What terms to express the categories of natural sciences in the *Dictionary of Medieval Scientific French*?

Cécile Le Cornec Rochelois

Centre de Recherche en Poétique, Histoire Littéraire et Linguistique
Université de Pau et des Pays de l'Adour, France

cecile.rochelois@univ-pau.fr

Fabrice Issac

Université Paris 13

fabrice@ldi.univ-paris13.fr

Abstract. This paper deals with several lexicographic problems arising from the choice of hypernyms in the *Dictionary of Medieval Scientific French* (Créalscience database). The generic names used by medieval scholars indicate categorizations which differ from modern classifications. We examine some cases selected from the medieval domain of natural sciences where we can notice a conceptual discrepancy between medieval and modern taxonomies. We first explain the kind of difficulties the editors encounter when they choose a generic name to give a medieval definition, by taking examples in the fields of botany (*arbre, herbe, courge*), mineralogy (*mineral, pierre, metal*) and above all, zoology (*poisson, coquille, ver, mouche, bête*). In order to avoid anachronisms and allow the users to understand conceptual gaps, several means are used by the editors of the *Dictionary*: indicating by an asterisk the words whose meaning has changed and specifying the identification corresponding to the modern categorizations at the end of definition or in an encyclopedic note. After presenting these lexicographic choices, we will wonder how Semantic Web can help to represent and understand this variable lexicon.

Keywords: anachronism; botanic; categorization; classification; conceptual discrepancy; generic names; hypernyms; lexicography; mineralogy; taxonomy; zoology.

Introduction

The *Dictionary of Medieval Scientific French* is a corpus based dictionary which aims to study the genesis of the scientific lexicon between the XIIth and the XVth centuries to reveal the lexical creations and semantic evolutions during this period.[1] It has been published on-line since 2014, with a first version of the letter C. The Crealscience website is based on a XML database, managed by the Basex engine of the University of Konstanz. The DTD, mostly TEI complient, is the result of joint work of the team, which includes linguists specialized in the Middle Ages language, historians of science and lexicographists. This project is behind a CMS dedicated to lexicography, Isilex, which has a consultation/edition interface but also a dictionary data validation module. Indeed, each writer of this collaborative platform is formally identified and a role is assigned. Before being published, any changes must be approved by an administrator. The tool developed uses standard technologies like (X)HTML/CSS/JavaScript on the client side and PHP on the server side. An XQuery engine allows database queries.

The corpus based on scientific texts in old French does not lend itself easily to lexicographic processing for linguistic and epistemological reasons: the identification of scientific terms and of the fields of knowledge themselves; the considerable diachronic variations which affect nomenclatures and taxonomies during the medieval period; the place to be granted to Greek, Latin and Arabic words and to their relations with this vernacular language which builds itself in a situation of diglossia such as scholars often have a Latin vocabulary larger than their French scientific vocabulary; the choice of lemmas among multiple variants. So many difficulties complicate the representation of the medieval scientific and technical lexicon, but the automatic data processing may bring more adapted solutions than a printed dictionary.

We have chosen to focus in this paper on a recurring problem in every scientific field: the choice of generic names to be used in the definitions in order to avoid anachronism.[2] Resorting to a category which does not belong to medieval scientific culture appears as inconsistent. Is it possible to use generic names which did not yet exist in French between the XIIth and the XVth centuries without betraying medieval scientific paradigms? While it is tempting to borrow the words used in the texts of the corpus, we soon realize the limits of this solution: the meaning of many words has changed. What terms should we choose then to express the categories through which the clerics of the Middle Ages saw the world? The choice made by the team was to avoid any reference to later classifications and to express medieval definition, even if formulations sound strange.[3]

---

[1] About the history of the project and its stakes, see Ducos, J., Salvador, X.L., « Pour un dictionnaire de français médiéval: le projet Crealscience », *Langages*, n° 183, septembre 2011, pp. 63-74.

[2] This difficulty was underlined in the introduction of the Lexicon of scientific language, the starting point of this lexicographical project, about medicine. *Lexique de la langue scientifique (Astrologie, Mathématiques, Médecine…). Matériaux pour le Dictionnaire du Moyen Français*, Jacquart, D., Thomasset, C. (dir.), Klincksieck, Paris (1997), p. IV.

