



**HAL**  
open science

## Projet Anamètre : présentation, limites et avancées

Eliane Delente, Richard Renault

### ► To cite this version:

Eliane Delente, Richard Renault. Projet Anamètre : présentation, limites et avancées. International Conference - Plotting poetry : on mechanically-enhanced reading, Oct 2017, Bâle, Suisse. pp.73-92. hal-04072815

**HAL Id: hal-04072815**

**<https://hal.science/hal-04072815v1>**

Submitted on 28 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Projet *Anamètre* : présentation, limites et avancées<sup>1</sup>

Éliane Delente, Richard Renault

CRISCO (EA 4255), Université de Caen Normandie

## ABSTRACT

*Our purpose is to present the project Anamètre, whose object is the automated processing of French metrical poetry and drama. The corpus extends from the early seventeenth to the early twentieth century and comprises more than 700,000 lines of verse.*

*First, we distinguish rhythm as it is felt by the reader and as it is analysed automatically. Then, we describe the different steps of the automated processing. Tools are able to characterize all metrical regularities: metre, rhymes and stanzas. Any pattern, provided it is regular, even if not traditional, is recognised as metrical. Thus, stereotypical metres as 6+6 (alexandrin) or 4+6 (décasyllabe) as well as less frequent forms as 5+5 or some new forms that appeared around 1860 such as 4+4+4, 8+4 and 4+8 are all properly processed. Besides, rhymes are automatically associated and assessed, and stanza patterns satisfyingly identified. These metrical regularities are analysed both in prescribed forms (sonnets, terza rima...) and in periodical sequences.*

*In the second part, we consider the limits of this processing. Unsurprisingly, these mainly concern the end of the nineteenth century, when traditional constraints become increasingly flexible and new meters appear; and some verses are still discussed among metrists.*

*In the last part, we highlight some advances. The “metricometrie” method allows us to take into account historical evolution without specific dating. Moreover, we develop a way of thinking the relationship between rhythm and meaning, an outstanding feature in French poetry.*

*Key-words: French metrics, automated processing, metre, rhythm, metric repertory, poetic corpus, database.*

---

1. Les parties de « Le formatage XML du corpus » jusqu'à la fin de « Les limites du traitement automatique » ont été rédigées par R. Renault. Les autres parties l'ont été par E. Delente. La conclusion a été rédigée en commun.

## INTRODUCTION

Le traitement automatique appliqué à la poésie versifiée peut encore parfois être vu comme un balayage aveugle de la machine qui ne saurait rendre compte de certains aspects difficiles à appréhender de la création poétique. C'est une position qui mérite pourtant d'être nuancée. Car les propriétés métriques mises en évidence par le traitement automatique sont une part du discours rythmé. Le discours poétique versifié se présente toujours comme déjà rythmé et les rythmes réguliers qui le caractérisent constituent un aspect fondamental de sa structure.

Les textes versifiés représentent un matériau idéal pour le traitement automatique puisque les propriétés métriques qui les structurent intégralement sont généralement périodiques et systématiques. Un tel traitement, appliqué à de larges corpus, peut produire des analyses métriques utiles aux chercheurs, linguistes ou littéraires, en assurant une assise fiable et indispensable à des approches sémantiques, pragmatiques, statistiques ou stylistiques de la poésie.

La première partie de cet article tente de distinguer deux façons d'appréhender les régularités rythmiques, l'activité de lecture et le traitement automatique. Dans la deuxième partie, nous présenterons le projet *Anamètre* qui a pour objet le traitement automatique des formes métriques dans la poésie littéraire française et dans le théâtre classique. Dans une troisième partie, nous exposerons les limites du projet lui-même et les avancées seront envisagées dans une dernière partie.

## TRAITEMENT AUTOMATIQUE DU RYTHME RÉGULIER

Il convient de mentionner le travail pionnier de J. Roubaud qui a participé à l'élaboration de l'outil *Dynastie*<sup>1</sup>, premier programme français d'analyse du vers, puis, dans une perspective semblable, le *Métromètre*<sup>2</sup> de Valérie Beaudouin. Ces premiers travaux font place aujourd'hui à de nouvelles approches<sup>3</sup>. Mais une problématique commune à toute entreprise de traitement automatique de la métrique mérite d'être soulignée, d'autant qu'elle est rarement abordée. Que produisent exactement ces outils ?

Une manière assez exacte de répondre à cette question serait de dire que de tels outils fournissent, à des degrés de pertinence, de précision et d'ampleur variables, des analyses automatiques des propriétés caractéristiques des textes versifiés, pertinentes pour rendre compte du rythme régulier dans les textes versifiés.

- 
1. Roubaud (1986) et (1988).
  2. Beaudouin & Yvon (1994) et Beaudouin (2002).
  3. En dépit d'affirmations imprudentes telles que « [...], l'automatisation complète d'une chaîne de traitement qui va du texte brut à son analyse rythmique, proposée par Valérie Beaudouin sous le nom de *métromètre*, n'a pas eu de successeur depuis près de 15 ans », M. Bernard (2017, 137). Concernant ces nouvelles approches, voir *Langage*, n° 199 (2015).

Ces analyses automatiques produisent ainsi des données à interpréter et à l'occasion — c'est d'ailleurs le cas le plus fréquent — elles peuvent proposer une interprétation métrique des vers et des strophes. Une telle interprétation est relativement aisée à condition de disposer d'un corpus homogène, reflétant un système métrique stable. Appliqués à des périodes charnières comme la période pré-classique<sup>1</sup> ou la fin du XIX<sup>e</sup> siècle, il s'avère beaucoup plus délicat pour ces outils de fournir une interprétation exhaustive et sûre. Il suffit de garder à l'esprit que l'interprétation métrique automatique des vers et des strophes, quand elle est proposée 1) ne constitue jamais qu'une analyse partielle des régularités métriques d'un poème, 2) que l'analyse des propriétés métriques d'un poème a peu à voir avec les rythmes mêmes du poème tels qu'ils peuvent être perçus par un lecteur.

La plupart du temps, les propriétés intrinsèques au vers sont insuffisantes pour permettre une lecture métrique du vers. Soit ce vers en contexte 6-6 :

(1) On sort du mariage habitueux. C'est comme

(Verlaine, *Jadis et Naguère*, « L'Impénitence finale »)

Le nom « mariage » doit-il être traité avec diérèse ? Il semble que oui dans tous les cas. Le peu usité « habitueux » sera-t-il traité comme disjonctif ? Cela implique que le *e* féminin, post-tonique de « mariage », sera sélectionné par le lecteur, que « habitueux » sera traité sans diérèse et que ce vers sera considéré comme un vers non classique, avec une féminine récupérée par la seconde sous-mesure. Aux hémistiches {*On sort du mariage*} {*habitueux. C'est comme*} le lecteur associe les séquences de voyelles suivantes [ō ɔ y a i a] [ə a i œ ε ɔ].

