014), in the case of the numerical humanities, the numerical can have three meanings: "The numerical as a research tool; the numerical as a communication tool; and the numerical as a research object," making up a complex "called on to redefine all of the research fields in the human and social sciences" (6). According to the authors, in the French case, there are numerical humanities players "but none of them is structured as a "center" as we now understand it" (43). But behind this problem of the polysemy of "numerical" there also lies a fundamental epistemological question which Meunier (2017) sums up well when stating that "there are no so-called numerical humanities if there is no formal modelling containing computable mathematical symbolic systems that can be translated into algorithms." For him, "any serious Numerical Humanities project, by using computers, implicitly or explicitly implements some formal modelling." It is this implementation that needs to be highlighted because it is in these formal

modelling processes that the links between the humanities and computing are played out via numerical technology. The definition put forward by Dacos and Mounier shows that the numerical is an epistemic construct, an object, and/or a tool. But in none of its senses can it therefore be confused with computing or the "digital". The challenge of NH is perhaps to consider that the numerical both describes and categorizes the humanities by the way in which it mobilizes knowledge and methods, and induces a different conceptualization of its object from that proposed by the humanities on the one hand, and by computing on the other. This point is a complex subject which is part of ongoing reflection, but the fact remains that this characterization of the numerical leads us to rethink certain dichotomies which are sometimes caricatures of the relationship between the numerical and the humanities.

### 4.3. *Qualitative/Quantitative, Conceptual/Formal, Semiotic/Numerical*

On these epistemological issues, the work carried out by Jean-Guy Meunier at UQAM is important. Meunier (2014) traces the genealogy of the numerical humanities but adds a comparative dimension based on spheres of emergence:

> In the 2000s there emerged, in conjunction with the emergence of computing technologies, an original and innovative research program deftly called the *Digital Humanities* [in English in original]. For the English-speaking world, this label took research in a more general direction than allowed by the previous label of *Computers and the Humanities* [in English in original]. In the French-speaking world, the English phrase which was translated as "*humanités numériques*" ["numerical humanities"] is more recent and raises questions not so much about the qualifier "digital" as about the noun "humanities." As we know, the terms "*humanités*" and "humanities" do not cover the same disciplinary areas in the two languages.

He explains that in English, "humanities" is "a traditional term covering a part of the complementary set of the so-called "hard" sciences", and in the francophone, Italian, and Spanish worlds, the humanities "are generally more likely to refer to an intellectual and even ethical tradition of an Erasmian-inspired humanist nature rather than an academic discipline". The label "Computers and the Humanities" is interesting because it reverses the order in which the French phrase "*humanités numériques*" presents the Humanities and Numerical Technology/Computing. Even if this may seem a little counterintuitive, it is possible to grasp the numerical humanities directly through the prism of computing, that is, study objects can be grasped from the perspective of the data sciences, algorithms, and information processing. This is what, for example, many researchers of the 27th section of the *Conseil National des Universités* or CNU (National Council of Universities, a French national institution) do: their study objects are objects that are equally dealt with by the humanities. Often, in order to contrast these approaches, debates focus on qualitative analyses done by the humanities versus quantitative analyses done by computing.

However, the separation between quality and quantity does not overlap with the distinction between the humanities and computing. In these fields, there are intellectual traditions, trends, schools of thought, and approaches which are particularly based on specific conceptions of knowledge, science, and the representation of knowledge. All this makes it possible to address the different models that can govern academic work. This is described by Meunier (2018) and leads to considerations that change our relationship to computing. Indeed, if computing as it is understood by the numerical humanities is based on a computational model, this computational model is based on a formal model and the choice of a specific formal model is not a neutral one:

> a formal model does not necessarily have a quantitative dimension. Thus, we can have logical, geometrical, topological, grammatical, etc. formalisms. Some of them use iconic symbols (graphs, images, etc.). In all these formalisms, various types of symbols can be found, such as constants, variables, operators, etc.

Thus, what could distinguish a computing approach to the numerical humanities from an approach that prioritizes the humanities is not necessarily a primary focus on computers, but rather the choice of

an appropriate formal model that makes computation possible. Therefore, for the numerical humanities to exist and put forward a coherence, signifier data need to be included in the outlook of the chosen formal model which will make their computation by a computer possible. Indeed, "these formal mathematical models are omnipresent in both the natural sciences and the human and social sciences." Thus, when looking at the difficulty of the notion of computationality in semiotics, Meunier notes that "a computational type of semiotic theory can only exist if it draws on formal mathematical (not necessarily quantitative) modelling whose statements, formulas, or equations allow for computability. Only then can a semiotics be computational."

