ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES

ANNEE 2000

THESE

Pour l'obtention du grade de

DOCTEUR DE L'EHESS

SCIENCES DU LANGAGE

présentée et soutenue publiquement

par

Valérie BEAUDOUIN

Le 24 janvier 2000

RYTHME ET RIME DE L'ALEXANDRIN CLASSIQUE

Étude empirique des 80 000 vers du théâtre de Corneille et Racine

Sous la direction de M. Jacques ROUBAUD

JURY

M. Bernard CERQUIGLINI

Mme Sophie FISHER

M. Benoît HABERT

M. Pierre LUSSON

M. François RASTIER

M. Jacques ROUBAUD

REMERCIEMENTS

SOMMAIRE

INTRODUCTION.....9

I. DONNEES HISTORIQUES ET THEORIQUES 21

II. LE CORPUS.....51

III. LA RIME.....69

IV. DIERESE, E "MUET", LIAISON : TROIS LIEUX DE VARIATION PHONOLOGIQUE.....139

V. MARQUAGES POSSIBLES, MARQUAGES CHOISIS 193

VI. LE METROMETRE : L'OUTIL ET SES LIMITES 225

VII. SON, SYNTAXE ET PROSODIE DE L'ALEXANDRIN « ORDINAIRE » .263

VIII. LA CONTRIBUTION DU RYTHME A LA FORME POETIQUE 341

CONCLUSION.....397

BIBLIOGRAPHIE.....409

INTRODUCTION

L'objet de ce travail est le rythme de l'alexandrin classique. Nous avons opté pour une approche empirique qui porte sur l'œuvre théâtrale complète de Corneille et Racine, soit près de 80 000 vers. C'est la taille de ce corpus et son exploitation systématique (aux niveaux phonétique, lexical et morpho-syntaxique) qui distingue notre travail des recherches antérieures.

Jusqu'à nos jours, en effet, les analyses empiriques menées sur l'alexandrin classique ont porté sur des corpus qui ne dépassaient pas le millier de vers. La plus importante, que l'on doit à Jacques Roubaud, s'appuie sur les mille premiers vers de *Bérénice* [1988]. Dans sa thèse consacrée au vers de la seconde moitié du XIXe siècle, Jean-Michel Gouvard [1994] explore, à titre comparatif, 500 alexandrins classiques du XVIIe siècle.

Si d'autres travaux empiriques ont porté sur le vers du XIXe siècle, ils visaient principalement à explorer des phénomènes spécifiques. Citons le « desserrement » des contraintes dans *La Vieillesse d'Alexandre* de Roubaud [1978, 1988], à propos du vers de Hugo, Nerval, Baudelaire, Rimbaud et Mallarmé ; la répartition de six propriétés syntaxico-phonétiques sur les douze positions de 1810 vers de Rimbaud, Verlaine et Mallarmé dans *Théorie du vers* de Benoît de Cornulier [1982b] ; la répartition de propriétés similaires à celle de Cornulier sur 35 000 vers de la seconde moitié du XIXe siècle, dans la thèse de Gouvard [1994]. Des analyses portant sur des points particuliers, comme la césure, ont donc pu être réalisées sur des volumes de vers importants, mais jusqu'ici la lourdeur du travail interdisait une approche systématique.

C'est grâce à l'outil informatique que nous avons pu traiter systématiquement les phénomènes rythmiques sur un corpus aussi important que le nôtre. La démarche est simple : il s'agit de marquer les composantes rythmiques de chaque vers, de rassembler tous ces vers marqués dans une grande base de données, puis de dégager les grandes lois ou tendances par

des calculs statistiques. C'est pour accéder à cette approche statistique que nous avons travaillé sur un grand corpus. La statistique a le mérite de fournir des représentations assez parlantes des phénomènes de distribution temporelle ou spatiale, comme c'est le cas du rythme. Nous pouvions donc légitimement espérer accéder par là à des résultats nouveaux et les décrire de manière quantifiée et reproductible. C'est bien le cas, comme on le verra.

Ce programme simple s'est pourtant heurté à des difficultés plus importantes que prévu. Nous pensions en effet à l'origine pouvoir nous appuyer sur les définitions existantes du rythme pour décrire la forme particulière de celui du vers alexandrin. Mais la notion de rythme, pour évidente qu'elle soit à la perception, est loin d'être claire ; la littérature en fournit des définitions pour le moins diverses, qui ne sont pas aisément formalisables, et ne résistent pas aux exigences d'explicitation de la programmation informatique[1]. À la réflexion, cela ne saurait étonner s'agissant d'un concept qui met en rapport les notions difficiles de « forme » et de « temps ».