[3] http://www.crealscience.fr/DFSM/fr/Projet [consulté le 01/03/2015]

Our purpose in this paper is to deal with these choices and their implications by considering cases which involve gaps between medieval and modern taxonomies, in domains which are a matter of natural science: botany, mineralogy and above all, zoology. We will first explain the kind of difficulties the editors encounter when they choose a generic name to give a medieval definition. Then we will set out the lexicographic solutions that have been proposed and the way Semantic Web can help to represent and understand this variable lexicon.

1. Lexical anachronisms, cultural anachronisms: the stumbling blocks of definition

The definitions including unattested terms in medieval French and referring to anachronistic notions are obviously to be excluded. So, for plants or animals, Linnaean definitions from the dictionaries of modern French are not suitable. The *coloquinte* was certainly not for the medieval clerics a "plante grimpante de la famille des Cucurbitacées, originaire de la Méditerranée orientale et dont le nom savant est *Citrullus colocynthis*" (TLFi). It does not seem wise to define fly as an *insecte*, oyster as a *testacé*, crab as a *crustacé*, dolphin as a *mammifère* or frog as a *batracien*: not only because these terms were not a part of the French scientific lexicon before the XVIth century[4], but furthermore, the corresponding notions do not seem to be relevant taxonomic criteria in the Middle Ages. It does not mean that the information which allows us to define these zoological categories today was ignored. So, the ancient knowledge on viviparity in dolphins and whales or on their udders was available in the medieval encyclopedia, because their authors did not ignore the aristotelian heritage. Nevertheless, these criteria which lead us today to distinguish fish from marine mammals did not define a special class.[5] If viviparity as well as the presence of udders is mentioned among other characterizations, they do not appear at the beginning of encyclopedic articles and give rise only to occasional links. The medieval texts in Latin as in French leave no doubt on this matter: dolphins and whales are fishes.

The modern distinction between *crustacés*, *testacés* and *céphalopodes* does not seem more relevant to define the concerned aquatic creatures in the *Dictionary*. The outlines of these categories seem nevertheless clearly drawn by Aristote: among the

[4] The first attestation is found in 1542 for *insecte* (FEW, vol. 4, p. 710a, *insecta*) and it does not yet correspond to the current sense because it includes gastropods, amphibians and lizards, as it will be the case until XVIIth century. *Testacé* in its zoological sense is used for the first time in French in 1578 (FEW, vol. 13, p. 282b, *testaceus*). Finally, it is necessary to wait until 1713 for *crustacé* (TLFi), 1791 for *mammifère* (FEW, vol. 6,1, p. 134b, *mamma*) and 1806 for *batracien* (TLFi).

[5] As Aristote never supplies normative definition of the genre of *cètes*, he does not introduce any exclusion between *cètes* and fishes. The cetaceans of the modern science do not exist as class different from fishes before XVIIIth century. See Zucker, A., « Étude épistémologique du mot κητος », *Les Zoonymes : actes du colloque international tenu à Nice les 23, 24 et 25 janvier 1997*, Publications de la Faculté des lettres, arts et sciences humaines de Nice (Centre de recherches comparatives sur les langues de la Méditerranée, 38), Nice (1997), pp. 425-454.

animals without blood, as opposed to animals with blood such as fish, the Greek scholar distinguishes soft animals (μαλακόι), close to our cephalopods, from animals with a soft shell as our *crustacés* (μαλακοστράκοι) and animals with scaly shell (ὁ στρακόδερμα) which correspond to our *testacés*. This antique taxonomy has left marks in the learned works of the XIIIth century. We find the expressions *omnis piscis et animalia mollis teste* used by Thomas of Cantimpré[6] or *animalia durae testae marina* by Vincent of Beauvais.[7] The group of the animals without blood is explicitly mentioned in the *Speculum naturale*, which copies Aristote, but also Pline who had already compiled the zoological books of Stagirite.[8] However, following the example of Pline, the medieval encyclopedists do not use this distinction.[9] It occurs in a heterogeneous chapter of Vincent of Beauvais about the diversity of fish, on the same level as the opposition between marine and freshwater fish or the category of fish which carry a stone inside their head. Blood animals are the subject of no peculiar chapter and do not constitute a visible group in catalogs of species. Even if the Dominican encyclopedists occasionally remind us that certain species are included in one of the three categories of bloodless animals, this information has no consequence on their classification: they are placed most of the time between two spindle-shaped fish. Encyclopedists generally include soft animals with *pisces*. For example, chapter 18 of the book XVII of the *Speculum naturale*, where bloodless animals are evoked, is entitled *De diversis generibus piscium* and *malaciae* are called a *genus piscium*. When we turn to French texts, the names used for these categories of animals seem to move further away from the Aristotelian model, especially as the uses vary. In 1267, Brunet Latin provides the following definition for *coquille*:

Coquille est un poissons de mer enclos en charsoiz come une escavris, et est toute raonde; mes ele l'ovre et clot quant ele viaut, et son manoir est au fon[t] de la mer.[10]

---

[6] Thomas of Cantimpré, *Liber de natura rerum*, Boese, H. (éd.), Walter de Gruyter, Berlin-New York (1973), VII, 1, p. 251 : *Omnis piscis et animalia mollis teste modicum dormiunt.* "All fishes and animals with soft head sleep little."

[7] Vincent of Beauvais, *Speculum naturale, Bibliotheca Mundi Vincentii Burgundi… Speculum quadruplex*, éd. de Douai 1624 (repr. Anastatique Akademische Druck- und Verlagsanstalt, Graz, 1965), 4 vol., XVII, 18, col. 1262.

[8] *Ibid.*: *Aristoteles* [...] *In quibusdam marinis non est sanguis, ut est saepia, et karabos, et omnia quae plures habent quatuor pedibus.* Plinius ubi supra. *Tria genera sunt aquatilium sanguine carentium. Primum, scilicet quae appellantur mollia. Deinde crustis tenuibus contecta, postremo testis duris conclusa.* "*Aristote* […] Some animals have no blood as the cuttlefish, the spiny lobster and all those who have more than four legs. *Pline*: There are three genres of aquatic animals without blood. At first those who are called soft animals, then those protected by thin crusts, finally those locked in hard shells."

See Pline, *Histoire naturelle*, IX, 83, de Saint-Denis, E. (éd.), Les Belles Lettres (CUF), Paris (1955), p. 64.

[9] About the warping of the Aristotelian classification of *crustacés* and *testacés* in Pline's works, see *Hortus sanitatis: Livre IV*, Fishes, chapter 16, notes 1. https://www.unicaen.fr/puc/sources/depiscibus [Site consulted on 02/03/2015].

[10] Brunet Latin, *Tresor*, Beltrami, P. G., Squillacioti, P., Torri, P., Vatteroni S. (éd.), Giulio Einaudi, Turin (2007), I, 133, p. 236.

As a matter of fact, this excerpt deals with the oyster. We nevertheless understand throughout the text that the word *coquille* is also suitable as hypernym for the *pourpre* (*Murex* Linnaeus, 1758) and the crab.[11] *Testacés* and *crustacés* are thus mixed-up under this metonymical name, which gives the animal the name of the shell. In 1372, Jean Corbechon's translation of the *Liber of proprietatibus rerum* gives two more precise generic expressions for *testacés*: "oistres, molles et aultres poissons qui ont forte escaille" translates the latin *ostreae et alii quidam pisces in conchis degentes* and "une manière de poissons en ostree ainsi comme oystre" is used as a substitute for the two names of species *murices* and *conchylia* which are found in the text of Bartholomaeus Anglicus. In spite of a generalization attempt, the translator does not seem to worry about naming a defined class.[12] While both excerpts deal with oyster and similar creatures, the formulation varies: within a few lines, "poissons qui ont forte escaille" become "poissons en ostree ainsi comme oystre," which seems to indicate the absence of stable terminology in French for this category.

The modern zoological terms of classification which we have just evoked are linked to scientific data which were often collected in the learned works of the Middle Ages without being classification criteria. The lexical anachronism then involves abstract anachronism.

