

*L'analyse textuelle du Bateau Ivre de Rimbaud avec le logiciel Lexico3:  
ressources actuelles et possibilités de développement*

Séminaire de Littérature et Informatique

Giordano Righetti  
Università degli Studi di Bologna

## Introduction

Cette recherche examine les possibilités offertes par un logiciel informatique d'analyse textuelle appliqué à l'analyse d'un texte littéraire. Le logiciel en question est Lexico3, un outil lexicométrique présent dans le bouquet *O.l.f.a.c.t.*, un bouquet qui contient plusieurs logiciels pour le traitement et l'analyse de textes.

Tous ces outils, ainsi que Lexico3, sont projetés spécifiquement pour des corpus textuels assez longs: des romans, des collections des contes ou des histoires, des anthologies, etc. Lexico3 est en fait utilisé le plus souvent pour l'analyse de ces grands corpus textuels, et il est en ce contexte qu'il fournit les résultats plus précises et satisfaisants, en vue des quels il avait été développé.

Dans l'analyse d'un grand corpus littéraire, les applications dont dispose Lexico3 recherchent une expression souvent répétée, une forme récurrente, une concordance; ensuite le logiciel élabore tous ces données en relation avec leur majeure ou mineure occurrence dans le corpus et conséquemment formule des hypothèses concernant des caractères du corpus tels que la forme, le style et le contenu sur la base d'un numéro consistant d'occurrences, de façon que ces hypothèses puissent être affirmées et vérifiées avec un suffisant degré de fiabilité.

Le défi que se pose pour un outil pareil regarde son applicabilité à un type de texte différent des corpus de prose auxquels il est habitué: pour exemple, un texte poétique. Ce dernier est par nature un texte considérablement plus court et synthétique, et il prévoit des formes et des structures que lui sont propres. Premièrement, le style de la syntaxe du texte poétique présente des différences sensibles par rapport à celui de la prose (parataxe plutôt que hypotaxe); deuxièmement, la densité du sens exprimé est radicalement différente pour chaque expression et pour chaque partie d'une poésie ou d'un poème par rapport à des romans ou des contes.

La question est donc si un outil lexicométrique comme Lexico3 peut être utilisé pour analyser un texte poétique et comment et combien on peut appliquer ses fonctionnalités et applications dans la lecture et dans la réflexion autour d'un poème. Plus en général, est-ce qu'il est possible d'analyser un texte poétique par moyen d'un logiciel lexicométrique? Et comment conviendrait-il de organiser l'analyse, avec quelles prémisses et avec quels objectifs? Combien il faut s'attendre par les résultats de cette opération et quel degré de fiabilité elle peut garantir?

Pour répondre à ces interrogatives on a pris en considération un texte poétique assez long dans son genre: *Le Bateau Ivre* de Arthur Rimbaud. Il a été soumis à une analyse lexicométrique et formelle avec le logiciel Lexico3, précisément avec la nouvelle version de ce programme, présente dans le bouquet *O.l.f.a.c.t. 2003*. On envisage ainsi de réaliser une analyse relative à plusieurs variables

importantes pour ce texte poétique et de voir comment implémenter ces variables dans le logiciel informatique. A la fin d'une telle opération, on peut observer quels résultats on a obtenu: en particulier, si le programme est capable de confirmer avec ses dates quelques connaissances qu'on possédait déjà sur *Le Bateau Ivre*, mais surtout s'il est également capable de fournir des indications intéressantes et fiables pour découvrir des aspects nouveaux de ce texte poétique complexe.

### ***Le Bateau ivre : structure et style du texte***

Comme je descendais des Fleuves impassibles,  
Je ne me sentis plus guidé par les haleurs:  
Des Peaux-Rouges criards les avaient pris pour cibles,  
Les ayant cloués nus aux poteaux de couleurs.

J'étais insoucieux de tous les équipages,  
Porteur de blés flamands ou de cotons anglais.  
Quand avec mes haleurs ont fini ces tapages,  
Les Fleuves m'ont laissé descendre où je voulais.

Dans les clapotements furieux des marées,  
Moi, l'autre hiver, plus sourd que les cerveaux d'enfants,  
Je courus ! Et les Péninsules démarrées  
N'ont pas subi tohu-bohus plus triomphants.

La tempête a béni mes éveils maritimes.  
Plus léger qu'un bouchon j'ai dansé sur les flots  
Qu'on appelle rouleurs éternels de victimes,  
Dix nuits, sans regretter l'œil ni ais des falots !

Plus douce qu'aux enfants la chair des pommes sures,  
L'eau verte pénétra ma coque de sapin  
Et des taches de vins bleus et des vomissures  
Me lava, dispersant gouvernail et grappin.

Et dès lors, je me suis baigné dans le Poème  
De la Mer, infusé d'astres, et lactescent,  
Dévorant les azurs verts ; où, flottaison blême  
Et ravie, un noyé pensif parfois descend ;

Où, teignant tout à coup les bleuités, délires  
Et rythmes lents sous les rutillements du jour,  
Plus fortes que l'alcool, plus vastes que nos lyres,  
Fermentent les rousseurs amères de l'amour !

Je sais les cieux crevant en éclairs, et les trombes  
Et les ressacs et les courants : je sais le soir,  
L'Aube exaltée ainsi qu'un peuple de colombes,  
Et j'ai vu quelquefois ce que l'homme a cru voir !