Un tout autre traitement est envisageable : si « habitueux » est traité comme un mot jonctif, sur le modèle de « habituel », alors le *e* féminin de « mariage » n'est pas sélectionné et « habitueux » est traité avec diérèse. Aux hémistiches {*On sort du mariage*} {*habitueux. C'est comme*} le lecteur associe les séquences de voyelles suivantes [ō ɔ y a i a] [a i y œ ε ɔ]. Pourvu que le contexte de ce vers soit traitable mentalement de façon similaire, le lecteur retiendra probablement cette interprétation car c'est celle qui lui permet de sentir ce vers comme rythmiquement semblable à son contexte, c'est-à-dire de le sentir comme métrique. C'est ensuite l'affaire de l'analyste, aidé de ses doigts ou d'outils automatiques, de le caractériser comme un 6+6 classique.

De tels cas, loin d'être exceptionnels, représentent plutôt les conditions ordinaires d'une lecture rythmique pour les périodes de non-stabilité d'un système métrique donné et touchent aussi bien les formes de vers que les schémas strophiques. Si bien que lire métriquement un texte versifié, c'est parfois se livrer à un véritable calcul interprétatif fondé sur de nombreux paramètres.

---

1. Marot en est un bon exemple, voir Hervé (2011).

L'interprétation métrique à laquelle se livre le lecteur met en jeu des paramètres aussi divers que la langue des vers (le statut rythmique distinct des *e* masculins et des *e* féminins, l'hiatus, l'éllision métrique, les diérèses, les conventions phono-graphiques de la rime, etc.), le type de vers (simple ou composé), l'époque, le contexte (les vers voisins dans le poème), la familiarité du lecteur avec la tradition, avec l'auteur, avec le poème, la linéarité de la langue, enfin la nature temporelle et dynamique tant du développement linguistique (l'énoncé) que du développement des expressions métriques (vers, modules de strophes et strophes).

L'analyse automatique a donc peu à voir avec l'activité du lecteur pour qui lire métriquement un texte versifié constitue un événement consistant à traiter les caractéristiques rythmiques du texte de telle manière que ce traitement lui permette de sentir des impressions rythmiques semblables en dépit des différences linguistiques qui caractérisent les expressions métriques. Il ne s'agit nullement ici de déplorer que la machine ne simule pas la perception des régularités métriques. Il s'agit, plus crucialement, de soutenir que les formes métriques n'existent qu'en tant qu'elles sont perçues, reconnues, construites par le lecteur, autrement dit, que sentir des impressions rythmiques semblables constitue une expérience mentale, et qu'en dehors de cette expérience, on parle d'autre chose que du rythme. Mais la clarification de ce point ne discrédite nullement les résultats de la machine, car le repérage de propriétés est un aspect constitutif du rythme.

#### LE PROJET ANAMÈTRE

Le projet *Anamètre* a pour objet le traitement automatique des formes métriques dans la poésie littéraire française et dans le théâtre classique du début du XVII<sup>e</sup> au début du XX<sup>e</sup> siècle. Il comporte deux volets : la constitution d'une base de données de textes poétiques et théâtraux annotés pour la période considérée et l'élaboration d'outils d'analyse métrique automatique<sup>1</sup>.

À ce jour, le corpus XML se compose de 12670 poèmes et 102 pièces de théâtre couvrant ainsi 87 auteurs et totalise 704 385 vers annotés en XML (TEI).

Les programmes testés sur l'ensemble du corpus sont aptes à déterminer le mètre des vers (mètres simples et mètres composés) ainsi qu'à identifier les rimes (contraintes phono-graphiques), les éléments excédant la rime et ce qui la caractérise. Ils déterminent également les schémas strophiques (suites périodiques de strophes et formes fixes), la forme globale du poème et enfin repèrent les signes de ponctuation.

Les textes soumis à nos programmes, ayant préalablement été reconnus comme poèmes, font donc toujours l'objet d'une sélection humaine. Cette intervention en amont n'est toutefois pas vue comme une faiblesse des programmes.

---

1. Au projet *Anamètre* est associé le site web *Métrie en ligne* : <http://www.crisco.unicaen.fr/Verlaine/>

En effet, le lecteur lui-même, selon qu'il s'apprête à lire un texte de prose ou un poème versifié, met en œuvre un mode de lecture spécifique ajusté à l'objet. Nous retenons comme métrique tout poème qui présente soit une forme fixe, soit une périodicité métrique des vers et/ou des schémas de strophe. Tout texte ne répondant pas à l'une de ces conditions est écarté.

### LE FORMATAGE XML DU CORPUS

Le texte fourni en entrée du traitement est un fichier au format XML-TEI conforme à un schéma de validation défini pour l'ensemble du corpus *Malherbe*<sup>1</sup> du projet *Anamètre*. Le formatage initial des textes versifiés ne contient que deux éléments XML-TEI qui relèvent d'un balisage sémantique : le poème et le vers. Poème et vers sont donc les deux unités métriques introduites préalablement au traitement automatique. L'analyse automatisée n'a pas pour objet d'identifier une portion de texte comme étant un poème, ni de reconnaître un vers dans un flux de mots. Ce formatage initial se fait très facilement au regard de la disposition typographique du texte, mais l'analyse humaine est cependant mise à contribution lorsque le découpage en poèmes en tant que pièces métriques ne correspond pas à la mise en forme du texte et à la table des matières. C'est notamment le cas lorsque des sonnets ou des triolets sont regroupés sous une même entrée dans la table des matières. L'identification des schémas rimiques, strophiques et des formes fixes implique de restituer un découpage en poèmes. La même démarche analytique s'applique préalablement au traitement automatique pour les vers, car si dans la plupart des cas le vers correspond à une ligne, il peut cependant être réparti sur plusieurs lignes comme c'est souvent le cas dans les pièces de théâtre. Il appartient alors au préparateur du texte de restituer l'unité du vers en introduisant un attribut identifiant les différentes parties du vers. Une unité métrique telle que la strophe ne fait pas l'objet d'une analyse préalable, car nous distinguons la strophe typographique qui est un simple regroupement de lignes délimitées par l'élément <lg> de la strophe métrique qui est une unité métrique définie notamment par un schéma de rimes. Le regroupement de lignes n'acquiert le statut de strophe métrique (par exemple : <lg type="quatrain" schema="abab">) qu'après le traitement des rimes et des formes strophiques. Strophe typographique et strophe métrique ne coïncident pas toujours dans la mesure où il est possible d'avoir : 1) un découpage typographique sans structure strophique qui lui corresponde (ex. André de Chénier, *Poésies antiques*, « L'Aveugle »), 2) une séquence de strophes métriques sans découpage typographique (ex. Voltaire, *La Pucelle d'Orléans*), ou bien encore, 3) un découpage typographique et une structure stro-

---

1. Le corpus *Malherbe* est disponible sur le serveur GitLab de l'université de Caen : <https://git.unicaen.fr/malherbe/corpus> ; ainsi que sur le site web du projet <http://www.crisco.unicaen.fr/verlaine/index.php?navigation=corpus>

phique qui ne coïncident pas (ex. Anna de Noailles, *Poème de l'amour*, I). Dans tous les cas, nos programmes analysent correctement le niveau strophique.