The choice of the chronological border between the relevant terms and the terms considered as anachronistic should be questioned. First of all, why would terms not attested in French before 1500 not have their place in the *Dictionary*? For instance, the concept corresponding to the Linnaean family of *Cucurbitaceae* has obviously existed since medieval times. Nothing challenges the equivalence between species grouped under the hypernym *courge* and the *Cucurbitaceae* of the modern botany, also named *courge* in modern French. However, the simple mention in the definition of a term referring to the Linnaean classification would be enough to distort things and *courge*, which still has a generic value in French, seems more adequate. A word like *amphibien* is undeniably anachronistic from a lexical point of view: it is Rabelais who introduces this Hellenism in French in 1553. It does not prevent medieval scholars from mentioning the customs of the animals which, according to the expression of Jean Corbechon, "vivent partie en eaue et partie en la terre" and "nagent et vont sus la terre si comme font les cocodriles et les chevauls d'yaue et mout d'autres qui vivent en terre et en yaue". The phrase compensates for the absence of an adjective; by excluding on principle *amphibien* from the elements of definition, the editors of the *Dictionary* condemn themselves to a circumlocution which can seem curious to readers who know it as a common term today. The precautionary principle consisting in excluding the unattested words before 1500 is certainly an inconvenience: common terms like *amphibien*, *carnivore* or *migrateur* are excluded even though, conceptually, they do not seem anachronistic.

---

[11] *Ibid.*: "Une autre coquille est en mer qui a nom morique, et li plusor l'apellent oistre, por ce que quant ele est taillie environ lui il en issent larmes, de quoi l'en taint les porpres; et cele tainture est de son charcois. Une autre coquille est que l'en apele cancre, por ce que il a jambes et est raonde ;"

[12] We notice moreover that this process is not systematic: the Latin *cancri et huiusmodi* is translated by "escrevices, escrevices de mer".

On the other hand, we can wonder if certain words attested before 1500 are suited to the state of the knowledge during the four centuries which preceded? Is it advisable for example to use the adjective *mineral* in the *Dictionary*? According to etymological dictionaries, the first occurrence of the noun meaning "corps inorganique qui se trouve dans l'intérieur de la terre ou à sa surface" would be in 1538 and the use of the adjective in 1516. But an earlier occurrence is mentioned by the *Dictionary*, in the works of Nicolas Panis in 1478 about arsenic:

Arcenic et orpigment, ce sont mineralx et sont sublimes et sont chaulx ou tiers, secs ou second et oultre [...]. (Nicolas Panis, Guidon,1478, tr.VII, doct.1, chap.7)

Thus, "matière minerale" seems appropriate to define arsenic, whose classification is all the more delicate as its modern definition requires knowledge of chemistry.[13] The reference to Nicolas Panis provides a contemporary scientific guarantee. But does that justify extending the use of the noun or the adjective to definitions of the other terms referring to stones or metals? A systematic use of the expression "matière minerale" is not relevant because the categories "pierre" and "métal" are more precise. Furthermore they seem to correspond better to the representations of medieval scholars. The distinction between stones and metals are very clearly formulated by Jean Corbechon.[14] We might be tempted to choose "minéral" for materials which, following the example of arsenic, are neither metals nor stones, such as antimony. However, antimony is mentioned in medical works which do not propose definitions or classifications. The question is whether antimony is part of the class of minerals such as medieval scholars conceived it. When the word *antimoine* was used on 1256 by Aldebrandin of Siena in his *Régime du corps*, was it already considered by learned contemporaries to be a mineral? In the time of Nicolas Panis, was there a common scientific idea of what constitutes a mineral? The use of "minéral" in the definitions supposes a lexicological work on the meaning of this term when it is used by scholars like Nicolas Panis and, more widely, on the meanings of the Latin word *mineralis* and their evolution throughout the Middle Ages.

Even if the chronological limit is of course debatable, by excluding as a matter of principle terms unattested before 1500, we limit to a certain extent the references to paradigms later than the Middle Ages. As regards zoology, the date of 1500 allows us in particular to not include terms which appear in French through naturalists like Pierre Belon du Mans or Guillaume Rondelet, in a pivotal period when zoological knowledge fundamentally evolves. We do not claim to define terms as medieval scholars would have and this choice creates difficulties in the formulation of definitions. But it seems essential to adopt these chronological restrictions in order to avoid anachronism and to help the editors to harmonize their definitions. Nevertheless this

[13] This is the définition found in TLFi (*Trésor de la langue française informatisé*): « Corps simple solide, de symbole As, d'aspect métallique, de couleur gris acier possédant à la fois des propriétés de métal et de métalloïde. »

[14] Jean Corbechon, *Livre des propriétés des choses*, Paris, BnF, fr. 16993, XVI, 75, f. 236ra.

precaution is not enough to avoid notional anachronisms because numerous terms were left in French with important semantic evolutions.