J'ai vu le soleil bas, taché d'horreurs mystiques,  
Illuminant de longs figements violets,  
Pareils à des acteurs de drames très antiques  
Les flots roulant au loin leurs frissons de volets !

J'ai rêvé la nuit verte aux neiges éblouies,  
Baiser montant aux yeux des mers avec lenteurs,  
La circulation des sèves inouïes,  
Et l'éveil jaune et bleu des phosphores chanteurs !

J'ai suivi, des mois pleins, pareille aux vacheries  
Hystériques, la houle à l'assaut des récifs,  
Sans songer que les pieds lumineux des Maries  
Pussent forcer le mufle aux Océans poussifs !

J'ai heurté, savez-vous, d'incroyables Florides  
Mêlant aux fleurs des yeux de panthères à peaux  
D'hommes ! Des arcs-en-ciel tendus comme des brides  
Sous l'horizon des mers, à de glauques troupeaux !

J'ai vu fermenter les marais énormes, nasses  
Où pourrit dans les joncs tout un Léviathan !  
Des écroulements d'eaux au milieu des bonaces,  
Et des lointains vers les gouffres cataractant !

Glaciers, soleils d'argent, flots nacreux, cieux de braises !  
Échouages hideux au fond des golfes bruns  
Où les serpents géants dévorés des punaises

Choient, des arbres tordus, avec de noirs parfums !

J'aurais voulu montrer aux enfants ces dorades  
Du flot bleu, ces poissons d'or, ces poissons chantants.  
– Des écumes de fleurs ont bercé mes dérades  
Et d'ineffables vents m'ont ailé par instants.

Parfois, martyr lassé des pôles et des zones,  
La mer dont le sanglot faisait mon roulis doux  
Montait vers moi ses fleurs d'ombre aux ventouses jaunes  
Et je restais, ainsi qu'une femme à genoux...

Presque île, ballottant sur mes bords les querelles  
Et les fientes d'oiseaux clabaudeurs aux yeux blonds.  
Et je voguais, lorsqu'à travers mes liens frêles  
Des noyés descendaient dormir, à reculons !

Or moi, bateau perdu sous les cheveux des anses,  
Jeté par l'ouragan dans l'éther sans oiseau,  
Moi dont les Monitors et les voiliers des Hanses  
N'auraient pas repêché la carcasse ivre d'eau ;

Libre, fumant, monté de brumes violettes,  
Moi qui trouais le ciel rougeoyant comme un mur  
Qui porte, confiture exquise aux bons poètes,  
Des lichens de soleil et des morves d'azur ;

Qui courais, taché de lunules électriques,  
Planche folle, escorté des hippocampes noirs,  
Quand les juillets faisaient crouler à coups de triques  
Les cieus ultramarins aux ardents entonnoirs ;

Moi qui tremblais, sentant geindre à cinquante lieues  
Le rut des Béhémots et les Maelstroms épais,  
Fileur éternel des immobilités bleues,  
Je regrette l'Europe aux anciens parapets !

J'ai vu des archipels sidéraux ! et des îles

Dont les cieux délirants sont ouverts au vogueur :  
– Est-ce en ces nuits sans fonds que tu dors et t'exiles,  
Million d'oiseaux d'or, ô future Vigueur ?

Mais, vrai, j'ai trop pleuré ! Les Aubes sont navrantes.  
Toute lune est atroce et tout soleil amer :  
L'âcre amour m'a gonflé de torpeurs enivrantes.  
O que ma quille éclate ! O que j'aïlle à la mer !

Si je désire une eau d'Europe, c'est la flache  
Noire et froide où vers le crépuscule embaumé  
Un enfant accroupi plein de tristesse, lâche  
Un bateau frêle comme un papillon de mai.

Je ne puis plus, baigné de vos langueurs, ô lames,  
Enlever leur sillage aux porteurs de cotons,  
Ni traverser l'orgueil des drapeaux et des flammes,  
Ni nager sous les yeux horribles des pontons.

"Je dis qu'il faut être voyant, se faire voyant. Le poète se fait voyant par un long, immense et raisonné dérèglement de tous les sens". Rimbaud expose son programme poétique dans sa lettre du 15 mai 1871 à Paul Demeny.<sup>1</sup> *Le Bateau ivre*, écrit la même année, apparaît comme la transposition allégorique de ce programme.

Il est un poème de cent vers organisés en vingt cinq quatrains. Les cinq premiers constituent une première partie, relativement homogène, consacrée à l'évocation des événements ayant entraîné le naufrage du bateau. La suite contera l'odyssée de l'épave, une fois gagnée la haute mer. Les cinq premières strophes racontent comment un bateau rompt ses amarres : c'est le poète rompant avec les normes de la poésie, les conventions de la morale, l'idéologie dominante de la société. Les strophes 6 à 17 évoquent les aventures maritimes étourdissantes de l'épave à la dérive : c'est le poète arrivant "à l'inconnu". Enfin, les strophes 18 à 25 décrivent l'épuisement du narrateur et sa nostalgie du vieux monde : c'est le moment où, "affolé", le "voyant" doit se résigner à toucher terre, abandonner ses visions, avec la consolation de *les avoir vues*.

