

The challenge of a **conceptualized** and **stratified** water lexicon : the **LEXEAU** project*

Mutual comprehension in the domain of water is a real challenge. A scientific language about water has been constituted by scientists. It is also used by journalists and engineers in a **para scientific** apprehension of the world.

This may cause **language distortion** and generate **misundertsanding** in other fields of human activity with a different water concern. It may cause **socio-economic damages** and nourish international conflicts.

The LEXEAU project proposes a scientific apprehension of this distortion in a « bottom-up » approach (fig. 1), based on the collective construction of an ontologic substrate.

*Communication of Jean-Louis Janin (PhD student at Bordeaux Montaigne University) presented in Amsterdam at the 1st European Ontology Network (EUON) Workshop on the 25/09/2014.

An **ontological** substrate of classes

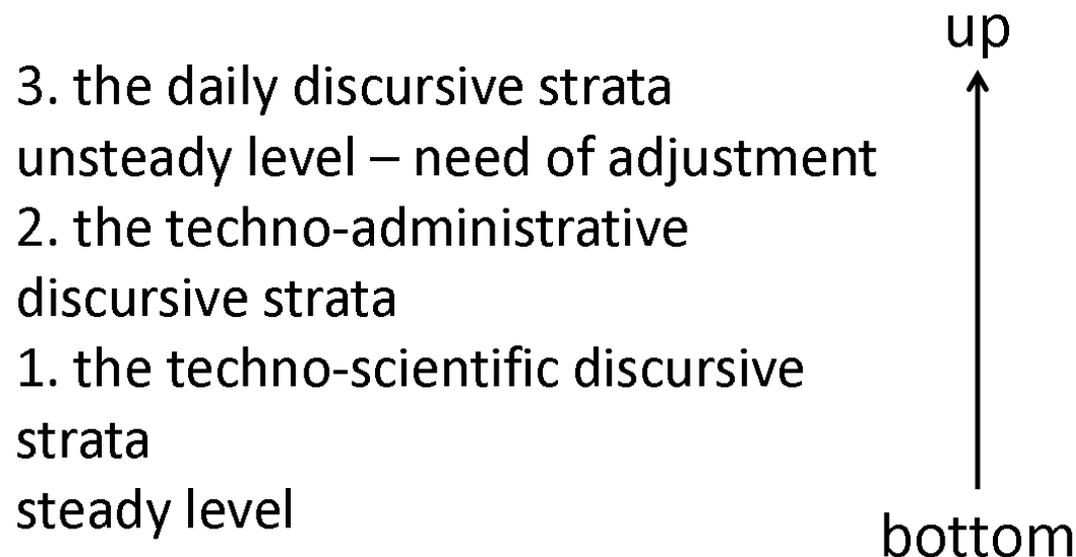
The ontology includes :

- a **denomination base** (DB) with a **class** called **Named entity**, and sub-classes of **persons**, **works**, **equipement**, **spatial entities** (water bodies ...), **dates**, **periods**, **substances**, **documents** of the **corpus** ...
- a **knowledge base** (KB) of water related **concepts** such as **administrative agreement**, **water flow**, **water movement**, **natural process**, **industrial process**, **quantization**, **water concern**, **water use** ...
- a **lexical base** (LB) of **lexemes** and **acceptations**, to define and illustrate the concepts of the DB and LB and treat other significant entries of the **corpus**.

A discursive stratification

The terms of a discourse about water are employed and understood in a given **discursive strata**, depending on the auditory, either **techno-scientific (TS)** (scientific books, articles and communications), **techno-administrative (TA)** (laws, regulations, authorisations, contracts, treatises) or used **daily**, in the **current discourse (CD)**, in the **press** and on **television**, which requires an **adjustment** process of the **CD**, as suggested by Henri Portine (fig. 1).