2. Lexicographical expressions of the conceptual discrepancy

One of the main goals of the *Dictionary* is to show the discrepancy between medieval and modern taxonomies. Three primary means were selected to indicate semantic evolutions and relationships with current terminologies: the asterisk behind the terms used in a medieval sense, the addition of a common name in modern French at the end of definitions and encyclopedic notes. The asterisk marks the terms whose meaning has evolved in comparison with modern language. It allows formulating definitions by means of taxemes which changed extension. Such is the case of *poisson** (fish), used in the broad sense of "créature qui vit dans l'eau" throughout the Middle Ages, and which appears for this reason in the definitions of dolphin or crab. The *coquille* mentioned by Brunet Latin will thus be defined as a "poisson* dont le corps est protégé par une coquille ou une enveloppe rigide."[15] Because of the formulations of certain authors of the corpus, the attention of the editors of the *Dictionary* is often drawn to taxonomic discrepancies, which incite them to use asterisks. We hesitate for instance to follow our first idea by defining the *artemisia* or the *aloe* as plants when Jean Corbechon presents them as *herbes*. Out of caution, we thus prefer *herbe*, accompanied with an asterisk which shifts to the article dedicated to this generic name the question of the relationship between the medieval meaning of *herbe* and the modern meanings of *plante* and *herbe*. If we refrain from using the name *insecte*, anachronistic, we can turn to the terms *mouche* (fly) or *ver* (worm) used by the clerics, since their generic value is well specified in the *Dictionary*.

This process which consists in tracing the medieval lexicon of the scholars not to deform their concepts has something reassuring; but what to make when all the authors of the corpus do not use the same hypernym? Is the medieval crab rather a *poisson**, as suggested by a majority of texts, or a *coquille**, as used by Brunet Latin? Yet this is only a hesitation on the extension of the hypernym: since a *coquille** is a *poisson**, there is no contradiction. Another more complex case reveals the difficulties that the editor can encounter choosing a generic name because of the instability of medieval terminology: the crocodile. Jean Corbechon presents it – rather logically from our point of view – as a *poisson*, next to the *cheval d'eau*, or, in other words, the hippopotamus. But Brunet Latin, who uses the term *poisson* to introduce the hippopotamus and mentions the crocodile within his inventory of aquatic creatures, prefers for the latter the more general term of *animal*. It is the word *beste* that appears in the quotation extracted from the *Chirurgie* of Henri de Mondeville and at the beginning of the note "La cocodrille" in the long version of the *Bestiaire* attributed to Pierre de

[15] The word *coquille* meaning "stiff shell" is already used in the Middle Ages and is thus in the definition without an asterisk.

Beauvais.[16] Further, we read in this last note that the crocodile is a "serpent marage". According to the texts of the Créalscience base, the medieval crocodile is thus at the same time an *animal*, a *beste*, a *poisson* and a *serpent d'eau.*[17] What category is it then advisable to select in the definition? Is it better to favor the majority use, by taking the risk of seeing this choice questioned by new reports, or to adopt the nomenclature of an author like Jean Corbechon, who defines most of the hypernyms and strives to be consistent? Another parameter must be taken into account, that of the coherence among entries. To define the hippopotamus as a *poisson** and the crocodile as a *serpent** would be inconsistent. The lexicographer cannot adopt a terminology as elastic and heterogeneous as his diverse sources. In this particular case, it seems reasonable to explain the taxonomic variations due to the hybrid nature of this animal in the note and to choose in the definition the term *animal* because it is the most neutral and most general, and the closest to modern use.

The asterisk actually raises another issue: its multiplication harms the legibility of the definition. Yet its presence could be justified after a large number of terms if we use it as soon as a discrepancy exists between medieval and modern knowledge. As regards gourds, as the list of the species quoted by medieval texts (*concombre*, *citrule*, *courge sauvage*, that is *courge d'Alexandrie*, that is *coloquinte*) does not correspond exactly to the species a modern botanist would recognize as gourds, and in the absence of an explicit medieval definition, the asterisk is imperative by caution: how can we be certain that medieval clerics were referring the same gourds that we are? Should we then put an asterisk after *corbeau*, *cheval* or *chien* because the medieval representation of these animals differs from the modern definition? To avoid the proliferation that the application of this principle could lead to, we prefer to limit the use of the asterisk to the terms which involve taxonomic gaps or whose semantic evolution can be confusing. So, besides *mouche*, *ver*, *poisson*, *herbe*, *arbre* and other generic names borrowed from medieval scholars, we append asterisks to terms like *lièvre* (which can be a rabbit) or *ongle* (used for the hooves of ungulate mammals).