Si l'on en reste à la lettre de ce qui est dit, à l'histoire racontée, le début du *Bateau Ivre* est tout

---

<sup>1</sup> A. Rimbaud. *Oeuvres poétiques et lettres choisies, Dossier du professeur*, par Thierry Méranger, pages 24-28. Paris, Hachette, 1998

simplement le récit d'un naufrage : Les strophes 1 et 2 racontent dans quelles circonstances le bateau a été attaqué par des « peaux-rouges criards » qui ont tué ses « haleurs ». On comprend qu'il s'agit d'un bateau à fond plat, ceux qui étaient utilisés pour transporter des marchandises (« porteur de blés flamands et de cotons anglais ») sur les rivières (le mot « fleuves » est répété à deux reprises) et qui devaient être guidés par des cordages depuis les berges. Privé de ses « haleurs », le bateau est donc désormais à l'abandon, à la dérive, porté par le courant du fleuve vers la mer où la violence des vagues le fera couler. Ce choc avec l'océan est le sujet des strophes 3, 4,5. Les strophes 3 à 5 portent jusqu'à l'ivresse cette joie d'être libre. C'est l'ivresse de la liberté, pourrait-on dire pour reprendre le mot du titre. Ivresse au sens de joie intense, euphorie. La rencontre violente avec la mer est saluée comme une fête sauvage.

A une première lecture, ce poème impressionne pour sa richesse lexicale et sa complexité de signification: il contient beaucoup des termes suggestifs, beaucoup de formes répétées, beaucoup des univers sémantiques spécifiques. Un très riche vocabulaire "marin", formidable par rapport au fait que Rimbaud, que avait dix-sept ans quand il écrivait ce poème, n'avait jamais encore vu le mer: il écoutait avec attention les récits de ces amis que l'avaient vu et y faisait confiance.

On y remarque aussi plusieurs formes typiquement poétiques, telles que similitudes, métaphores, hyperboles, invocations, souvent répétés dans le texte. Ces figures rhétoriques enrichissent la description du voyage du bateau à travers les mers et les océans.

### **«Je»: le bateau**

Il faut continuer ensuite analysant rapidement la forme morpho-syntactique de l'oeuvre. En ce qui concerne les structures grammaticales et stylistiques, le poème se présente comme un récit à la première personne. Celle-ci peut se présenter sous diverses formes grammaticales: le pronom personnel sujet, la forme d'insistance « moi », le pronom complément « me », l'adjectif possessif singulier ou pluriel.

En général, dans les textes poétiques, la première personne désigne le poète lui-même. La particularité de ce poème vient de ce que la première personne ne peut pas renvoyer à l'auteur, sujet réel de l'énonciation, mais désigne un sujet fictif qui est le bateau du titre. En effet, certains groupes grammaticaux ne sauraient renvoyer à un sujet humain. L'expression "ma coque de sapin" (v.18) par exemple se rapporte nécessairement au bateau. Il en est de même de l'adjectif « porteur » (vers 6), construit en épithète détaché, apposé au pronom sujet « j' » (vers 5) : seul un bateau peut être dit «porteur de blés flamands et de cotons anglais». Ce n'est pas non plus l'auteur mais le pont du bateau

que l'océan "lava" de ses "tâches de vin bleu et des vomissures" (vers 19).

Cependant, dans la majorité des cas, les indices de première personne sont associés à des verbes désignant des actions humaines : « je courus ; j'ai dansé », à des verbes de sentiment ou de volonté : « je ne me sentis plus ; j'étais insoucieux ; je voulais », qui tendent à personnifier le bateau. Le lecteur est donc encouragé à considérer le bateau comme une représentation métaphorique du poète.

### **Lexico3 et l'analyse lexicométrique des corpus textuels**

Après cette analyse synthétique de la structure et des formes du poème, nécessaire prémisses à l'implémentation du texte de Rimbaud dans l'outil informatique, il est possible d'empoigner ce dernier en y découvrant les fonctions et les potentialités. Lexico3 est un outil lexicométrique pour le traitement et l'analyse des documents textuels: il implique donc des logiciels créés pour calculer le nombre d'occurrences de certaines formes et expressions et leur distribution dans un texte.

Toutefois, il faut préventivement préciser que ce logiciel, comme la totalité des ressources lexicométriques informatiques, est projeté pour l'analyse de textes longs, pour la plupart politiques: à savoir, des corpus textuels assez longs, avec beaucoup des formes répétées et avec une densité et complexité de sens relativement maigre. D'ailleurs, les outils lexicométriques ont été déjà appliqués depuis quelques années à des documents textuels d'autre genre: essai sociopolitiques et sociopolitiques, entretiens de télévision, et enfin, textes littéraires.

Parmi ces derniers, les programmeurs se sont concentrés au début sur des corpus assez larges: anthologies des contes, romans des plusieurs volumes, oeuvres complètes ou correspondances de certains écrivains. En tout ce cas, les potentialités et les résultats calculés normalement sur les textes politiques avaient donné des résultats intéressants et comparables; avec la mise à jour continue des programmes, donc, les textes choisis pour l'analyse informatique ont devenu peu à peu plus brefs, jusqu'au single roman ou récit.