Curieusement, comme l'a montré Benvéniste [*1951*, 1966], la plupart des définitions du rythme s'appuient sur une interprétation inexacte de l'étymologie, qui associe le concept au mouvement régulier des flots. Or le verbe *couler* (ῥεῖν) dont dérive *rythme* (ῥυθμός) en grec, n'a jamais été utilisé pour désigner le mouvement de la mer. Benvéniste revient aux usages premiers du terme. Le mot *rythme* en grec serait utilisé par Démocrite pour désigner la « forme », « en entendant par là la forme distinctive, l'arrangement des parties dans un tout » [p. 330]. Il désigne des configurations toujours susceptibles de changer, d'où l'importance de l'étymologie, *couler*. Platon modifie le sens du terme et lui donne son sens actuel : « configuration des mouvements ordonnés dans la durée » [p. 335]. Comme le soulignait Fraisse [1956] « il y a un sens général du mot rythme avec deux pôles principaux : périodicité et structure ». Le rythme désigne la répétition d'un phénomène et la structure d'une suite d'événements ou de stimulations. Les deux aspects du rythme sont étroitement mêlés puisque la répétition d'un phénomène implique que ce phénomène ait une structure reconnaissable. Les travaux de Fraisse [1956] sur la psychologie du rythme se sont attachés à montrer cette double dimension du rythme. Cette

[1] Pierre Getzler a recensé dans différentes langues plus de 500 définitions du terme rythme contradictoires, imprécises, relevant de théories et idéologies très diverses. Ce matériau devra être étudié d'une manière approfondie.

distinction entre périodicité et structure rejoint en partie celle qui est faite entre le mètre et le rythme, où le mètre désigne la périodicité et le rythme l'organisation [p. 112].

Nous retiendrons donc du mot *rythme* qu'il désigne la répétition de formes ou configurations dans la séquence temporelle. Ce sont ces configurations que nous allons explorer. Toutefois, comme à chaque fois que l'on examine une évidence, nous avons rencontré un problème ardu, qui pourrait s'exprimer ainsi : qu'est-ce que la répétition, autrement dit qu'est-ce que le *même* par rapport au *différent* ? Quelles sont les dimensions sous lesquelles nous devons, du point de vue du rythme, comparer des objets pour juger de leur identité ; à partir d'où considère-t-on qu'il y a différence ? Plus précisément : puisqu'il est évident que deux vers successifs ne sont pas rigoureusement identiques et qu'au sein du vers les mots et sons varient également, à quel niveau de contenu et de détail devons-nous décrire le vers pour appréhender les événements qui se répètent ?

Pour décrire ces formes qui se répètent, notre approche empirique a été guidée par la théorie du rythme développée par Pierre Lusson [1973, 1998], qui propose un formalisme général pour l'étude des phénomènes rythmiques, et par les applications qui ont pu en être faites au Centre de Poétique Comparée. Dans ce cadre, le rythme de l'alexandrin admet plusieurs niveaux d'analyse : celui de la position ou syllabe métrique, celui du vers (et de l'hémistiche) et celui défini par la rime. Nous les avons systématiquement explorés. Il nous a fallu, avant de décrire chacun de ces niveaux, en définir les éléments de manière suffisamment opératoire pour être compris même par une machine.

Sur ces principes, nous avons mis au point un outil, le METROMETRE, qui permet :

- de segmenter chaque vers concret en syllabes qui correspondent de manière univoque aux douze positions métriques du modèle ;
- de caractériser chaque vers concret par des marquages phonétiques, lexicaux, morpho-syntaxiques et prosodiques, qui permettront de construire une figure générale et multiple du vers et des hémistiches.

Repérer les positions métriques

Si nous parvenons sans peine à identifier qu'un vers est un alexandrin à la simple lecture, la tâche est nettement moins simple quand elle est confiée à l'ordinateur. Pour identifier les positions métriques dans le vers, il faut mettre à jour les principes et les règles qui permettent le passage de la séquence du vers à un modèle abstrait qui, à ce niveau de l'analyse, est décrit comme une séquence concrète de douze positions.