Le traitement automatique produit trois types de résultats : 1) des fichiers XML-TEI enrichis par l'ajout d'éléments et d'attributs relevant de l'analyse métrique et linguistique, 2) des tables analytiques qui alimentent une base de données de relevés métriques, et 3) des pages HTML qui permettent la visualisation des poèmes analysés.

#### LES ÉTAPES DU TRAITEMENT AUTOMATIQUE

Le traitement automatique consiste à appliquer plusieurs programmes qui se présentent sous la forme de modules indépendants mais néanmoins reliés pour certains d'entre eux du fait que, par exemple, le calcul du mètre des vers présuppose notamment l'identification des noyaux syllabiques, des dièses et des différents types de *e* instables. De même, la détermination de la forme globale du poème requiert les schémas de rimes. Dans ce qui suit, les différentes étapes sont illustrées à partir d'un sonnet de Baudelaire.

La première étape, de nature strictement linguistique, consiste à segmenter le texte en mots. La reconnaissance des mots est utile à plusieurs titres : lorsque le recours à une ressource lexicale est nécessaire (dictionnaire des dièses et des mots jonctifs), pour l'identification des *e* masculins, des *e* féminins, des clitiques et des prépositions monosyllabiques<sup>1</sup>, pour le placement de la césure<sup>2</sup> et pour l'extraction des mots-rimes<sup>3</sup>. Le découpage en mots est obtenu à partir d'expressions régulières, et par conséquent, les mots composés et les mots phonologiques réunis par un trait d'union ou une apostrophe sont traités comme plusieurs unités.

La seconde étape consiste à identifier les voyelles et les semi-consonnes. À cette étape, trois types d'unité sont reconnus : les voyelles stables (vs), le *e* instable (vi) et les semi-consonnes (sc). La conversion des graphèmes en phonèmes se fait au moyen de règles ordonnées reposant également sur des expressions régulières.

L'étape suivante (étape 3) a pour objet le traitement des *e* instables. Cinq types sont ici utilisés : *e* masculin, *e* féminin, *e* élidé, *e* ignoré et *e* écarté. Les trois

- 
1. Cf. la métricométrie (Cornulier 1982, 1995, 2010 et Gouvard 2000).
  2. Nous faisons la distinction entre coupe métrique qui est une propriété attachée à une voyelle métrique et césure qui est le report de cette propriété sous forme d'un élément XML-TEI auto-fermant (<caesura/>) en fin de mot du sous-vers. Dans ce vers de Baudelaire (*Les Fleurs du mal*, « Le Tonneau de la haine ») : « La Haine est un ivrogne | au fond d'une taverne. » la coupe est en 6 mais la césure est placée après la voyelle élidée, à la fin du mot.
  3. Ou plus précisément : des unités linguistiques porteuses de la rime, c'est à dire du mot phonologique (voir par exemple la rime *collège::sais-je* (Alfred de Musset, *Poésies nouvelles*, « Dupont et Durand »). De même pour l'extraction des PGTC (voir note 16), comme dans l'exemple suivant : *désir, Idées::des iridées* (Stéphane Mallarmé, *Poésies*, « Prose », 1899),

premiers types sont usuels tandis que les deux derniers sont justifiés par le traitement automatique. Un *e* ignoré est un *e* instable graphiquement présent, et qui n'est pas pris en compte par le calcul métrique. Cela concerne les terminaisons *aient* et *oient* de l'imparfait et du conditionnel, ainsi que le subjonctif du verbe *être* (*soient*). Un *e* écarté est un *e* instable placé entre une voyelle et une consonne et maintenu dans la forme graphique (*dénouement*, *reniement*, *gaieté*, *niera*), alors que les mots concernés font généralement l'objet d'une licence à l'époque classique (*devoûment*, *renîment*, *gaîté*, *nîra*). Ce type de *e* n'est pas métriquement pris en compte dans les textes de notre corpus.

Lors de la quatrième étape, les semi-consonnes sont, soit assimilées à des consonnes, soit traitées comme des voyelles (diérèse). Par défaut, le statut des semi-consonnes est donné par une ressource lexicale dont les entrées sont les mots fléchis avec diérèse obtenus à partir des formes lemmatiques du dictionnaire *Littre* dans sa version XML<sup>1</sup>.

La cinquième étape est l'aboutissement des précédentes. L'identification des voyelles métriques rend possible le calcul de la longueur métrique des vers à partir du nombre de voyelles dans le vers (voyelles stables, *e* masculin et *e* féminin).

La sixième étape — une des plus lourdes avec le traitement de la rime et des formes strophiques — a pour objet le calcul du mètre des vers. Pour tous les vers dont la longueur métrique est inférieure ou égale à 8, le mètre est identique à la longueur métrique. Pour tous les vers dont la longueur métrique est supérieure à 8 voyelles, le calcul du mètre des vers repose sur la mise en œuvre de la métricométrie de B. de Cornulier et sur la détermination d'un profil métrique<sup>2</sup>. Les propriétés PCMF<sup>3</sup> de la métricométrie sont tout d'abord distribuées sur le poème. La distribution de ces propriétés linguistiques (PC) et métriques (MF) permet de déterminer le profil métrique du poème : 6+6, par exemple pour un poème constitué uniquement de vers classiques de 12 voyelles, mais 6–6 pour un poème qui contient au moins un vers avec une propriété PC en 6. Ce profil métrique est par ailleurs indispensable pour les poèmes en vers de 10 syllabes afin de déterminer leur mètre : 4+6 ou 5+5. Une fois que le profil métrique des différentes longueurs métriques du poème est connu, il est alors possible d'établir le mètre de chaque vers. Dans le sonnet de Baudelaire, le profil métrique est 6+6 ; ce qui signifie que tous les vers sont des alexandrins classiques 6+6 (voir illustration 1) :

- 
1. <https://www.littre.org>
  2. Pour le détail de la mise en œuvre de cette procédure, voir Delente & Renault 2015.
  3. P = préposition monosyllabique, C = proclitique (pronoms et déterminants), M = voyelle masculine prétonique, F = voyelle féminine.