## 5. Conclusions

The mutual transformation of the relations between the humanities and computing is therefore a profound one, covering issues that are deeper than they seem. Thus, Jean-Guy Meunier (2019b) points out that "some humanities experts believe that computing technology merely provides new types of access to the field of the humanities whilst their analysis continues to be interpretive in nature." Thus, the work of the humanities can consist in commenting on the results produced thanks to a computer and "computing simply appears as a new tool that helps deal with semiotic objects through digitization, archiving, mining, etc." This would suggest that "computing does not seriously affect the integrity of classic humanities practices." However, the great epistemological importance of Meunier's work lies in the fact that it brings the articulation between the humanities and computing to another level. This shift converges towards a reorganization of the fields/disciplines I have already proposed by confronting the humanities with computing/numerical technology. Indeed, if these computing approaches "seen as mainly quantitative in nature are only a reshuffling of the cards for the human and social sciences," thereby confirming "a clear opposition between computing and the humanities,"[4] I believe it is possible to see things in terms of a co-construction at the level of the very conceptualization of research objects and issues. And in this way, discourse analysis has an important role to play: computational linguistics has important limitations when it is applied to discourse analysis: important subjects such as opinions, ideologies, and political distinctions must be characterized by their linguistic and semiotic features, and must be interpreted thanks to the theoretical and methodological hypothesis.

If discourse analysts can understand all the parameters invested in the numerical humanities, and maintain coherence between the different planes of their research, they will be able to produce significant results based on controlled, chosen calculations, which will provide a precise answer to their questions. Their interpretations will then be able to bring new knowledge on complex semiotic objects, while advancing fundamental research on the methodological and theoretical level.

**Conflicts of Interest:** The author declares no conflicts of interest.

---

[4]　Regarding the encounter between semiotics and computing, Meunier reminds us that "several formulations, often synthetic but sometimes simplistic or inadequate, can express this encounter between semiotics and computing as an opposition between: quantitative and qualitative, descriptive and interpretive, experimentation and interpretation, natural sciences and human sciences, *Naturwissenschaften* and *Geisteswissenschaften*, etc."

## Appendix A

| | | |
|---|---|---|
| heureusement | 137 | adv |
| inquiétude | 137 | nom |
| tendance | 137 | nom |
| électorat | 137 | nom |
| compétence | 136 | nom |
| exercice | 136 | nom |
| immobilier | 136 | adj |
| naturellement | 136 | adv |
| pollution | 136 | nom |
| présence | 136 | nom |
| soumettre | 136 | ver |
| unique | 136 | adj |
| confronter | 135 | ver |
| correspondre | 135 | ver |
| ennemi | 135 | nom |
| ficher | 135 | ver |
| former | 135 | ver |
| moitié | 135 | nom |
| oser | 135 | ver |
| évoluer | 135 | ver |
| banlieue | 134 | nom |
| carte | 134 | nom |
| coalition | 134 | nom |
| cohérent | 134 | adj |

**Figure A1.** Lemmas of the Complete Corpus.

## References

Cadiot, Pierre, and Yves-Marie Visetti. 2001. *Pour une Théorie des Formes Sémantiques. Motifs, Profils, Thèmes*. Paris: PUF.

Dacos, Marin, and Pierre Mounier. 2014. *Humanités Numériques. État des Lieux et Positionnement de la Recherche Française dans le Contexte International*. Paris: Institut Français/Ministère des Affaires Étrangères pour L'action Culturelle, Available online: http://www.institutfrancais.com/sites/default/files/if_humanites-numeriques.pdf (accessed on 15 November 2019).

Doueihi, Milad. 2015. Quelles humanités numériques? *Critique* 819–20: 704–11. [CrossRef]

Evans, Vyvyan. 2018. Conceptual vs. inter-lexical polysemy. An LCCM theory approach. In *Language Learning, Discourse and Cognition*. Studies in the Tradition of Andrea Tyler. Amsterdam: John Benjamins Publishing Company.

Flament, Claude. 1962. L'analyse de similitude. *Cahiers du centre de recherche opérationnelle* 4: 63–97.

Flament, Claude. 1981. L'analyse de similitude: une technique pour les recherches sur les représentations sociales. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition* 1: 375–95.

Ganascia, Jean-Gabriel. 2015. Les *big data* dans les humanités. *Critique* 819–20: 627–36. [CrossRef]

Le Deuff, Olivier. 2015. Les humanités digitales précèdent-elle le numérique? Les humanités digitales précèdent-elle le numérique? Jalons pour une histoire longue des humanités digitales. H2PTM 15. Available online: https://archivesic.ccsd.cnrs.fr/sic_01220978 (accessed on 15 November 2019).

Longhi, Julien. 2015. La Théorie des objets discursifs: Concepts, méthodes, contributions. HDR Thesis, Cergy-Pontoise University, Cergy, France.