Fig. 1 The « bottom-up » adjustment process of the current dicourse

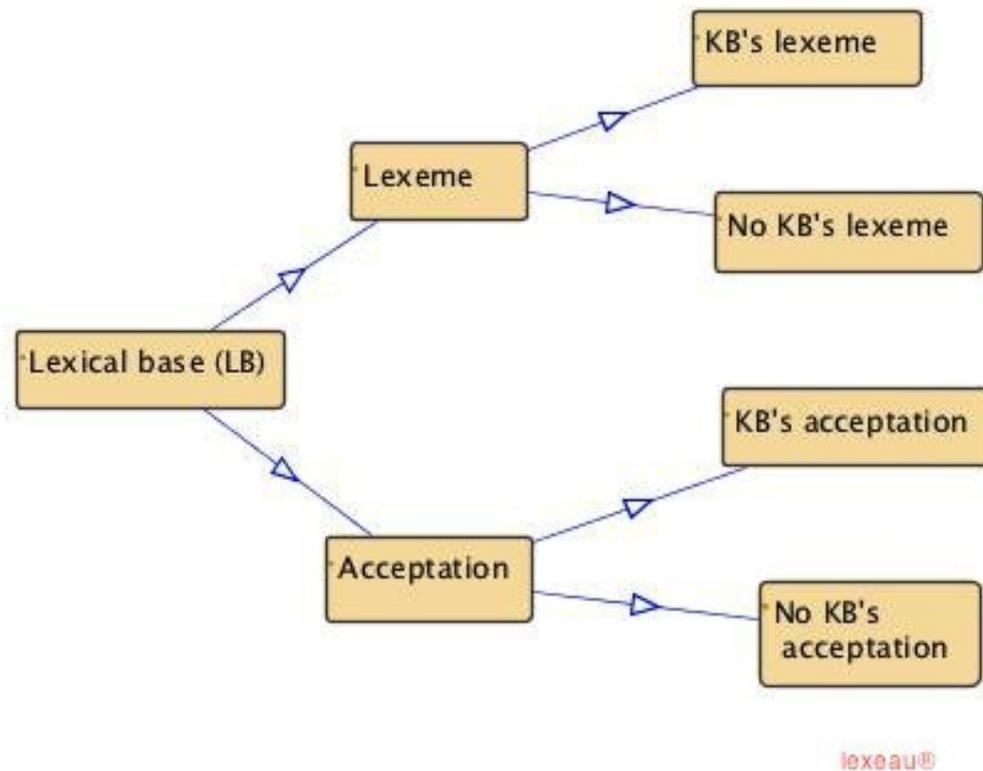


From DB/KB's **concepts** to **lexical acceptations**

- In the DB/KB's development process, **concepts** involved in **class names**, **slots** and **instances** will be defined and illustrated in the corresponding LB's **lexeme articles**.
- The **acceptations** will usually belong to the **techno-scientific** or **techno-administrative** discursive strata.
- With **current discourse** acceptations of such lexemes, the question to the linguist is : how does the locutor deals with this **porosity** between TS/TA and CD discursive strata, to prevent misunderstanding in the auditory ?
- What sort of **porosity** between TS and TA acceptations ?
- How to deal with **synonymy** in the lexicon ?
- This is to be analysed in the full text of the **documents**.

A lexical base of lexemes and acceptations derived from the DB and KB structure and content

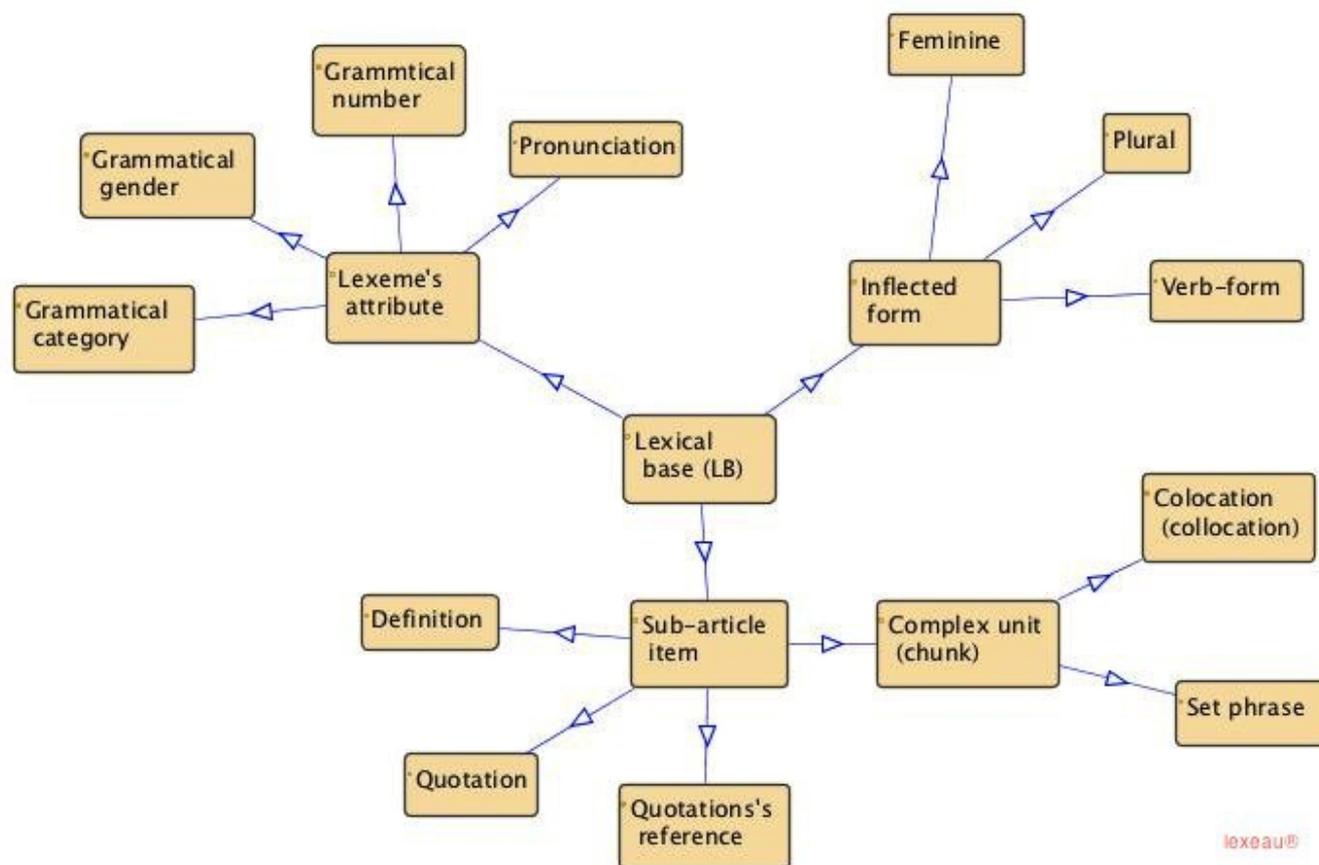
Fig. 2 Lexeme and Acceptation classes in the lexical base class tree