Pour tous ces types de texte, les fonctions principales que Lexico3 permet de appliquer à un texte sont:

- le "dictionnaire": la liste des toutes les termes ordonnées par rapport au calcul du nombre des occurrences de chaque *forme* (=terme) dans le texte
- les "segments répétés": l'occurrence conjointe d'un certain terme avec des autres qui les précèdent ou les suivent (Ex. "j'ai vu", "les arbres")
- les principales caractéristiques lexicométriques de chaque section (numéro des occurrences; numéro des formes, numéro de *hapax*: un hapax est un segment présent que une fois; fréquence maximale, etc.)

- la "carte des sections" avec la fréquence de chaque forme ou segment dans chacune des sections qu'on a défini au début
- les "group de formes": cette importante fonction recherche la présence de partie de texte à l'intérieur de mots et expressions qui ne correspondent pas nécessairement au terme recherché (Ex. "oiseau+" trouvera toutes le fréquences singuliers et pluriels; "+ant" trouvera tous les participe présent, etc.)

Tous ces résultats sont visualisables sur la fenêtre de Lexico3 distribués par sections par rapport à la division en sections qu'un a opéré au début, au moment de baliser le texte pour l'adapter au logiciel.

Dans ce contexte, les potentialités d'analyse d'un texte littéraire sont moins envisageables que les difficultés annexes: dans cette perspective, donc, un défi ultérieur peut être représenté par le cas qu'on prend en considération ici, à savoir la prise en main d'un texte poétique. Ce type de texte est au dehors de ce qu'en générale prévoient les logiciels lexicométriques, et Lexico3 en particulier. Il s'agit d'un texte assez long dans son genre, mais très court par rapport à la longueur moyenne des textes soumis à l'analyse informatique. De plus, le sens du récit de voyage de Rimbaud est très sophistiqué, il sous-entend nombreux soustextes et des doubles lectures difficiles à toucher et comprendre complètement. Etant donnée cette complexité stylistique et cette "étrangeté" du texte par rapport aux modèles lexicométriques, il faut tout d'abord se poser la question de comment préparer-en terme informatique, "baliser"-le texte au mieux pour le traiter aisément avec Lexico3.

### **Préparation du *Bateau Ivre* pour le logiciel: division des sections et balisage**

Avec le concept de "balisage" d'un texte on se réfère à sa prédisposition pour l'analyse lexicométrique informatique: en pratique, il s'agit de la division des textes en sous-sections qui pourront être traitées et comparées par le logiciel et de l'introduction de quelques "balises" et marque qui permet de manipuler le document plus facilement. Les textes longs, tels que les romans, sont souvent partagés en sections par rapport à leur division en chapitre ; s'il s'agit de récit plus courts, par rapport au sous-chapitre, au aussi au paragraphe.

Le poème de Rimbaud qui cette recherche analyse n'a évidemment pas de chapitres ou des paragraphes: toutefois, comme il semble naturel s'agissant d'une poésie, on peut reprendre la division en strophe pour identifier les sections.<sup>2</sup> On obtient avec ce système 25 sections pour l'analyse.

---

<sup>2</sup> Même si pas tous les strophes contiennent un discours fini et relativement autonome: dans des cas une phrase ou un discours peuvent terminer dans la strophe successive

Un autre système de division cherche de répondre au problème de la brièveté du texte et du nombre conséquemment limité de chaque occurrence: il suffit de multiplier pour dix le texte. D'une telle façon, on aura un nombre dix fois plus grand pour chaque occurrence spécifique: les résultats finales seront proportionnellement les mêmes, mais avec un numéro plus consistante on se propose de stigmatiser plus efficacement des caractéristiques du texte.

Dans le cas du *Bateau ivre*, il est préférable de maintenir le même numéro des strophes, pour voir clairement combien on peut opérer et obtenir d'un poème long et complexe comme celui-ci. La forme du texte partagé en section et balisé avec les marques pour les majuscules résulte ainsi:

Sda=1871> <S03=0>

§ le \*bateau ivre

<S03=1>

§ comme je descendais des \*fleuves impassibles

je ne me sentis plus guidé par les haleurs:

des \*peaux-\*rouges criards les avaient pris pour cibles

les ayant cloués nus aux poteaux de couleurs.

<S03=2>

§ j'étais insoucieux de tous les équipages,

porteur de blés flamands ou de cotons anglais.

§ quand avec mes haleurs ont fini ces tapages

les \*fleuves m'ont laissé descendre où je voulais.

<S03=3>

§ dans les clapotements furieux des marées

moi l'autre hiver plus sourd que les cerveaux d'enfants,

je courus! et les \*péninsules démarrées

n'ont pas subi tohu-bohus plus triomphants.

<S03=4>

§ la tempête a béni mes éveils maritimes.

§ plus léger qu'un bouchon j'ai dansé sur les flots

qu'on appelle rouleurs éternels de victimes,

dix nuits, sans regretter l'oeil niais des falots!

<S03=5>

§ plus douce qu'aux enfants la chair des pommes sures,  
l'eau verte pénétra ma coque de sapin  
et des taches de vins bleus et des vomissures  
me lava, dispersant gouvernail et grappin.

<S03=6>

§ et dès lors, je me suis baigné dans le \*poème  
de la \*mer, infusé d'astres, et lactescent,  
dévorant les azurs verts; où, flottaison blême  
et ravie, un noyé pensif parfois descend;

<S03=7>

où, teignant tout à coup les bleuités, délires  
et rythmes lents sous les rutillements du jour,  
plus fortes que l'alcool, plus vastes que nos lyres  
fermentent les rousseurs amères de l'amour!