Il s'agit d'établir « en vertu d'une règle de transformation (*mapping rule*) (...) une correspondance biunivoque entre les syllabes d'un vers et les X [éléments] de la structure métrique » [Jakobson & al., 1972, p. 104]. Pour Halle, il n'existe en gros pour le français qu'une règle de transformation :

> La règle de transformation qui concerne le type de poésie française le plus important s'énonce donc ainsi :
> Chaque élément du mètre (chaque X) doit correspondre soit à un *e* muet suivi d'une consonne, soit à une voyelle autre qu'un *e* muet indépendamment du contexte.
>
> [Jakobson & al, 1972].

Malheureusement, les choses ne sont pas aussi simples que semble le croire Halle et la construction des règles de transformation qui permettent de passer de la séquence graphique au modèle abstrait du vers ne va pas de soi.

Nous rencontrons en effet différentes difficultés, classiques en phonologie :

1. Qu'est-ce qu'une syllabe, comment découper les syllabes dans la séquence continue des phonèmes ?
2. Qu'est-ce qu'un noyau vocalique de position métrique ? Peut-on utiliser les seize phonèmes vocaliques du français ?
3. Que faire des séquences de voyelles adjacentes ? Combien de positions occupent-elles ? Comment identifier de manière systématique diérèse et synérèse ?
4. Quel traitement adopter pour les *e* muets ? Quand sont-ils maintenus, quand sont-ils élidés ?
5. Faut-il modéliser les phénomènes de liaison ? Si oui, comment ?

Nous présenterons dans le détail les réponses que nous avons choisies ou élaborées à chacune de ces interrogations. Nous traiterons donc successivement la diérèse, le *e* muet et la liaison.

Décrire les positions métriques

Le second objectif (le premier étant l'identification des positions métriques) est celui de la représentation du contenu de chaque position dans notre système, afin de pouvoir construire la figure rythmique du vers. Quels descripteurs retenir pour appréhender la question du rythme du vers ? Ce choix est dangereux, car il va limiter les régularités que nous pourrons percevoir ; et nous rendra aveugle aux régularités des phénomènes non marqués. Pour limiter ce risque, nous avons fait le choix de projeter plusieurs marquages relevant de composantes linguistiques différentes : phonétique, syntaxe et prosodie. Ces choix résultent en partie d'un arbitrage fondé sur la difficulté technique de mise en œuvre. Nous examinerons les marquages adoptés, mais aussi d'autres que nous aurions aimé pouvoir mettre en œuvre.

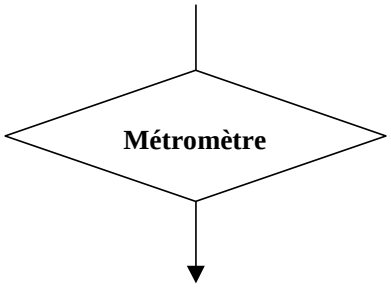
Le métromètre

Le métromètre est une réponse, encore imparfaite, à notre projet de construire un outil d'analyse systématique du vers classique. Cet outil informatique répond au premier objectif : il identifie les positions métriques à partir de la séquence graphique dans n'importe quel vers de type « classique » et en donne une représentation dans l'alphabet phonétique international. Il répond également au second objectif, puisqu'il produit des descriptions phonétiques, morpho-syntaxiques et prosodiques de chacune des positions du vers, comme on peut le voir sur la figure de la page 14. Le métromètre lit chaque vers, en fait une analyse syntaxique et produit une transcription phonétique, organisée en syllabes métriques. En sortie, le métromètre attribue à chacune des positions du vers un certain nombre de marquages : la syllabe métrique, son noyau vocalique, sa position dans le mot (intérieur ou fin de mot), sa catégorie syntaxique, le fait qu'elle porte ou non un accent et enfin le nombre de positions du vers.

Précisons que la représentation phonétique produite par le métromètre ne prétend aucunement définir ce que devrait être une bonne diction du vers. Il s'agit seulement d'une représentation abstraite conforme au modèle du vers, que l'acteur ou le lecteur reste libre ou non d'adopter.

De cette nuit, Phénice, as-tu vu la splendeur ?

Racine, *Bérénice*, vers 302.