Charles Baudelaire

LES FLEURS DU MAL

1857-1861

SPLEEN ET IDÉAL

LXXVI

Le Tonneau de la haine

	La Haine est le tonneau   des pâles Danaïdes ;	6+6
	La Vengeance éperdue   aux bras rouges et forts	6+6
	A beau précipiter   dans ses ténèbres vides	6+6
	De grands seaux pleins du sang   et des larmes des morts,	6+6
5	Le Démon fait des trous   secrets à ces abîmes,	6+6
	Par où fuiraient mille ans   de sueurs et d'efforts,	6+6
	Quand même elle saurait   ranimer ses victimes,	6+6
	Et pour les ressaigner   ressusciter leurs corps.	6+6
	La Haine est un ivrogne   au fond d'une taverne,	6+6
10	Qui sent toujours la soif   naître de la liqueur	6+6
	Et se multiplier   comme l'hydre de Lerne.	6+6
	— Mais les buveurs heureux   connaissent leur vainqueur,	6+6
	Et la Haine est vouée   à ce sort lamentable	6+6
	De ne pouvoir jamais   s'endormir sous la table.	6+6

profil métrique : 6+6

*Illustration 1.*

*Profil métrique du poème, mètre des vers et distribution des propriétés PCMF*

La septième étape, consacrée à l'appariement des vers en couples rimiques (appel-écho) permet ensuite de construire les schémas de rimes indispensables à la reconnaissance des formes strophiques et de la forme globale du poème. Le rapprochement entre les schémas de rimes et des modèles de formes fixes (sonnet, ballade, triolet, terza rima...) ou de schémas de strophes aboutit à la reconnaissance de la forme globale du poème. Dans l'exemple de référence (voir illustration 2), il s'agit d'un sonnet non classique<sup>1</sup> étant donné que le schéma de rimes des deux quatrains est abab et non abba et que ces rimes ne se font pas sur 2 timbres.

1. Sonnet irrégulier pour Morier, 1961.

**Charles Baudelaire**  
**LES FLEURS DU MAL**  
 1857-1861  
 SPLEEN ET IDÉAL  
 LXXVI  
**Le Tonneau de la haine**

La Haine est le tonneau des pâles Danaïdes ;	6+6	a
La Vengeance éperdue aux bras rouges et forts	6+6	b
A beau précipiter dans ses ténèbres vides	6+6	a
De grands seaux pleins du sang et des larmes des morts,	6+6	b
5 Le Démon fait des trous secrets à ces abîmes,	6+6	c
Par où fuiraient mille ans de sueurs et d'efforts,	6+6	d
Quand même elle saurait ranimer ses victimes,	6+6	c
Et pour les ressaigner ressusciter leurs corps.	6+6	d
La Haine est un ivrogne au fond d'une taverne,	6+6	e
10 Qui sent toujours la soif naître de la liqueur	6+6	f
Et se multiplier comme l'hydre de Lerne.	6+6	e
— Mais les buveurs heureux connaissent leur vainqueur,	6+6	f
Et la Haine est vouée à ce sort lamentable	6+6	g
De ne pouvoir jamais s'endormir sous la table.	6+6	g

**sonnet non classique**

genre = f  
 type = écho  
 numéro = 7

*Illustration 2. Rimes, schémas strophiques et forme globale du poème*

La dernière étape, en l'état actuel du projet, porte sur l'évaluation de la « richesse » et de la « qualité » des rimes. Nous nous écartons de l'usage traditionnel de ces deux termes, en adoptant une approche strictement formelle et restrictive<sup>1</sup> : l'évaluation de la « richesse » d'une rime est de nature essentiellement quantitative et repose sur la distinction entre rime et PGTC<sup>2</sup>. Cette évaluation de la richesse de la PGTC porte donc sur l'équivalence éventuelle des phonèmes en amont de la rime. L'évaluation de la « qualité » de la rime est de nature qualitative et ne concerne que la rime (voyelle masculine et ce qui suit) et non ce qui pré-

1. Dans cette approche, l'évaluation de la « richesse » de la PGTC n'inclut pas les éventuelles consonnes qui suivent la voyelle et qui relèvent de la rime.
2. Les traités de versification s'accordent à définir la rime comme l'identité/équivalence de la dernière voyelle masculine et de ce qui la suit, quel que soit la terminologie adoptée notamment concernant ce qui suit la voyelle : phonèmes ou graphèmes. Dès lors que l'évaluation de la « richesse » de la rime porte sur ce qui la précède, il ne peut être question de « richesse de la rime » puisque ce qui précède est extra-rimique. Nous empruntons à Cornulier (1997, 1999) le terme de PGTC (Plus Grande Terminaison Commune) pour désigner la partie du vers en relation d'identité/équivalence qui contient pour le moins le matériel linguistique de la rime et qui participe à la perception de la relation rimique. Par ailleurs, les notions d'extension de la rime (Peureux, 2009) ou d'élargissement de la rime (Cornulier, 1999) sont trompeuses, s'agissant de désigner la partie extra-rimique de la PGTC, dès lors que ce qui précède la rime ne relève pas de la rime.

cède. Il s'agit de distinguer les rimes classiques des rimes non classiques, telle que *gisent::kirghyses*<sup>1</sup> dans laquelle l'identité ou équivalence des terminaisons n'est pas conforme à la tradition. De même dans l'exemple *Glascow::cou*<sup>2</sup> où l'équivalence phonétique des voyelles ne va pas de soi<sup>3</sup>.

Dans ce sonnet de Baudelaire, excepté les terminaisons de vers *liqueur::vainqueur* et *lamentable::table* qui comportent une partie extra-rimique, toutes les autres PGTC se réduisent au matériel linguistique de la rime. Le schéma de l'illustration 3 explicite la relation entre PGTC et rime pour l'une des relations rimiques du sonnet :

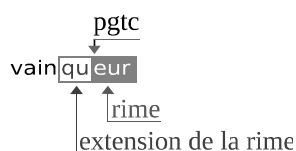


Illustration 3. PGTC et rime

En affectant une valeur chiffrée à la PGTC et à la rime, on obtient ainsi des valeurs calculées qui caractérisent le poème :

- Qualité de la rime : 9/9 (rimes classiques)
- Longueur de la PGTC :
  - Mode de la PGTC : 0 (5 pour 7 rimes)
  - Médiane de la PGTC : 0.0
  - Valeur maximale de la PGTC : 2
  - Valeur moyenne de la PGTC : 0.57
  - Écart-type de la PGTC : 0.9
  - Exemple de PGTC de poids fort (2) : *vain[queur]::li[queur]*
  - Exemple de PGTC de poids faible (0) : *vict[imes]::ab[îmes]*

### LES LIMITES DU TRAITEMENT AUTOMATIQUE

Ces limites sont de deux types. Il y a d'une part, d'inévitables interventions lors du traitement dues à une approche déterministe par règles, d'autre part, des limites liées à l'état d'avancement du projet. Ces limites concernent également l'évaluation de la concordance rythme / sens dues à l'absence d'intégration de composantes syntaxique, sémantique, pragmatique, voire discursive.