<S03=8>

§ je sais les cieux crevant en éclairs, et les trombes  
et les ressacs et les courants: je sais le soir,  
l'aube exaltée ainsi qu'un peuple de colombes,  
et j'ai vu quelquefois ce que l'homme a cru voir!

<S03=9>

§ j'ai vu le soleil bas, taché d'horreurs mystiques,  
illuminant de longs figements violets,  
pareils à des acteurs de drames très antiques  
les flots roulant au loin leurs frissons de volets!

<S03=10>

§ j'ai rêvé la nuit verte aux neiges éblouies  
baiser montant aux yeux des mers avec lenteurs,  
la circulation des sèves inouïes,  
et l'éveil jaune et bleu des phosphores chanteurs!

<S03=11>

§ j'ai suivi, des mois pleins, pareille aux vacheries  
hystériques, la houle à l'assaut des récifs,

sans songer que les pieds lumineux des \*maries  
pussent forcer le mufle aux Océans poussifs!

<S03=12>

§ j'ai heurté, savez-vous, d'incroyables \*florides  
mêlant aux fleurs des yeux de panthères à peaux  
d'hommes! Des arcs-en-ciel tendus comme des brides  
sous l'horizon des mers, à de glauques troupeaux!

<S03=13>

§ j'ai vu fermenter les marais énormes, nasses  
où pourrit dans les joncs tout un \*léviathan!  
§ des écroulements d'eaux au milieu des bonaces  
et les lointains vers les gouffres cataractant!

<S03=14>

§ glaciers, soleils d'argent, flots nacreux, cieux de braises!  
échouages hideux au fond des golfes bruns  
où les serpents géants dévorés des punaises  
choient, des arbres tordus, avec de noirs parfums!

<S03=15>

§ j'aurais voulu montrer aux enfants ces dorades  
du flot bleu, ces poissons d'or, ces poissons chantants.  
§ - des écumes de fleurs ont bercé mes dérades  
et d'ineffables vents m'ont ailé par instants.

<S03=16>

§ parfois, martyr lassé des pôles et des zones,  
la mer dont le sanglot faisait mon roulis doux  
montait vers moi ses fleurs d'ombre aux ventouses jaunes  
et je restais, ainsi qu'une femme à genoux....

<S03=17>

§ Presque île, ballottant sur mes bords les querelles  
et les fientes d'oiseaux clabaudeurs aux yeux blonds  
et je voguais, lorsqu'à travers mes liens frêles  
des noyés descendaient dormir, à reculons!

<S03=18>

§ or moi, bateau perdu sous les cheveux des anses,  
jeté par l'ouragan dans l'éther sans oiseau,  
moi dont les \*monitors et les voiliers des \*hanses  
n'auraient pas repêché la carcasse ivre d'eau;

<S03=19>

§ libre, fumant, monté de brumes violettes,  
moi qui trouais le ciel rougeoyant comme un mur  
qui porte, confiture exquise aux bons poètes,  
des lichens de soleil et des morves d'azur;

<S03=20>

qui courais, taché de lunules électriques,  
planche folle, escorté des hippocampes noirs,  
quand les juillots faisaient crouler à coups de triques  
les cieus ultramarins aux ardents entonnoirs;

<S03=21>

§ moi qui tremblais, sentant geindre à cinquante lieues  
le rut des \*béhénots et les \*maelstroms épais,  
fileur éternel des immobilités bleues,  
je regrette l'\*europe aux anciens parapets!

<S03=22>

§ j'ai vu des archipels sidéraux! et des îles  
dont les cieus délirants sont ouverts au vogueur:  
§ -est-ce en ces nuits sans fonds que tu dors et t'exiles,  
million d'oiseaux d'or, ô future \*vigueur?

<S03=23>

§ mais, vrai, j'ai trop pleuré! les \*aubes sont navrantes.  
toute lune est atroce et tout soleil amer:  
§ l'âcre amour m'a gonflé de torpeurs enivrantes.  
§ o que ma quille éclate! o que j'aille à la mer!

<S03=24>

§ si je désire une eau d'\*europe, c'est la flache

§ noire et froide où vers le crépuscule embaumé

§ un enfant accroupi plein de tristesses, lâche

§ un bateau frêle comme un papillon de mai.

<S03=25>

§ je ne puis plus, baigné de vos langueurs, ô lames,

enlever leur sillage aux porteurs de cotons,

ni traverser l'orgueil des drapeaux et des flammes,

ni nager sous les yeux horribles des pontons.

### Les fonctionnalités de Lexico3 appliquées au poème

Une fois qu'on a introduit le texte balisé on lance le calcul lexicométrique: cette fonction présentera les principales caractéristiques lexicométriques, au niveau général, du texte. Ensuite, on lance la recherche des occurrences, des segments répétés, des groups de formes, et on parcourt les résultats avec un oeil aux dates statistiques et l'autre au texte original.

Principales caractéristiques lexicométriques	
Nombre des occurrences	834
Nombre des formes	471
Fréquence maximale	38
Nombre des hapax	390

Ces premières dates statistiques nous confirment clairement et efficacement un des caractères principaux du poème, sa richesse lexicale: en fait, on le remarque du fait que sur un total des 834 occurrences des syntagmes, plus que la moitié sont des formes spécifiques différentes. De plus, les *hapax*, à savoir les mots que le poète utilise seulement une fois dans tout le poème, sont presque 400, un numéro énormément élevé sur un total de 800.