Représentation multidimensionnelle :

Syllabes métriques	d ə	s ɛ	t $ə_f$ [2]	n ɥ i	f e	n i s	a	t y	v y	l a	sp l ɑ̃	d œ r
Voyelles métriques	ə	ɛ	ə $_f$	ɥ i	e	i	a	y	y	a	ɑ̃	œ
Repérage des fins de mots	1	0	1	1	0	1	1	1	1	1	0	1
Catégories syntaxiques	7	6	6	0	4	4	1	6	2	6	0	0
Marquage accentuel	0	0	0	1	0	1	0	1	1	0	0	1
Nb syllabes	12											

Clef de lecture du tableau : La cinquième syllabe métrique correspond à la séquence phonémique /f e/, elle n'est pas située en fin de mot (*cf.* repérage des fins de mots = 0) et appartient à un nom propre (*cf.* catégorie syntaxique = 4).

Repérage des fins de mot :

0 = syllabe interne	1 = syllabe finale

Catégories syntaxiques (selon l'analyseur syntaxique Sylex)

0 = Nom	5 = Sigles[3]
1 = Verbe	6 = Déterminant, pronom
2 = Adjectif, participe passé	7 = Préposition
3 = Adverbe	8 = Conjonction
4 = Nom propre	9 = Pronom relatif

Marquage accentuel

0 = syllabe ne portant pas d'accent de mot	1 = syllabe portant un accent de mot

Figure 1. Fonctionnement du métromètre et représentation métrique

[2] e_f : désigne un e muet « faible », susceptible d'élision selon son environnement.

[3] Cette catégorie syntaxique n'est évidemment pas pertinente pour le vers classique : nous ne l'indiquerons plus.

L'outil n'existait pas, il a donc fallu le construire : le métromètre a été construit à partir d'un phonétiseur du français contemporain réalisé par François Yvon [1993], qui s'appuyait lui-même sur un analyseur syntaxique développé par Patrick Constant [1991]. Le métromètre a été mis au point en collaboration avec François Yvon en ajoutant au phonétiseur un ensemble de règles supplémentaires adaptées au cas du vers. La diérèse, le *e* et la liaison, qui sont les trois lieux de variation phonologique, ont donné lieu à des traitements spécifiques [Beaudouin et Yvon, 1994 et 1996].

C'est la première fois qu'un outil d'analyse systématique de grands corpus de vers est ainsi construit.

Révéler la figure du vers

Cependant, l'objectif n'était pas de fabriquer l'outil, mais de l'utiliser pour explorer les phénomènes métrico-rythmiques à l'œuvre dans les vers. Le découpage en positions métriques, tout comme les marquages, sont conçus pour être soumis à des traitements statistiques. Nous avons donc « préparé » un corpus important qui puisse servir de matériau d'investigation : les œuvres théâtrales complètes de Corneille et Racine, qui sont typiques de la période « classique » de l'alexandrin et qui sont composées essentiellement en alexandrins. Ce corpus a été édité sous une forme standardisée afin de pouvoir être « passé au métromètre ».

À l'issue de cette opération, chaque vers est décrit par un ensemble de variables. Plus précisément, à chaque position du vers sont associées autant de variables qu'il y a de marquages. Ainsi organisées, les données sont directement accessibles pour des traitements statistiques qui vont nous permettre une approche empirique approfondie et systématique. C'est là que se voit récompensé l'investissement conséquent que représentent l'harmonisation éditoriale du corpus, la construction du métromètre et la structuration en bases de données. Nous serons ainsi à même d'explorer le vers à une échelle à ce jour inédite.

Toutes les informations produites par le métromètre seront tour à tour mobilisées pour construire les composantes phonétiques, morpho-syntaxiques et prosodiques de la figure métrico-rythmique de l'alexandrin classique, mais aussi pour explorer des phénomènes rythmiques et métriques spécifiques. Ces figures sont obtenues par l'analyse statistique de la base de vers. On verra qu'il se dégage bien de ces analyses, grâce à la

taille importante de notre corpus, une « forme » du vers. Le rythme y apparaît comme une structure « feuilletée » à laquelle participent les différents niveaux de la langue.

La forme même des structures ainsi révélées va nous amener à voir différemment le rythme, en même temps qu'elle clarifie les relations entre mètre et rythme. Le grand nombre rendra évidente la question de la forme du rythme, mais ne nous dira pas grand chose sur la perception de la forme sur des vers isolés, où seul le mètre est évident.

Enfin, puisque la stylistique est l'étude des corrélations entre niveaux (phonétique, lexical, syntaxique, sémantique et rythmique) [Rastier, 1989, p. 10], une meilleure description du rythme des vers devrait permettre de mieux intégrer la dimension rythmique dans les analyses stylistiques. Qu'apporte la forme du vers à la construction de l'interprétation, à la lecture du texte ? Nous avons essayé de mettre à jour quelques corrélations entre le rythme et le sens dans les pièces de Corneille et Racine. Ces résultats restent certes parcellaires, mais l'on verra qu'ils ébauchent une perspective plus générale de la forme poétique, dont le rythme serait une des composantes.