- 
1. Jules Laforgue, *L'Imitation de Notre-Dame la lune*, « Nobles et touchantes divagations sous la lune », 1886.
  2. Victor Hugo, *La Légende des Siècles*, « L'aigle du casque », 1877.
  3. Si l'on se réfère à la prononciation usuelle (Oxford University Press) : [glɑ:zɤø], les voyelles [ø] et [u] sont en relation d'équivalence dès lors qu'elles sont phonétiquement proches, à la différence près que le français n'a pas de diphtongue ni de voyelle postérieure quasi-fermée [ɔ̃]. [u] est donc la voyelle la plus proche.

Tout d'abord il faut parfois intervenir lors de l'identification des noyaux syllabiques comme dans les exemples suivants :

- (2) Et d'affreux **curaçaos** troublants, et des absinthes

(Théodore de Banville, *Dans la Fournaise*, 1892, « Consommation »)

- (3) Et l'amiral \*\*\* — Ce n'est pas matelot !

(Tristan Corbière, *Les Amours jaunes*, 1873, « Matelots »)

- (4) Qu'on se lise entre amis ce chant tranquille et **fier**,

(Charles Cros, *Le Coffret de Santal*, 1879, « Le Fleuve »)

Le mot *curaçaos*, avec diphtongue, doit être traité en 3 syllabes et non en 4. Pour le vers de 12 syllabes en (2), il convient d'introduire 2 voyelles virtuelles pour obtenir la longueur métrique du vers conforme au contexte. En (3), une intervention est nécessaire pour lever l'ambiguïté sur la voyelle du mot *fier*, d'autant plus qu'il y a diérèse dans un cas (*fier* = verbe) mais pas dans l'autre (*fier* = adjectif) et que le mot est à la rime.

Une intervention est également parfois nécessaire dans le cas du traitement du *e* instable. Dans la majorité des cas le mot *oui* est jonctif (traitement par défaut) mais parfois il est disjonctif comme dans ce vers de Hugo :

- (5) Un petit prince est-il un petit homme ? **Oui**.

(Victor Hugo, *L'Art d'être grand-père*, 1877, XIII, « L'Épopée du Lion »)

Même chose pour les dièses ; par défaut le mot *lierre* est traité en synérèse mais dans l'exemple suivant, il y a diérèse :

- (6) Le chardon aux longs dards, l'ortie et le **lierre**

(Théophile Gautier, *Albertus ou l'Âme et le Péché*, 1833)

L'exemple suivant de Corneille correspond à une intervention lors du calcul du mètre :

- (7) Un amant digne d'**elle** est trop digne de moi,

(Pierre Corneille, *Pulcherie*, 1672)

Le pronom *elle* est reconnu, à tort, comme pronom clitique (propriété C de la métricométrie) à cause de la séquence bien formée « elle est trop digne de moi ». Il faut donc intervenir pour marquer ce pronom comme pronom fort concluant le premier hémistiche « Un amant digne d'elle ».

Une intervention est parfois nécessaire dans le processus d'appariement des vers. Dans la terza-rima de Heredia<sup>1</sup> (voir illustration 4) :

---

1. José-Maria de Heredia, *Les Trophées*, « La Revanche de Diego Laynez », 1893.

	J'ai forcé ce ragot ; je t'en offre la h <u>ure</u> ! —	6+6	a
	Ruy dit, et tend le chef livide et h <u>é</u> ri <u>ss</u> é	6+6	b
30	Qu'il tient empoigné par l'horrible chev <u>u</u> re.	6-6	a
	Diego Laynez d'un bond sur ses pieds s'est d <u>re</u> ss <u>é</u> :	6+6	b
	— Est-ce toi, Comte infâme ? Est-ce toi, tête ex <u>an</u> gue,	6+6	c
	Avec ce rire fixe et cet œil convuls <u>é</u> ?	6+6	b
	Oui, c'est bien toi ! Tes dents mordent encor ta l <u>an</u> gue ;	6+6	c
35	Pour la dernière fois l'insolente a r <u>ai</u> ll <u>é</u> ,	6+6	a
	Et le glaive a tranché le fil de sa har <u>an</u> gue ! —	6+6	c
	Sous le col d'un seul coup par Tizona t <u>ai</u> ll <u>é</u> ,	6+6	a
	D'épais et noirs caillots pendent à chaque f <u>ib</u> re ;	6+6	b
	Le Vieux frotte sa joue avec le sang c <u>ai</u> ll <u>é</u> .	6+6	a

Illustration 4. Rime forcée

le programme identifie correctement la rime *hérissé::dressé*. Mais il faut intervenir pour forcer la rime suivante *dressé::convulsé*, car le programme considère la rime précédente *hérissé::dressé* comme saturée et procède ensuite à un appariement faux entre *convulsé* et *raillé*. Ce problème vient du fait qu'il y a ici 6 vers avec le même timbre rimique dans le même domaine local. Dans l'exemple précédent, les mots-rimes *convulsé* et *raillé* ne sont séparés que par un vers mais n'appartiennent pas à la même rime.

Toutes ces interventions représentent finalement peu de choses :

Corpus de 471 536 vers (version 1.9.1)				
Étape		Total	Interventions	%
2	Noyaux syllabiques	5 343 300	1 464	<b>0,03 %</b>
3	<i>e</i> élidables (= devant un mot jonctif ou disjonctif)	220 566	457	<b>0,21 %</b>
4	Diérèses	39 164	2 926	<b>7,47 %</b>
5	Propriétés métriques (PCMF)	306 117	200	<b>0,07 %</b>
7	Rimes <sup>1</sup>	35 165	68	<b>0,19 %</b>

<sup>1</sup> analyse partielle du corpus

Les interventions sur les noyaux syllabiques représentent seulement 0,03 % des noyaux syllabiques identifiés. Par contre, les interventions lors du traitement des diérèses (mots traités en diérèse ou en synérèse) représentent pas moins de 7,5 % des diérèses reconnues. Cette proportion importante au regard des autres chiffres s'explique par deux facteurs : 1) le nombre important de textes des XIX<sup>e</sup> et

XX<sup>e</sup> siècles dans notre corpus (78 % du corpus<sup>1</sup>) ; période pour laquelle on observe un affaiblissement des contraintes métriques<sup>2</sup>, 2) la proportion importante des noms propres (35 % des interventions sur les diérèses/synérèses) pour lesquels nous n'avons pas défini de règles par défaut.