The medieval definition can make certain familiar animals unrecognizable. It then proves useful to specify the identification of the animal defined by its current name. Would the reader be able to recognize the animal named *boterel* hidden under the definition "Ver* venimeux aux yeux rouges qui fréquente les lieux humides et subit une mue", if we do not mention that it is the *crapaud* (toad)? To make the consultation easier, this identification is added at the end of definition. However, it is not al-

[16] Henri de Mondeville*, Chirurgie,* 1314, éd. Bos, ch. 247, p. 71 ; *Bestiaire version longue, attribuée à Pierre de Beauvais),* 42, p. 194.

[17] Latin sources use hypernyms *animal*, *bestia* and *belua*. In Vincent of Beauvais's work, the crocodile appears in book XVII, that of the fishes, but in the section dedicated to the marine monsters, at the end, after fishes "qui pure naturam and speciem piscis habent". The crocodile is also found in Thomas of Cantimpré's book 6 of *De Natura rerum*, "De monstris marinis". Albert the Great places the crocodile in book 24 of *De Animalibus*, the inventory of aquatic species, among other *pisces*. He compares it with the lizard, without using any other generic name than *aquatici* or *belua*. It seems that the three Dominicans classify the crocodile among the aquatic species because it lives in water, even if their sources do not present it explicitly as a *piscis*.

ways convenient or possible. So the equivalence between the *cète* and the animal which we name today *baleine* is not at all obvious. At least this notation presents the advantage of making an identification possible for a reader knowing nothing of the antique or medieval *cète*. Editors of the *Dictionary* can use the encyclopedic note to explain the variable relationship between the textual creature and the real animal. Concerning *chamel leopard*, to assimilate the animal to the giraffe would be meaningless. Here is the definition which we can develop by compiling the medieval information: "Animal doux et beau, qui vit en Éthiopie et présente une tête semblable à celle du chameau, un cou de cheval, des pattes de buffle et un pelage tacheté comme celui du léopard". If this animal has given rise to such a strange description, as far as to look like a hybrid, a textual fancy, it is exactly because the connection with the real animal known in old French from the XIIIth century under the names *girafe* or *orafle*, inherited from the Arabic, was ignored. As it has been showed by Thierry Buquet, it will be necessary to wait until the end of the XVth century for the *camelopardalis* bequeathed by the antique knowledge and Deuteronomy to be identified with the giraffe.[18]

The *chamel leopard* raises another recurring problem. Its definition looks more like a description than a definition in compliance with lexicographical uses. It is due to the nature of the zoological knowledge (and to a certain extent botany) passed on by medieval works. The longest notes of bestiaries and encyclopedia proceed by accumulation of natural properties according to the compiled sources, so that heterogeneous elements from our point of views are mixed without explicit hierarchy. Let us use as an example the text dedicated by Brunet Latin to the crocodile in his *Tresor*, where we can find the following information: the crocodile is a four-legged animal of yellow color born in the Nile; it is twenty feet long and armed with big teeth and long claws; its skin is so resistant that it would not feel a blow from a stone; it lives on the ground in the daytime, in the river at night; it has no tongue and it is the only animal in the world able to move its upper jaw while keeping lower jaw immobile; it is a rival to the hydra. What should be selected in the definition? Collecting the elements in the corpus necessary to reconstitute a modern definition would mean deforming the representation expressed by the medieval text. If it is possible to distinguish striking properties, to select certain data either because they appear at the beginning of the notes, or because they are obviously recurring, we can understand the nature of a given animal, plant or stone in medieval culture and formulate organized definitions, reflecting the particularity of a medieval system of representations.

Thus for the eel, the link with the snake appears as essential information. Indeed the Isidorian etymology which connects *anguilla* to *anguis* is systematically mentioned and this comparison enlightens most of the medieval representations attached to this fish. This characteristic will thus have its place in the medieval scientific definition. However deciding which characteristics are striking is far from obvious. For instance, we are tempted to select elements which allow us to represent the crocodile such as we know it ("Animal vivant à la fois dans l'eau et sur la terre, au cuir jaune

[18] Buquet, T., "La girafe, belle inconnue des bibles médiévales. Camelopardalis : un animal philologique", *Anthopozoologica* 43 (2), 2008, pp. 47-68.