Longhi, Julien, ed. 2017. Humanités numériques, corpus et sens. *Questions de Communication* 31.

Longhi, Julien. 2018. *Du discours comme champ au corpus comme terrain. Contribution méthodologique à l'analyse sémantique du discours*. Paris: l'Harmattan.

Longhi, Julien, and André Salem. 2018. Approche textométrique des variations du sens. Paper presented at the JADT 2018 Conference, Rome, Italy, June 12–15; pp. 452–58. Available online: http://lexicometrica.univ-paris3.fr/numspeciaux/special8/tutoriel2.pdf (accessed on 15 November 2019).

Loubère, Lucie. 2016. L'analyse de similitude pour modéliser les CHD. Paper presented at the JADT 2016 Conference, Nice, France, June 7–10. Available online: http://lexicometrica.univ-paris3.fr/jadt/jadt2016/01-ACTES/83440/83440.pdf (accessed on 15 November 2019).

Marchand, Pascal, and Pierre Ratinaud. 2012. L'analyse de similitude appliquée aux corpus textuels: Les primaires socialistes pour l'élection présidentielle française (septembre-octobre 2011). Paper presented at the JADT 2012 Conference, Liège, Belgium, June 13–15; Available online: http://lexicometrica.univ-paris3.fr/jadt/jadt2012/Communications/Marchand,%20Pascal%20et%20al.%20-%20L\T1\textquoterightanalyse%20de%20similitude%20appliquee%20aux%20corpus%20textuels.pdf (accessed on 15 November 2019).

Mayaffre, Damon. 2007. L'entrelacement lexical des textes. Cooccurrences et lexicométrie. *Journées de Linguistique de Corpus*, 91–102.

Meunier, Jean-Guy. 2014. Humanités Numériques ou Computationnelles: Enjeux Herméneutiques. Sens Public. Available online: http://www.sens-public.org/spip.php?article1121 (accessed on 15 November 2019).

Meunier, Jean-Guy. 2017. Humanités numériques et modélisation scientifique *Questions de Communication* 31: 19–48. [CrossRef]

Meunier, Jean-Guy. 2018. Vers une sémiotique computationnelle? *Applied Semiotics* 26: 75–107.

Meunier, Jean-Guy. 2019a. Digital humanities: meaning engineering or material hermeneutics? Paper presented at Guest lecture of the Institute of advanced studies, Cergy-Pontoise, France, April 9.

Meunier, Jean-Guy. 2019b. La rencontre du sémiotique et du "numérique": Le rôle d'une modélisation conceptuelle. *Semiotica*. In press.

Morin, Edgard. 1994. Sur L'interdisciplinarité. Available online: http://ciret-transdisciplinarity.org/bulletin/b2c2.php (accessed on 15 November 2019).

Mounier, P. 2018. *Les Humanités Numériques. Une Histoire Critique*. Paris: Éditions de la Maison des Sciences de L'homme.

Ortigues, Edmond. 1979. SCIENCES HUMAINES. Encyclopædia Universalis. Available online: http://www.universalis-edu.com/encyclopedie/sciences-humaines/ (accessed on 15 November 2019).

Rastier, François. 2004. Doxa et Lexique en Corpus–Pour une Sémantique des Idéologies. Texto! Available online: http://www.revue-texto.net/Inedits/Rastier/Rastier_Doxa.html (accessed on 15 November 2019).

Reinert, Max. 1983. Une méthode de classification descendante hiérarchique: Application à l'analyse lexicale par contexte. *Les Cahiers de L'analyse des Données* VIII: 187–98.

Salem, André. 1988. Approches du temps lexical. *Mots. Les langages du politique* 17: 105–43.

Salem, André. n.d. a. Séries Textuelles Chronologiques. Available online: http://lexicometrica.univ-paris3.fr/numspeciaux/special8/tutoriel2.pdf (accessed on 15 November 2019).

Salem, André. n.d. b. Tutoriels pour L'analyse Textométrique. Available online: http://lexicometrica.univ-paris3.fr/numspeciaux/special8/tutoriel1.pdf (accessed on 15 November 2019).

Vergès, Pierre, and Bouriche Boumédine. 2001. L'analyse des données par les graphes de similitude. *Sciences Humaines*. Available online: http://www.scienceshumaines.com/textesInedits/Bouriche.pdf (accessed on 15 November 2019).

Visetti, Yves-Marie. 2003. Formes et Théories Dynamiques du Sens. Texto! Available online: http://www.revue-texto.net/Inedits/Visetti/Visetti_Formes1.html (accessed on 15 November 2019).

Visetti, Yves-Marie. 2004. Le Continu en Sémantique: Une Question de Formes. Texto! Available online: http://www.revue-texto.net/Inedits/Visetti/Visetti_Continu.html (accessed on 15 November 2019).