Armand Renaud	
Poésies 1860-1880	
LES NUITS PERSANES LA SOLITAIRE	
Le Puits	
DANS le jardin, assise au bord du puits	10
Qu'un soleil ardent   séchait de son hâle,	5+5
Je lui contais ma   tristesse depuis	5-5
Que j'ai vu passer   le cavalier pâle.	5+5
5 Je dis combien l'isolement m'abat,	10
Je dis ma révolte   avec mes alarmes.	5+5
Bien que nul pleur de   mes yeux ne tombât,	5-5
Le puits desséché   se remplit de larmes.	5+5
Au bord du puits je   vins le lendemain ;	5-5
10 J'aurais mieux aimé   la tombe profonde.	5+5
Je ne dis rien, mais   je posai la main	5+5
Sur mon cœur saignant,   en regardant l'onde.	5+5
Des dents de feu me   déchiraient le front,	5-5
Je songeais aux morts   en qui rien ne bouge.	5+5
15 Les pleurs sont peu pour   un cœur qui se rompt ;	5-5
L'eau blanche du puits   devint du sang rouge.	5+5
Hélas ! si lui, le   bien-aimé, voulait	5-5
Porter son amour   dans mon âme sombre,	5+5
Il changerait en   paradis complet	5-5
20 Ma nuit infinie   et mes feux sans nombre ;	5+5
Et le vieux puits, en   écoutant nos doux	5-5
Soupirs de pigeons   et rires de merles,	5+5
Transformerait, pour   faire comme nous,	5-5
Son sang en rubis,   ses larmes en perles.	5+5

Illustration 5. Un seul mètre reconnu pour une alternance de 4+6 et 5+5

Le programme qui détermine le profil métrique des poèmes ne propose qu'un seul profil par longueur métrique. Alors que dans cet exemple tout à fait marginal d'Armand Renaud<sup>3</sup>, il y a une alternance régulière de 4+6 et de 5+5 voir

1. Répartition du corpus : XVII<sup>e</sup> : 16 %, XVIII<sup>e</sup> : 6 %, XIX<sup>e</sup> : 67 %, XX<sup>e</sup> : 11 %.
2. Cet affaiblissement des contraintes s'observe de façon significative pour les diérèses : 2 interventions pour mille vers dans notre corpus du XVII<sup>e</sup>, 3 pour le corpus du XVIII<sup>e</sup> et 7 pour le corpus des XIX<sup>e</sup> et XX<sup>e</sup>.
3. Exemple présenté par Gouvard (1999, 134).

illustration 5).

Le programme a retenu seulement celui des deux qui est le plus favorable. Ce qui a pour conséquence que les vers 4+6 sont traités comme des 5+5 avec une entrave à la coupe ou comme des vers inanalysables.

La détermination du profil métrique repose sur au moins 2 vers, mais parfois c'est loin d'être suffisant, comme dans cet exemple d'Apollinaire où 4 vers ne permettent pas de reconnaître un profil métrique pour ces vers de 12 syllabes (illustration 6) :

**Guillaume Apollinaire**  
Bestiaire  
1908  
Orphée

Que ton cœur soit l'appât et le ciel, la piscine !	12
Car, pêcheur, quel poisson d'eau douce ou bien marine	12
Égale-t-il, et par la forme et la saveur,	12
Ce beau poisson divin qu'est JÉSUS, Mon Sauveur ?	12

*Illustration 6. Profil métrique indéterminé*

La présence d'une propriété P6 dans le vers 3 alors que les places 4 et 9 sont libres de toute entrave dans tous les vers, bloque la reconnaissance d'un profil 6.6<sup>1</sup>. On s'en tient donc ici à la longueur métrique. Dans les exemples suivants de poèmes constitués également de 4 vers, le profil 6+6 s'impose sans difficulté (illustrations 7 et 8) :

**Arthur Rimbaud**  
POÉSIES II  
1870-1872

L'étoile a pleuré rose   au cœur de tes oreilles,	6+6
L'infini roulé blanc   de ta nuque à tes reins	6+6
La mer a perlé rousse   à tes mammes vermeilles	6+6
Et l'Homme saigné noir   à ton flanc souverain.	6+6

*Illustration 7. Profil métrique 6+6 pour un poème de 4 vers (1)*

**André de Chénier**  
ŒUVRES POÉTIQUES  
tome I  
1790  
POÉSIES DIVERSES  
VI

Des monts du Beaujolais   aspect délicieux	6+6
Quand l'Azergue limpide,   enfant de ces beaux lieux,	6+6
Descendant sur les prés   et la côte vineuse,	6+6
Vient grossir de ses eaux   la Saône limoneuse.	6+6

*Illustration 8. Profil métrique 6+6 pour un poème de 4 vers (2)*

---

1. Dans notre notation, 6.6 est une mesure générique qui correspond à 6+6, 6-6, 6=6 ou 6÷6. Pour des explications, voir E. Delente & R. Renault (2015).

## AVANCÉES ET PERSPECTIVES

Les conditions d'écriture et de lecture des textes de Villon diffèrent de celles des textes de Baudelaire. C'est dire que la situation historique d'un texte versifié constitue un paramètre décisif de sa lecture rythmique. Techniquement, il n'est pas impossible de soumettre l'application des algorithmes à une datation précise. La difficulté tient plutôt à la poésie elle-même : les formes métriques nouvelles n'apparaissent jamais simultanément chez tous les auteurs d'une même époque. Le traitement automatique serait-il alors condamné à ne traiter qu'une période remarquable par la stabilité du système métrique à l'œuvre, comme peut l'être le XVII<sup>e</sup> siècle français, par exemple ? En réalité, la métricométrie est apte à repérer des évolutions dans l'écriture versifiée. La recherche dans le vers, non pas d'une propriété simple (P ou C ou M ou F) mais d'une propriété complexe permet d'intégrer à l'analyse automatique, autrement sourde à l'histoire, une dimension historique. Un bon exemple nous est fourni par le traitement rythmique du *e* féminin 7<sup>e</sup>.

### *Le e féminin 7<sup>e</sup> dans les 12v*

Durant la période classique, les deux hémistiches d'un vers composé sont l'objet d'un traitement rythmique autonome, c'est pourquoi on ne rencontre pas de vers dans lequel un *e* féminin commencerait une seconde sous-mesure comme en (8) :

(8) Mes amants n'appartiennent pas aux classes riches, F7

(Verlaine, *Hombres*)

Vers 1880, de tels vers commencent à apparaître. La distribution des propriétés métricométriques dans ces vers nous a conduit à constituer trois sous-corpus.

### *Les F7 libres en 4<sup>e</sup> et 8<sup>e</sup> voyelles*

Sur les 310 vers F7 de notre corpus<sup>1</sup>, la métricométrie automatisée en relève 160 dont les 4<sup>e</sup> et 8<sup>e</sup> voyelles sont libres, comme le suivant :

(9) Dans votre sein, sur votre cœur qui fut le nôtre, F7

(Verlaine, *Sagesse*, VI)

Le *e* féminin empêche une lecture 6–6 mais les 4<sup>e</sup> et 8<sup>e</sup> frontières syllabiques non entravées autorisent à supposer un rythme ternaire 4+4+4, et nous observons de fait que, dans la grande majorité des cas, ce rythme apparaît comme assez évident :

(9') {Dans votre sein,} {sur votre cœur} {qui fut le nôtre,}

---

1. À la différence de Cornulier (1982), la propriété F7 n'est pas retenue à la frontière de morphème : « Ce n'est pas la méchanceté, c'est la bonté » (Verlaine, *Sagesse*, III, 1881).