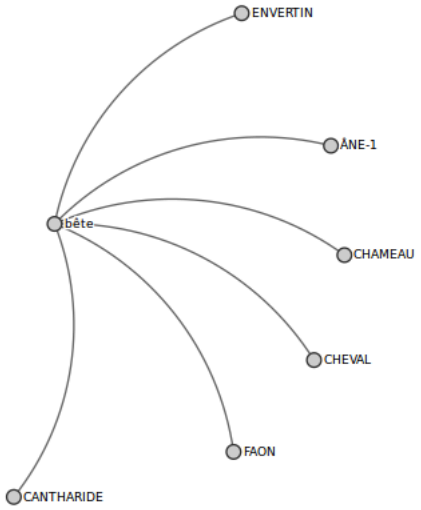
résistant, dangereux pour l'homme") and to relegate the properties ruled out of the definition to the encyclopedic note. Nevertheless, this solution is not satisfactory because the medieval point of view is falsified to a certain extent: in bestiaries and in the iconographic tradition, the fight against the hydra which devours it from within appears as a striking property of the crocodile. We can thus wonder if this characteristic, which may have been prominent for a medieval cleric, would not have its place in a zoological definition. And how might we justify the exclusion of the crocodile's tears or its opposing jaw which are so peculiar to this animal? It would be useless to look in these encyclopedias for an organization of the knowledge comparable to the one that will be proposed by naturalistic doctors of the XVIth century. Encyclopedists and medieval translators did not try to find the "marques" by which the scholars as Fuchs, Belon or Rondelet will structure their descriptions of plants and animals to allow their identification.[19] The medieval taxonomies are rather organized around prototypes from which the various species of the category are more or less distant. In order to give an exact definition, we need to know what is the best example of *poisson**, *ver** or *herbe**. The selection of the striking characteristics depends thus essentially on the appreciation of the editors of the *Dictionary*, that is on their representation of medieval culture.

3. The opportunities of Semantic Web

What advantage do they provide data Web resources to better represent and understand these taxonomic and/or lexical differences? As part of work on the definitions, the main advantage is in our opinion the opportunity to create links between the entries to build semantic networks. They can be established by different ways: create explicit links (through the synonym tag for instance); create implicit links based on the content of the entries (keywords in the definitions or in the encyclopedic notes); use the navigation history that shows the interests of users and target certain playback modes; or, finally, combine the previous strategies. The Créalscience already provides access to networks constructed from the contents of definitions and a graph of synonyms. It is then possible for someone who has no access to the vocabulary of the ancient language to navigate through the *Dictionary* and to find, for instance, kinds of tree without knowing their ancient name.

[19] Philippe Glardon, *L'Histoire naturelle au XVI^e siècle. Introduction, étude et édition critique de* La nature et diversité des poissons *de Pierre Belon (1555)*, Droz (Travaux d'Humanisme et Renaissance, n° CDLXXXIII), Genève, 2011, especially pp. 207-240. According to Philippe Glardon, the "marque" or "note" is a decisive cultural tool in the emergence of natural history in the XVIth century, "l'indispensable trait d'union entre le texte et la représentation graphique, l'élément discursif qui doit emporter la décision de l'identification, qui s'est resserrée autour de l'espèce réelle".

Besides the interest of this kind of representation for users, it definitely speeds up the work of editors: it helps to verify and coordinate formulations. The use of graphs allows for example to instantly measure the degree of polysemy of a term, to see what terms is assigned a hypernym, to cross-check in order to replace uniformly anachronistic terms and to ensure that the same generic term was used by all editors for specific terms relating to the same category.



In this case, the graph allowed us to notice a mistake: *cantharide* in a zoological sense was defined as a *ver**(which was right), but in a medical sense the generic name *bête** had been retained.

However any automatic processing applied to the definitions seems risky for the moment. Imposing such as hypernym *poisson** in all definitions which deals with an animal that lives in the water would be a contradiction. As we have said, the medieval *crocodile* is not only a *poisson** and other animals related to water as the *crapaud* clearly fall into the category of *vers**. At this stage of research, we cannot do without a trial and error linked to both the body and the lexicographical process itself.