On remarque après qu'il y a des formes avec plus d'occurrence qui appartiennent au "paysage marin": "oiseau/aux", "mer", évidemment "bateau", "fleuve/fleuves", "ciel/cieux", "soleil", etc. Pour tous ces termes, il faut le répéter, les occurrences ne sont pas en grand nombre, mais même sans multiplier pour dix le texte on note leur présence plus nombreuse des autres. En particulier, ils sont les premiers termes en ordre d'occurrence qu'on trouve après naturellement les prépositions et les articles.

Corpus : Bateau ivre	Concordance de : mer, bateau, oiseau, ciel, soleil (tri ordre du texte)
----------------------	---

Parfois, martyr lassé des pôles et des zones, la **mer** dont le sanglot faisait mon roulis doux montait vers . § o que ma quille éclate ! o que j ' aille à la **mer** ! § si je désire une eau d ' \*europe , c ' est la oyés descendaient dormir , à reculons ! § or moi , **bateau** perdu sous les cheveux des anses , jeté par l ' n enfant accroupi plein de tristesses , lâche § un **bateau** frêle comme un papillon de mai . § je ne puis s anses , jeté par l ' ouragan dans l ' éther sans **oiseau** , moi dont les \*monitors et les voiliers des \*hanses

ant sur mes bords les querelles et les fientes d ' **oiseaux** clabaudeurs aux yeux blonds et je voguais , sans fonds que tu dors et t ' exiles , million d ' **oiseaux** d ' or , ô future \*vigreur ? § mais , vrai , j ' ai de panthères à peaux d ' hommes ! Des arcs - en - **ciel** tendus comme des brides sous l ' horizon des mers

t , monté de brumes violettes , moi qui trouais le **ciel** rougeoyant comme un mur qui porte , les rousseurs amères de l ' amour ! § je sais les **cieux** crevant en éclairs , et les trombes et les ressacs § glaciers , soleils d ' argent , flots nacreux , **cieux** de braises ! échouages hideux au fond des golfes juillets faisaient crouler à coups de triques les **cieux** ultramarins aux ardents entonnoirs ; § moi qui i vu des archipels sidéraux ! et des îles dont les **cieux** délirants sont ouverts au vogueur : § - est - ce en efois ce que l ' homme a cru voir ! § j ' ai vu le **soleil** bas , taché d ' horreurs mystiques , illuminant de confiture exquise aux bons poètes , des lichens de **soleil** et des morves d ' azur ; qui courais , taché de bes sont navrantes . toute lune est atroce et tout **soleil** amer : § l ' âcre amour m ' a gonflé de enivrantes tains vers les gouffres cataractant ! § glaciers , **soleils** d ' argent , flots nacreux , cieux de braises !

On a recherché des termes qui puissent appartenir au contexte "marin" et de la nature, et on en a considéré la concordance, à savoir la position dans le texte ensemble avec d'autres éléments grammaticaux. Dans ce cas, on peut observer que presque tous ces termes, qui renvoient à des éléments du paysage marin, sont précédés ou suivis par un attribut. Ainsi les bateau (perdu, frêle), les oiseaux (clabaudeurs, d'or), les cieux (rougeoyant, crevant), le soleil (bas, amer, d'argent). Il est donc prévalente la dénotation, le caractère descriptif du récit: ces composants du paysage riche d'adjectifs favorisent le niveau descriptif et visuel du texte, plus que celui narratif. Elle est une prérogative typique des textes poétiques, plus que de la prose, mais en ce texte elle est particulièrement accentuée.

Une autre forme qui résulte parmi les plus fréquents est "enfant/enfants": ça peut aussi être un signal significatif si on le met en relation au thème de l'enfance et des enfants pour Rimbaud, un thème central

dans son oeuvre poétique autant plus qu'il écrit ses poèmes dans le coeur de son adolescence (on a déjà souligné qu'au moment du *Bateau ivre* il avait dix-sept ans).

Après on pourrait calculer la présence de certaines formes syntactiques particulières: pour exemple l'occurrence des participes présents, que dans la première lecture du texte ressemblaient très significatifs. On peut utiliser là une fonction plus sophistiquée de Lexico3, qui recherche le "Group de formes". On peut taper une série de lettre incomplète, et indiquer si elle est

- "exactement ce que je recherche"
- "le début de ce que je recherche"
- "la fin de ce que je recherche"

et lancer le calcul. Pour trouver le participes, ça suffit donc de taper dans la masque de recherche le group de formes "ant" avec la troisième option ("la fin de ce que je recherche"): "+ant". On obtient ainsi la longue liste des participes.

ayant  
ballottant  
cataractant  
crevant  
dévorant  
dispersant  
enfant  
fumant  
illuminant  
mêlant  
montant  
rougeoyant  
roulant  
sentant  
teignant

En observant cette liste, on y trouve surtout des participes qui se réfèrent au mouvement et au voyage, plus précisément au voyage marin du bateau/poète: il y a ici la dimension du voyant, dont Rimbaud avait fait le propos programmatique de sa poésie. La liste des participes "voyants"("ballottant","cataractant","dispersant","montant","roulant", etc.) qu'on a obtenu peut donc nous transmettre fortement la continuité et l'ivresse du voyage de ce poème.