La plupart des vers F7 de ce sous-corpus sont ternaires. Quelques-uns seulement offrent une meilleure convergence en tant que semi-ternaires. Il suffirait alors d'ajouter une règle du type de celles énoncées pour traiter les exemples de 15 à 18.

Quelques vers plus préoccupants, tous de Verlaine, inviteraient plutôt à une lecture 6-6, comme ceux-ci :

- (10) Le dernier coup de vêpres a sonné : l'on tinte  
(Verlaine, *Liturgies Intimes*, « Vêpres Rustiques »)
- (11) Du pépiement des autres « hôtes de nos bois »  
(Verlaine, *Dédicaces*, « Le pinson d'E\*\*\* »)
- (12) Et d'une main fiévreuse, mais honnête, dame  
(Verlaine, *Biblio-sonnets*, « L'arrivée du catalogue »)
- (13) Et les rinceaux gothiques, fins d'apothéoses  
(Verlaine, *Poèmes Divers*, « Oxford »)

La régularité observée dans ce sous-corpus permet à un algorithme de proposer une analyse ternaire, voire semi-ternaire, satisfaisante pour environ 150 vers sur 160 avec cependant une petite dizaine de vers qui restent très problématiques.

#### **Les F7 non libres en 4<sup>e</sup> et 8<sup>e</sup>**

Ce sous-corpus comporte 27 vers F7, qui présentent tous des 4<sup>e</sup> et 8<sup>e</sup> frontières syllabiques entravées, empêchant toute lecture ternaire et forçant la lecture binaire 6-6. Un exemple :

- (14) Et ce soin de se faire soi-même cocu            F7  
(Verlaine, *Femmes*, « Hommage dû »)

Excepté pour un ou deux vers, l'analyse en 6-6 apparaît très satisfaisante. Il apparaît encore une fois que les régularités observées autorisent l'application d'une règle.

#### **Les F7 et PCMF4 ou PCMF8**

Reste 123 vers qui présentent cette propriété complexe et qui peuvent donc être des 6-6 ou 8+4 ou 4+8. Ces vers, difficiles à interpréter, ne nous permettent pas de décider si la mesure binaire l'emporte sur la mesure semi-ternaire, ni même s'il est possible de répondre à cette question sans examiner chaque vers dans son contexte. Si l'on ajoute à cela que ce sous-corpus réunit 14 poètes sur une période s'étendant de 1891 à 1922 et que l'interprétation métrique de ces vers est sujette à discussion entre métriciens eux-mêmes, il paraît plus raisonnable pour l'instant

d'adopter une attitude prudente consistant à afficher selon les cas 6-6 et 8+4 ou 6-6 et 4+8.

Un autre enjeu d'importance concerne la prise en compte des relations entre rythme et sens. La métricométrie est une méthode fiable pour le calcul du mètre puisqu'elle nous fournit des observations solides sur lesquelles fonder une argumentation en faveur d'une analyse métrique donnée. Les propriétés P, C et M permettent de repérer des cas de divergence rythme / sens fortement marqués. Mais la seule métricométrie automatisée est inapte à saisir tous les paramètres assurant une consistance satisfaisante des expressions métriques et produit parfois des analyses contestables. Les vers suivants analysés comme 4+4+4 (mètres de substitution ternaires<sup>1</sup>) offrent des sous-vers divergents :

- (15) N. de D. ! J'ai | rien voyagé | d'puis mon dergnier  
(Verlaine, *Autres vieux Coppées*, VIII)
- (16) Une goutte, et | qu'il faudrait boire | à pleine amphore !  
(Verlaine, *Poèmes Divers*, « Meliora »)
- (17) Désormais, Reine | Légitime | de mes sens  
(Verlaine, *Poèmes Divers*, « À ma femme »)
- (18) Faute de rente | ou d'achat bien | vus, son crédit ?  
(Verlaine, *Biblio-Sonnets*, « Désappointement »)

Dans les vers (15) à (17), le premier sous-vers est divergent à la finale : {N. de D. ! J'ai} ; {Une goutte, et} ; {Désormais, Reine} alors qu'en (18), c'est le troisième sous-vers qui est divergent à l'initiale {vus, son crédit ?}. Une analyse semi-ternaire paraît préférable. Or, ces vers présentent tous la propriété commune d'avoir un signe de ponctuation à la frontière syllabique qui précède ou qui suit immédiatement l'une des deux coupes ternaires, 4<sup>e</sup> ou 8<sup>e</sup>.

La ponctuométrie<sup>2</sup> peut alors nous être utile. Cette méthode consiste à repérer les signes de ponctuation vus comme indicateurs de l'organisation syntaxique/sémantique des expressions métriques. La requête consiste à repérer les propriétés complexes suivantes :

- une coupe 4<sup>e</sup> et un signe de ponctuation 4<sup>e</sup> ± 1
- une coupe 8<sup>e</sup> et un signe de ponctuation 8<sup>e</sup> ± 1<sup>3</sup>

Dans tous ces exemples, le signe de ponctuation 4<sup>e</sup> ± 1 est un indice de coupe 4<sup>e</sup> mal-convergente, défavorisant la coupe ternaire 4+4+4 et favorisant la coupe semi-ternaire 8<sup>e</sup> :

- 
1. Voir Cornulier (1982) et (1995, 98-102).
  2. Cornulier (1995, 265)
  3. « 4<sup>e</sup> ± 1 » signifie donc un signe de ponctuation 3<sup>e</sup> ou 5<sup>e</sup>, « 8<sup>e</sup> ± 1 », un signe de ponctuation 7<sup>e</sup> ou 9<sup>e</sup>.

- (19) {N. de D. ! J'ai rien voyagé} {d'puis mon dergnier} 8+4  
 (20) {Une goutte, et qu'il faudrait boire} {à pleine amphore !} 8+4  
 (21) {Désormais, Reine Légitime} {de mes sens} 8+4

Alors que le signe de ponctuation  $8^e \pm 1$  est un indice de coupe  $8^e$  mal-convergente, défavorisant la coupe ternaire 4+4+4 et favorisant la coupe semi-ternaire  $4^e$ .

- (22) {Faute de rente} {ou d'achat bien vus, son crédit ?} 4+8

Ces règles n'ont pas encore été testées.

Dans d'autres cas, la prise en compte des relations rythme / sens enrichirait nos analyses. Dans la deuxième moitié du XIX<sup>e</sup>, l'équivalence métrique entre les vers s'affaiblit progressivement avant de disparaître totalement. Il faut ici souligner la difficulté de soumettre une telle production au traitement automatique. Les algorithmes doivent en effet traiter avec la même efficacité le théâtre classique, la poésie de Hugo, celle de Verlaine et certains poèmes d'Apollinaire. Durant cette période, apparaissent des 12v avec une préposition univocalique, un déterminant ou un proclitique en 6<sup>e</sup> voyelle, manifestant un enjambement qui, pourtant, n'empêchait pas de percevoir ces vers comme 6-6 en contexte 6+6<sup>1</sup> :

- (23) Sans violence, sans | remords et sans envie :

(Verlaine, *La Bonne Chanson*, IV.)