La rime

Le modèle métrique de l'alexandrin comprend des niveaux d'organisation qui dépassent le vers : la paire de vers rimant ensemble et le quatrain d'alexandrins à rimes plates avec alternance en genre. Ces niveaux d'organisation de la structure métrique sont définis par la rime (équivalence des fins de vers et alternance en genre). Toujours à la recherche de structures qui se répétaient, nous avons cherché à mieux décrire cet élément. Nous avons ainsi examiné les conditions graphico-phonétiques nécessaires pour que deux vers riment ensemble et montré comment ces conditions variaient selon la nature de la terminaison commune. Nous avons organisé en catégories de mots pouvant rimer ensemble (« rimèmes ») tous les couples de mots-rimes attestés dans le corpus, pour dégager quelques caractéristiques de la rime classique.

La question de la rime sera abordée dans ce texte avant le métromètre, parce qu'elle a une certaine indépendance par rapport au reste du modèle métrique de l'alexandrin, mais surtout parce que le travail sur la rime nous a permis de traiter différemment la question des noyaux vocaliques des

positions métriques, en lui donnant le soubassement empirique qui lui faisait défaut. Plus généralement, l'approche systématique sur un corpus important nous a ici encore ouvert des perspectives nouvelles et fourni des résultats inédits, même si nous n'avons pu présenter ici une théorie définitive de la rime.

D'autres voies pour explorer le rythme du vers

Cependant, l'approche empirique a ses limites, et notre travail ne prétend pas épuiser la question. Ainsi avons-nous privilégié une représentation particulière du vers pour en faire l'analyse, et avons-nous ignoré des dimensions du phénomène qui pourtant existent. Parlons-en tout de suite.

Pour l'analyse du rythme de l'alexandrin, notre point de départ a été la forme écrite, car c'est la représentation qui subsume le mieux la variabilité des réalisations du vers. On est donc assuré de travailler sur la meilleure représentation canonique, et de résoudre ainsi en partie les difficultés inhérentes à la variabilité des réalisations orales. C'est d'ailleurs l'approche la plus couramment adoptée. Elle a ses limites, qui ne sont pas dirimantes. Entre la manière dont les mots sont écrits et leur prononciation, l'écart est souvent important, il nous a donc paru utile d'adopter une orthographe minimale, celle que propose l'alphabet phonétique international[4] (API), pour transcrire nos vers (nous aurons l'occasion de détailler ce point). C'est donc sous forme graphique que nous avons étudié le rythme, phénomène qui exige une représentation phonétique, qu'elle soit mentale ou sonore.

Cependant, pour appréhender la question du rythme dans le vers, trois autres points de vue auraient également pu être adoptés : celui de la production (composition de vers), celui de la déclamation, celui de la réception (lecture ou audition des vers), points de vue que nous exposons ici.

Concernant la production, il s'agirait de comprendre comment le poète ou l'écrivain écrit ses vers, comment il les compose. Y a-t-il un modèle, un gabarit rythmique sur lequel il se règle ? Comment se représente-t-il les règles, le rythme ? Mais, quel que soit le procédé de création, sa

[4] Créé par l'association phonétique internationale à partir de 1886, cet alphabet est régulièrement mis à jour. Le site de l'association peut être consulté à l'adresse suivante : http://www2.arts.gla.ac.uk/IPA/ipa.html

connaissance à travers le poète n'est pas forcément nécessaire, puisque, à l'évidence, nous percevons quand même le rythme dans les vers produits, alors même que nous sommes bien incapables de les écrire. Si rythme il y a, il est donc *aussi* ailleurs que dans le processus d'écriture. Dans le cadre de ce travail, nous n'irons pas plus loin dans cette voie.