A **DB/KB's lexeme** can have a **No DB/KB's acceptance** (an acceptance not valid in the DB/KB's ontology), which differs from a **No DB/KB's acceptance** of a **No DB/KB's lexeme**.

A lexical base attested in the corpus

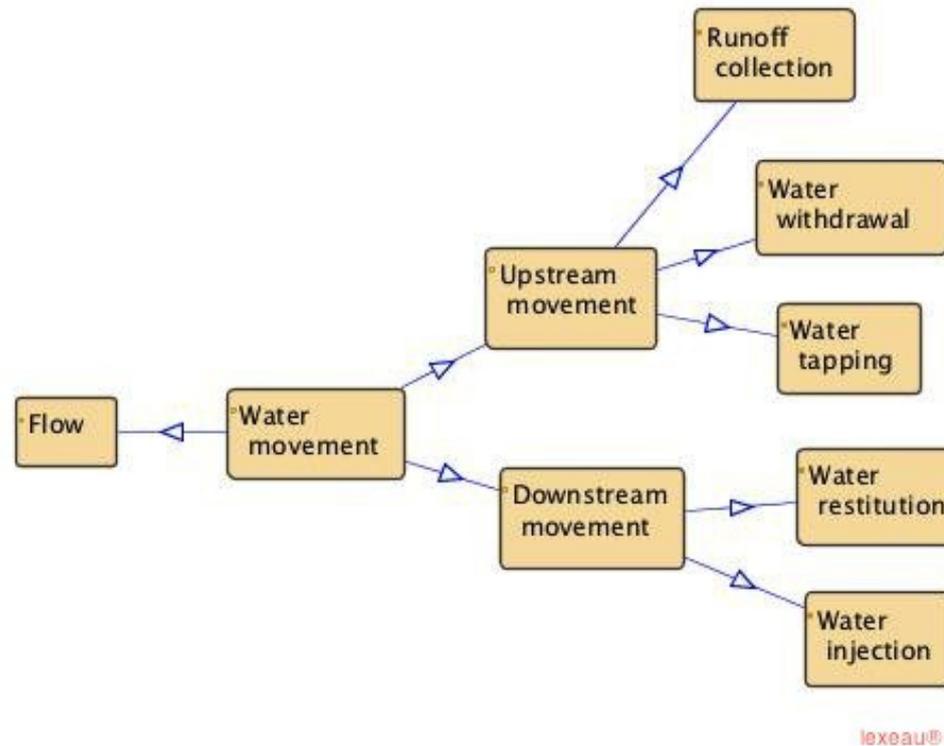
Fig. 3 Other classes in the lexical base class tree



The LB sub-article items are the definitions, quotations and chunks of the lexeme's acceptations attested in the corpus

An extract of the **knowledge base** : **modeling water movements**

Fig. 4 Class tree of the water movements and water flows in the knowledge base



The KB's model allows to register the flow data (maximum discharge, water volume, pollutant concentration ...) due to human dated and localised actions on water ressource.

Programm of work, contact

Since october 2013, the project is linked with a PhD thesis on lexical semantics prepared at Bordeaux Montaigne University.

After the presentation in Paris, at the invitation of *Académie de l'eau* and *Société hydrotechnique de France*, on the 07/07/2014, the project is to create a *Dictionnaire de l'Académie de l'eau*. The lexical treatment will take place step by step, on a succession of themes (water movements, industrial processes, chemical substances ...). The edition of the project ontology will change from Protégé 3.4.8 (Frame) to Protégé 4.3 (OWL). An internet mailing site is envisaged.

The French partners of the project are looking for academic and professional (native) English speaking institutions and contributors to set up a bilingual project consortium.

Contact :

- Jean-Louis Oliver, Paris, academie@oieau.fr
- Jean-Louis Janin, Bordeaux, turenejlj@orange.fr