Un autre élément qui peut être recherché est la présence de déictiques, à savoir de syntagmes qui se réfèrent et s'adressent à la première, deuxième ou troisième personne. On peut organiser cette recherche par moyen d'un groupe de formes pour la première personne ("je"+"j"+"moi"+"me"+"m"), un pour la deuxième, etc. On obtient le classement suivant:

je	12
j	12
moi	6
me	3
m	3
t	1
tu	1

Il paraît assez claire que ce poème exploite en grand mesure la première personne, comme on avait vu déjà dans la première lecture: ce calcul montre en plus qu'il n'y a que deux occurrences de la deuxième personne, et aucune occurrence de "il", "elle", "lui". Cette particulière distribution des déictiques nous communique à quel point le poète est seul sur le bateau, ou, plus littéralement, le bateau voyage seul sur la mer devant la nature qu'on a découverte si riche. En plus, si on approfondit l'analyse avec les ventilations des déictiques par sections, on voit une intensification de la présence du "je" poétique dans les dernières strophes (18-25), où on avait localisé, dans la schématisation préliminaire, l'épuisement du voyage et une certaine nostalgie du vieux monde. Le poète continue donc à être le seul protagoniste du voyage jusqu'à la fin de celui-ci, où sa présence devient plus marquée.

### **Les possibles développements du logiciel pour le traitement des textes poétiques**

A la fin de ces opérations on peut en tirer de conclusion mais aussi, surtout, envisager les possibilités d'implémentation de nouvelles fonctionnalités qui enrichissent Lexico3 et qui soient focalisées sur les textes poétiques. Dans ce sens, on pourrait penser de définir de façon générale mais inclusive des

champs sémantiques récurrents concernant les thématiques et les éléments du langage poétique.

En ce poème, pour exemple, on a mis en évidence l'émergence de l'univers sémantique du mer et de la nature: on en a trouvé les composantes en ajoutant l'un à l'autre des termes que normalement appartiennent à ce domaine (tels que "mer", "fleuve", "bateau", "ciel", etc.). Mais pour dépasser cette opération mécanique il sera probablement souhaitable de créer des champs sémantiques déjà "confectionnés"(au moins les plus fréquents) et y soumettre le texte; évidemment avec ce système on obtiendra aussi des termes et expressions que, bien que normalement appartiennent au champ défini, dans le texte qu'on analyse seront utilisés avec un autre sens ou valeur. Mais il suffit de contrôler la liste compréhensive des résultats des occurrences du champ sémantique choisi et éliminer aisément (avec le commande "Supprimer" de la fonction "Groupes de formes") les occurrences qui font exception au champ sémantique. Ainsi, on peut pour exemple définir le champ sémantique "univers marin" ou "éléments de la nature", qui comprendront la mer, les fleuves, les bateau, les marinières, etc. ou le ciel, le soleil, etc. et lancer la recherche directement sur tous ces termes, au singulier et au pluriel: une fois obtenus beaucoup des résultats, il sera facile, en contrôlant les concordances, éliminer ceux qui ne répondent pas au cas recherché. Dans le langage poétique, en fait, il y a des univers sémantiques et des vocabulaires spécifiques beaucoup plus fréquents et limitables que dans la prose, ce qui facilite le développement d'une telle fonctionnalité et en maximiser l'utilité.

Deuxièmement, on peut penser de appliquer la même opération, après le lexique, aussi aux formes typiques du langage poétique. Il y a des figures rhétoriques et des formes stylistiques qui sont propre de la poésie, en ce qu'elles sont beaucoup plus fréquentes que dans la prose mais surtout qu'y assument une fonction et une signification beaucoup plus dense. C'est le cas des similitudes, des métaphores, des comparaisons, des invocations. Pour certaines de ces formes on ne peut pas envisager des logiciels capables de les reconnaître: une métaphore est impossible à définir (jusqu'à aujourd'hui) en terme informatiques, soit lexicographiques soit statistiques. Toutefois, pour d'autres figures on peut introduire quelques artifices qui commencent à aider l'analyse lexicographique à les localiser dans les poésies. Dans le cas de similitudes, même si les possibilités poétiques sont infinies, il y a des formes récurrentes. Ce poème est exemplaire: si on lance la recherche des similitudes et comparaisons avec quelques mots typiques on obtient une liste assez riche:

§ j ' ai suivi , des mois pleins , **pareille** aux vacheries hystériques , la houle à  
illuminant de longs figements violets , **pareils** à des acteurs de drames très - antiques

nts : je sais le soir , l ' aube exaltée **ainsi** qu ' un peuple de colombes , et j ' ai vu  
bre aux ventouses jaunes et je restais , **ainsi** qu ' une femme à genoux . . . . § Presque

des \*fleuves impassibles je ne me sentis **plus** guidé par les haleurs : des \*peaux - \*rouges  
s furieux des marées moi l ' autre hiver **plus** sourd que les cerveaux d ' enfants , je  
démarrées n ' ont pas subi tohu - bohus **plus** triomphants . § la tempête a béni mes éveils  
tempête a béni mes éveils maritimes . § **plus** léger qu ' un bouchon j ' ai dansé sur les  
regretter l ' oeil niais des falots ! § **plus** douce qu ' aux enfants la chair des pommes  
mes lents sous les rutillements du jour , **plus** fortes que l ' alcool , plus vastes que  
s du jour , plus fortes que l ' alcool , **plus** vastes que nos lyres fermentent les rousseurs  
comme un papillon de mai . § je ne puis **plus** , baigné de vos langueurs , ô lames , enlever