Il pourrait sembler naturel de s'appuyer pour l'analyse du rythme sur les vers déclamés. En effet, le rythme, en tant que composante du discours et de la prosodie, est spontanément associé à une exécution sonore. C'est le deuxième point de vue possible. Ce fut la voie choisie par la phonétique expérimentale et, en particulier, la position adoptée par Lote [1913], qui dans sa thèse *L'Alexandrin français d'après la phonétique expérimentale* cherche à éclairer les questions du rythme à partir de vers déclamés. Cette approche semble saine. Elle pose cependant plusieurs types de problèmes. Pour des auteurs des siècles passés, comment retrouver la prononciation de l'époque ? Comme l'écrivait ironiquement Cornulier dans *Théorie du vers* [1982b], la thèse de Lote aurait dû porter comme titre : « *L'Alexandrin français d'après la déclamation contemporaine* ». Ensuite, des difficultés tiennent aux variabilités inter et intra individuelles. Comment se fier en effet à un seul ou à un nombre limité de récitants pour construire un modèle de vers, comment espérer qu'une voix particulière puisse à elle seule incarner l'alexandrin ? Plus techniquement, quelle taille d'échantillon faudrait-il retenir pour avoir une couverture raisonnable de la déclamation actuelle de l'alexandrin ? Reste le dernier obstacle, la variabilité des alexandrins eux-mêmes. En tant qu'élément de discours, chaque alexandrin est unique : quel échantillon de vers faudrait-il dès lors retenir pour couvrir toute la diversité ? Pour le théâtre classique tel qu'il a été mis en scène au XXe siècle, on pourrait se reporter à tous les enregistrements réalisés et conservés par l'INA. Mais on rencontrerait alors d'autres problèmes techniques puisque les outils d'analyse du signal, dans leur état actuel, en raison de tous les paramétrages qu'ils exigent et de la difficulté d'interprétation, ne permettent pas de traiter des volumes aussi importants que des pièces. Ensuite se pose la question de savoir comment on pourrait abstraire à partir de toutes ces variations le modèle de l'alexandrin. Dernière objection, et non des moindres : là encore, à la simple *lecture*, nous pouvons percevoir le rythme de l'alexandrin. La forme que nous cherchons semble indépendante de sa réalisation sonore, au moins dans une certaine mesure.

Le troisième point d'entrée est celui de la réception. Comment la structure du vers est-elle perçue par les lecteurs-auditeurs ? Cornulier [1982b], dans *Théorie du vers*, a cherché à tester la « capacité métrique » des individus. Il a fait lire ou entendre à un certain nombre de personnes une version altérée des *Djinns* de Hugo : il y avait dans chaque strophe un vers faux (ayant une syllabe de plus ou de moins que les autres vers de la strophe) et les deux strophes de décasyllabes avaient été remplacées par des strophes de vers de neuf syllabes, sans césure évidente. Cornulier conclut de cette expérience que la capacité métrique, c'est-à-dire la capacité à reconnaître l'égalité du nombre de syllabes entre segments métriques voisins, est variable selon les individus, mais qu'elle est limitée à huit syllabes[5]. Dans les strophes constituées de vers de neuf syllabes, quasiment personne n'est capable d'identifier le vers faux. Il rattache cette loi des huit syllabes aux travaux de psychologie sur la mémoire et la perception, en particulier à l'article de Miller [1954, cité par [Cornulier, 1982b]] : « Le nombre magique 7 + ou - 2 », qui fixe les limites de notre capacité à traiter l'information à sept unités environ (plus ou moins deux). Cornulier pose la relativité du nombre métrique : ce n'est pas le nombre de syllabes en tant que tel qui est reconnu, mais l'équivalence de ce nombre sur plusieurs séquences : je reconnais que deux séquences successives partagent le même nombre métrique, mais je n'identifie pas forcément le nombre lui-même. Cette limite cognitive à huit peut expliquer un fait bien connu en versification française : au-delà de huit syllabes (en fait à partir de dix, puisque les vers de neuf syllabes sont rares), les vers sont constitués de deux segments métriques de longueur fixe : 6 et 6 pour l'alexandrin, 4 et 6, 6 et 4, rarement 5 et 5 pour le décasyllabe...

Ce troisième point d'entrée est très sensible à la compétence « rythmique » des individus, au niveau socio-culturel, et à la place qu'occupe le vers à une époque donnée dans une société. Nous prenons, par exemple, peu de risques en faisant l'hypothèse que la reconnaissance des vers devait être bien meilleure, à niveau social équivalent, au XVIIe siècle qu'à la fin du XXe. Ce type d'expérimentation sur la réception des vers ne va pas de soi, et requiert, pour donner des résultats fiables, la mise en place de

[5] Notons cependant que la capacité de reconnaissance de l'équivalence entre séquences va bien au-delà si les événements constituants les séquences, les syllabes dans le vers par exemple, sont regroupés dans des unités de taille intermédiaire (le pied dans les métriques quantitatives et accentuelles, l'hémistiche…).

protocoles lourds pour contrôler les biais. Mais surtout, il ne nous renseigne pas sur la forme du rythme proprement dit. Il ne présente donc d'intérêt qu'en appui de résultats obtenus autrement par une approche descriptive du rythme.