Other applications of graphs can be considered for the future. As synonyms are tagged by the editors, their linking can provide a representation of onomastic fields. It will be interesting then to compare the networks of different related words: *arbre** and *herbe**, or on the other hand *poisson**, *bête**, *ver** and **coquille*. The addition of a "comparison" tag would integrate an important criterion of identification and classification in the field of natural sciences. Certain plants or animals are referred to as prototypes. For example, it could be interesting to study the place of the word *pomme* among the other fruits. Contrary to the generic Latin word *pomum*, the French word has a specific sense. But a comparison with the apple can still be used to describe another fruit[20]. A query with the word *serpent* would probably show a set of several related animals. Some associations could be unexpected because medieval sciences do not involve the same criteria as ours. Furthermore, the explanation of the relationship

[20] Lemon (*citron*) is described as a kind of apple ("une manière de pomme") for instance by Olivier de la Haye (*Poème sur la grande peste de 1348,* 1426, éd. H. Georg, p. 187).

with the Latin scientific lexicon could give interesting results. Indeed, a significant portion of corpus texts are translations of Latin works or take nomenclatures expressed at the same time in Latin. A specific tag under which one would notice the Latin equivalent would refine the data on lexical creation process.

Finally, the *Dictionary*, which already includes scientific senses identified in the AND (Anglo-Norman Dictionary), the DEAF (Dictionnaire Étymologique de l'Ancien Français) and DMF (Dictionnaire du Moyen Français), if it continues its expansion, should allow clear observation of semantic changes from the XIth to the XVth century. Concerning the confrontation with modern scientific terminology, it is already possible even if the operation is in its early trials. At more or less long term, it should be possible to compare the dictionary systematically to modern French nomenclature of specialized vocabularies. It is likely that some modern terms will not find resonance in the database, even including among the association criteria text notes. It will then be necessary to ask whether this absence is caused by a wrong description to correct in the database, a lexical creation that appeared after 1500 or an epistemological discrepancy that could then be explained in the note. The *Dictionary* would then allow queries from modern terms, by a non medievalist user curious about ancient representations and evolution of knowledge. It would thus give users the opportunity to question their own conceptual tools. There is much to be done before we reach a satisfactory model of scientific neologisms and semantic relationships between categorizing terms during medieval times, but we can assume it is worth meeting the challenge for linguists and historians of science.

Conclusion

Writing articles in the *Dictionary of Medieval Scientific French* raises permanent questions about the choice of generic names and the nature of definitions. Indeed, the lexicographic project leads us to forget the categories through which we think of nature and our reflexes of definition. The difficulty lies, on one hand, in the nature of the inherited knowledge, the fact that data are not selected yet and not organized into a hierarchy according to stable and homogeneous criteria from the XIIth to the XVth centuries, and on the other hand, in the semantic evolution of the lexicon. The meaning of generic words evolved at the same time as taxonomies, which entailed a lexical vacancy for certain concepts: there is no word in modern French for *coquilles** of Brunet Latin, for *poissons** in the medieval sense or for *herbes** as Jean Corbechon understands them. Furthermore, the contrast is great between the lexical freedom and the conceptual flexibility that the French corpus natural sciences lets us perceive and the necessary coherence of a dictionary, especially an electronic one. Far from the monosemic ambition of modern scientific language, medieval scientific works in French are characterized by a linguistic variety that their authors obviously did not try to reduce.[21] They allow us to see the evolution of thought or at least practices of trans-

[21] See, for instance, about the multiplicity of names for the same concept, Ducos, J., Salvador, X.L., « Pour un dictionnaire de français médiéval : le projet Crealscience », *Langages*, n° 183, septembre 2011, p. 65.

lation. Categorization is organized around a prototype, an exemplary species. The aim of the project is to show evolutions which are often not linear and which we cannot understand well without comparing them to the lexical uses of Latin and without placing it in a wide cultural context, by being careful to take into account modern knowledge which is the prism through which the user reads the *Dictionary*. To find the balance between allegiance to medieval scientific culture and coherence, homogeneity and legibility which are expected from a lexicographical database, the encyclopedic note offers a useful space to collect the essential philological, linguistic and scientific information to explain the relationship with modern nomenclature.

The use of a collaborative platform has already taken forward the project. We can hope that browsing by graphs, taking into account specific tags (such as keywords in definitions and notes, fields, synonyms, Latin equivalents or comparisons) will allow to improve the analysis of the birth of French scientific terminology and to connect this work with other web projects concerning the reconstruction of former scientific paradigms.