§ le \*bateau ivre § **comme** je descendais des \*fleuves impassibles je  
d ' hommes ! Des arcs - en - ciel tendus **comme** des brides sous l ' horizon des mers , à  
tes , moi qui trouais le ciel rougeoyant **comme** un mur qui porte , confiture exquise aux  
de tristesses , lâche § un bateau frêle **comme** un papillon de mai . § je ne puis plus ,

Il est clair, comme on a vu pour les termes lexicaux d'un certain champ sémantique défini, qu'aussi dans ce cas il faut contrôler les résultats et supprimer les mauvaises concordances (celles que ne représentent pas une similitude ou comparaison). Il y a des termes qu'on peut considérer souvent comme typiques des comparaisons en poésie ("pareil/ils que"; "plus que"; "ainsi que"): on peut en définir une liste générale des plus communs dans les similitudes poétiques, et lancer une recherche avec eux tous au même moment. Après on contrôlera les résultats, mais quand même on voit déjà ici que même un terme complexe tels que "comme", qui très souvent ne se réfère pas à une similitude, donne ici seulement un cas sur quatre qui ne correspond pas; et aussi dans le cas de "plus", les occurrences que n'expriment pas une comparaison sont deux sur un total de huit.

Il faut préciser que *Le bateau ivre* avec son langage créateur d'images offre une occasion idéale pour cette opération, mais il reste en tout cas possible en générale d'envisager une possible définition de certaines formes stylistiques typiques du langage poétique et y inclure les termes que plus fréquemment les expriment: ainsi on peut lancer des recherches plus compréhensives des figures poétiques au coût, acceptable, du contrôle des résultats.

Cette opération de formalisation d'une structure et d'inclusion de ses formes plus communes reste une

fonctionnalité approximative mais peut se révéler très utile dans la pratique: et il est facile d'imaginer comment et combien on peut éteindre ces discours à d'autres formes poétiques, pour exemple les invocations (recherche de "O" + ..., toujours prévoyant un contrôle), en développant peu à peu les outils lexicométriques dans l'espoir qu'ils arrivent à maîtriser des textes toujours plus courts et plus denses de signification.

### **Conclusions**

Cette recherche a prise en considération le poème de Rimbaud *Le Bateau ivre*, choisi en raison de sa longueur et de ses caractéristiques formelles, pour le soumettre à une analyse textuelle réalisée par l'intermédiaire du logiciel informatique Lexico3, logiciel lexicométrique d'analyse textuelle. Cet outil, comme la grande majorité des logiciels de son genre, est projeté pour fournir des résultats intéressants et fiables sur des grands corpus textuels, pour la plupart textes politiques.

La poésie n'est pas donc son domaine spécifique originaire ou idéale: proprement d'ici dérive le défi de le tester avec un objet nouveau, d'un coté pour voir ses potentialités réelles, de l'autre pour en dériver des indications pour le développement des nouvelles fonctionnalités projetées pour les textes poétiques. Une fois appliqué à une poésie suffisamment étendue comme *Le bateau ivre*, Lexico3 peut être quand même utile: d'une coté pour confirmer avec des dates statistiques et des fréquences lexicographiques les impressions et les hypothèses qu'on formule lors de la lecture et de la première analyse littéraire du texte. D'autre coté, il est précieux en ce qu'il met en évidence les formes stylistiques et syntactiques récurrentes du poème.

On a vu comment on peut reconnaître des formes répétées typiquement poétiques dans le texte de Rimbaud, et on a proposé des possibles systèmes, qui restent encore des artifices, pour les élaborer à l'intérieur des catégories plus compréhensives. Ces catégories ne sont pas propres seulement du poème de Rimbaud mais, avec les nécessaires modifications, peuvent être déduites et abstraites pour une première formalisation. Cette formalisation des structures typiques du langage poétiques qui peut être élaborée par les outils informatiques reste pour l'instant très approximative; mais avec quelques artifices et quelques spécifications pourra constituer une pointe de départ pour le développement futur de logiciels qui analysent les poésies.

## Bibliographie

- É. Noulet. *Le Bateau ivre*, dans *Le Premier visage de Rimbaud*, pages 81-98, Bruxelles, Palais des Académies, 1953.
- H. de Bouillane de Lacoste. *Recherches sur les sources du Bateau ivre et de quelques autres poèmes de Rimbaud*, dans *Mercure de France*. Paris, 15 août 1935.
- M.-P. Berranger. *Le Bateau ivre*, dans *12 poèmes de Rimbaud analysés et commentés*, pages 112-128. Paris, Marabout, 1993.
- T. Méranger. *Le Bateau ivre*, dans *Rimbaud, Oeuvres poétiques et lettres choisies*, pages 24-28, Paris, Hachette, 1998.
- P. Fiala, J.M. LeBlanc. *Lexico 3. Manuels d'utilisations*. Paris, Ceditec, 2002.