En conséquence, nous n'avons choisi, dans le cadre de ce travail, aucun de ces points d'entrée, ni la production de vers, ni leur déclamation, ni leur réception, et cela bien qu'à plus d'un titre ils auraient pu et pourraient, répétons-le, enrichir notre approche, qui est linguistique.

Déroulement

Nous commencerons par un premier chapitre consacré à une brève histoire de l'alexandrin puis nous présenterons les théories ou les travaux qui, à un titre ou à un autre, ont eu une influence sur notre démarche empirique. Il ne s'agit en aucun cas d'une recension critique de toutes les recherches menées dans le domaine qui nous occupe : ce travail d'érudition a déjà été réalisé à maintes reprises par d'autres. Il s'agira plutôt ici d'une présentation des « échafaudages » de notre analyse.

Nous présenterons ensuite dans le chapitre II le corpus de vers numérisés que nous avons utilisé, constitué par les œuvres théâtrales complètes de Corneille et Racine. Celles-ci ont été « éditées » sous une forme propice à l'analyse automatique du vers et de la rime. Nous aborderons dans le chapitre III la question de la rime, qui est partiellement autonome. Nous traiterons ensuite (chapitre IV) le problème de l'identification des positions métriques et du dénombrement. La diérèse, le *e* muet[6] et la liaison occuperont une place de choix. Nous exposerons alors (chapitre V) les marquages que nous avons adoptés pour décrire la figure rythmique du vers, avant de nous consacrer à la description du métromètre (chapitre VI). Nous explorerons (chapitre VII) les différentes composantes phonétiques, morpho-syntaxiques et prosodiques du rythme de l'alexandrin et nous les illustrerons par des représentations graphiques. Enfin, pour finir, nous mettrons au jour des liens entre la structure rythmique des vers et les univers lexico-sémantiques (chapitre VIII).

---

[6] Dans les différents ouvrages consacrés au vers, devant la lettre *e*, le pronom est tantôt sous sa forme complète, « le *e* muet », tantôt sous sa forme élidée, « l'*e* muet ». Nous avons opté pour la première solution, bien que l'interdiction du hiatus dans le vers eût dû nous faire préférer la seconde.

I. DONNEES HISTORIQUES ET THEORIQUES

Nous voulons décrire le rythme de l'alexandrin. On a vu dans l'introduction que la notion de rythme n'allait pas de soi. Celle de vers est moins difficile à cerner, mais il faut admettre cependant qu'elle est historiquement marquée. L'alexandrin classique est différent de celui des origines et de celui du XIXe siècle. Les vers évoluent, tout comme les modèles qui les décrivent. Par ailleurs, les vers ne sont pas toujours définis par les mêmes critères. Tantôt le vers est défini par des pieds où alternent des syllabes marquées (quantité ou accent) et des syllabes non marquées, tantôt il est défini par un nombre de syllabes. Parfois, plusieurs critères servent à le définir : nombre de syllabes et position des accents par exemple. Enfin, les notions de rythme et de mètre et les rapports que ces notions entretiennent avec la langue ne sont pas toujours clairs. Nous évoquerons quelques uns des facteurs qui rendent l'existence d'un modèle de vers problématique.

Nous commencerons ce chapitre par une brève histoire de l'alexandrin, depuis ses origines jusqu'à l'époque classique, occasion pour nous de nommer et de décrire quelques phénomènes métriques, en particulier ceux de la césure. Nous dresserons alors un tableau des types de versification, et rappellerons quelques-unes des positions adoptées par les métriciens pour décrire les aspects formels de l'alexandrin.

Nous examinerons ensuite l'apport de la métrique générative américaine à l'étude du vers, ainsi que celui d'autres théories plus actuelles. Enfin, nous exposerons certains concepts de la théorie du rythme développée par Lusson et Roubaud, et nous montrerons leur application à l'analyse de l'alexandrin.

Ces différentes considérations historiques et théoriques nous amèneront à présenter le projet du métromètre, dont la teneur sera esquissée en fin de chapitre.