- (24) De par la grâce le | sourire et la bonté

(Verlaine, *La Bonne Chanson*, IV.)

- (25) Je le sais bien, je le | sais bien, nous nous aimons

(Michaël, *Fragments et ébauches de poésies*, « La forêt sacrée ».)

Il n'est pas rare que ces vers présentent également une organisation syntaxique ternaire évidente :

- (23') {Sans violence,} {sans remords} {et sans envie :}

- (24') {De par la grâce} {le sourire} {et la bonté}

- (25') {Je le sais bien,} {je le sais bien,} {nous nous aimons}

Au rythme métrique 6-6 se superpose un rythme 4+4+4, voire un mètre 4+4+4 dont la périodicité interne vient comme compenser l'absence d'équivalence contextuelle. Petit à petit pourtant, s'instaure une équivalence contextuelle *acquise* 4+4+4 = 6.6. À ce jour, il est encore prématuré pour nous d'envisager les modalités de repérage de tels patrons syntaxiques et d'évaluer la

---

1. Pour une argumentation en faveur de cette analyse, voir Cornulier (1982, 192-204) et Gouvard (2000).

consistance des trois expressions métriques. Les outils proposent donc un mètre 6-6 sans rendre compte de cette ambivalence métrique.

## CONCLUSION

Le projet *Anamètre* dispose d'un corpus important de textes versifiés s'étendant sur plus de trois siècles. Les outils élaborés sont aptes à traiter les formes métriques produites dans le cadre d'un système métrique relativement stable, mais également les formes nouvelles apparues pendant la période d'affaiblissement de ce système (4+4+4, 8+4 et 4+8, 6-6 avec post-tonique récupérée ; mélange 4+6 et 6+4). La notion de profil métrique permet d'intégrer une dimension contextuelle dans l'analyse métrique des vers. Et enfin, les programmes analysent aussi bien les régularités métriques que strophiques, aussi bien les formes stéréotypées (vers simples de 6v, 8v etc. ; vers composés de 6+6 et 4+6 ; distiques, quatrains, sizains, sonnets réguliers, rondeaux...) que les formes complexes ou rares (3+6 et 4+5, 5+6, 5+8, formes variées du sonnet, pantoum...).

Par ailleurs, ces programmes sont également à même d'évaluer la qualité des rimes et la longueur de la PGTC. Cette évaluation à la fois qualitative (qualité de la rime) et quantitative (longueur de la PGTC) permet ainsi une étude plus fine de l'affaiblissement des contraintes sur les relations d'équivalence entre fins de vers pour une période qui s'étend du XVII<sup>e</sup> au début du XX<sup>e</sup> siècle.

L'intérêt de disposer d'un aussi large corpus et d'une base de données analytique importante est évident. À partir de ces données, qui sont le résultat de l'application des programmes d'analyse automatique, il est possible d'établir différents répertoires (mètres, strophes, formes fixes...) avec leur distribution en fréquences et pourcentages sur l'ensemble du corpus. Les données obtenues permettent d'appréhender la discipline au travers de l'évolution de ses formes, et ce aussi bien à l'échelle d'un auteur, d'un recueil ou d'une période. Ce qui rend possible également une caractérisation métrique exhaustive et détaillée de l'œuvre d'un auteur tout en retraçant son évolution.

L'amélioration des analyses et leurs extensions indiquent qu'une des voies à suivre est une étude portant sur les relations entre rythme et sens. Les poètes, tout en s'appuyant sur les principes rythmiques du français, projettent sur le flux du discours des régularités rythmiques systématiques. En quoi le développement d'un discours construit sur un mouvement rythmique régulier influe-t-il sur la construction de sa/ses significations ? Plus précisément, en quoi les principes rythmiques propres au français, soumis à des régularités systématiques, participent-ils à l'élaboration pas à pas (unité métrique après unité métrique) de la construction des significations dans le poème. Le projet *Anamètre* met à la disposition du chercheur des milliers d'unités métriques de chaque niveau. Il peut aider à cette réflexion par le repérage automatique des cas d'unités métriques discordantes (enjambement, rejet, contre-rejet) favorisant ainsi une observation

rarement pratiquée jusqu'à ce jour. Il s'agirait là d'une contribution importante à une meilleure compréhension du discours rythmé.

#### RÉFÉRENCES

- Beaudouin V. & F. Yvon 1994. « Le Métromètre et les douze positions alexandrines », *Mezura*, n° 32, (Paris : INALCO) pp. 1–31.
- Beaudouin V. 2000. *Rythme et rime de l'alexandrin classique, étude empirique des 80 000 vers du théâtre de Corneille et Racine*, thèse de doctorat. EHESS.
- Beaudouin V. 2002. *Mètre et rythmes du vers classique : Corneille et Racine* (Paris : Honoré Champion).
- Bernard, M. 2017, « Études métriques automatisées : où en est-on aujourd'hui ? », *Modèles et nombres en poésie*, dir. V. Magri-Mourgues et G. Purnelle (Paris : Honoré Champion), pp. 137–145.
- Cornulier, B. de, 1982. *Théorie du vers. Rimbaud, Verlaine, Mallarmé* (Paris : Seuil, coll. 'Travaux Linguistiques').
- Cornulier, B. de, 1995. *Art poétique. Notions et problèmes de métrique* (Lyon : Presses Universitaires de Lyon, col. IUFM).
- Cornulier, B. de, 1997. *Voyelle fondamentale, énonciation pendante et représentation de constituant*, tapuscrit (Nantes : CEM).
- Cornulier B. de, 1999, *Petit dictionnaire de métrique*, Université de Nantes (disponible ici : [http://www.crisco.unicaen.fr/verlaine/ressources/Dict\\_MET.pdf](http://www.crisco.unicaen.fr/verlaine/ressources/Dict_MET.pdf)).
- Delente E. & R. Renault. (éd.), 2015. « Traitement automatique des textes versifiés : problématiques et pratiques », *Langages* n° 199 (Paris : Larousse).
- Gouvard, J. M. 2000. *Critique du vers* (Paris : Champion).
- Gouvard, J. M. 1999. *La Versification* (Paris : PUF).
- Hervé N. 2011. *Structures et superstructures métriques dans l'œuvre de Clément Marot (1496–1544)*, thèse de doctorat, Université de Nantes.
- Morier H. 1989, *Dictionnaire de poétique et de rhétorique*, 4<sup>e</sup> édition, revue et augmentée (Paris : PUF, 1961).
- Peureux G. 2009, *La Fabrique du vers* (Paris : Éditions du Seuil).
- Roubaud J. 1986 et 1988, « Dynastie : études sur le vers français, sur l'alexandrin classique », *Cahiers de poétique comparée* n° 13 et n° 16 (Paris : INALCO).