Il nous a paru inutile de mener dans le cadre de ce travail une recension critique des travaux existant sur le vers. Nous renvoyons à Henri

Meschonnic [1982] qui présente un certain nombre de théories sur le vers, qui ne se limitent pas au vers français, et à Jean-Michel Gouvard [1996] qui donne quelques éléments sur l'histoire de la théorie du vers. Leurs analyses ont le mérite de couvrir la grande majorité des travaux sur le vers. Dans ce chapitre, nous ne donnons de place qu'aux articles et ouvrages qui nous ont été les plus utiles dans l'élaboration de notre approche empirique.

I.1. BREVE HISTOIRE DE L'ALEXANDRIN

Reprenons quelques éléments de l'histoire de l'alexandrin proposés par Roubaud[7]. Celui-ci considère comme alexandrins classiques les alexandrins composés entre 1630 et 1830 [1986, p. 49]. Certes, avant comme après, on a écrit des alexandrins conformes au modèle classique, mais c'est la seule période où *tous* les alexandrins produits respectent les mêmes contraintes formelles. Lors d'un cours consacré au vers alexandrin (1992-1993), Roubaud restreindra cette période aux années 1630-1789. Le tracé des frontières n'est pas toujours aisé, ce qui témoigne de la difficulté à reconnaître l'identité d'une forme.

La biographie du vers alexandrin fait sans cesse référence aux vers concurrents, parce que « la destinée de l'alexandrin est solidaire de l'ensemble de la versification française » [Roubaud, 1977, p. 5]. L'évolution de la césure dans les autres mètres a eu une incidence sur celle de

[7] Cette biographie de l'alexandrin est organisée par Roubaud, pour les besoins du récit et de l'objet, en douze branches. Les branches I à IV, dont le récit figure dans un article publié dans les *Cahiers de Poétique Comparée* (« Une vie : synopsis en douze branches. Branches I à IV : Des origines à la "Cléopâtre captive" (1553) » [Roubaud, 1977]) « racontent » la naissance du premier alexandrin, son succès, son déclin accompagné de sa théorisation et enfin sa renaissance sous une forme nouvelle au milieu du XVIe siècle. Roubaud a mis à contribution pour cette première partie de l'histoire de l'alexandrin un article de Kund Togeby [1968].

Les branches IX à XII, qui couvrent la période 1850-1975, sont racontées dans *La Vieillesse d'Alexandre* publié chez Maspéro en 1978. Quant aux branches centrales, de 1550 à 1850, celles de l'alexandrin préclassique, classique et romantique, « l'alexandrin ordinaire », un certain nombre d'éléments métriques et rythmiques en sont présentés dans un article de *Langue Française* [Lusson et Roubaud, 1974], dans *La Vieillesse* [*1978*, 1988, p. 98-103], mais surtout dans les deux articles consacrés au programme Dynastie : « DYNASTIE : études sur le vers français, sur l'alexandrin classique » [Roubaud, 1986 et 1988].

l'alexandrin, ce qui nous amène à définir les types de césures qui sont une des composantes de la signature du vers : la césure épique, lyrique, italienne et classique.

Dans le vers à ***césure épique***, ou vers épique, on admet une syllabe surnuméraire (non comptée) après la dernière voyelle accentuée du premier segment métrique. Dans le vers suivant, le *-re* d'*estoire* ne s'élide pas car il est suivi d'une consonne, mais il n'entre pas dans le décompte du vers car il se trouve juste après la dernière voyelle marquée du premier hémistiche :

> Qui vers de riche estoi**re** veut entendre et oïr,
> [Si vous voulez entendre une belle histoire en vers]
> Alexandre de Paris, *Le Roman d'Alexandre*, p. 70.

Dans le vers à ***césure lyrique***, ou vers lyrique, une syllabe post-tonique comptée (en français celle qui contient un *e* muet final) occupe la dernière position du premier segment métrique, par exemple la quatrième position pour le décasyllabe :

> Margueri**te** qui le dragon d'envie (...)
> Et Christi**ne** fut de roial lignie,
> Eustache Deschamps, in [Roubaud, 1998, p. 59].

Le vers à ***césure italienne***, l'hendécasyllabe italien ou espagnol par exemple, admet une position post-tonique après la césure qui entre dans le décompte. Ainsi, dans le vers suivant la césure se situe entre *cier-* et *-zo*, *-zo* occupant une position métrique. La césure est clairement définie par la position de l'accent de mot (sur *cier-*) :

> Antes que el cierzo de la edad ligera
> (a n)(t ẹ s)(k ẹ l)(θ j ẹ r) | (θ o)(d e)(l e)(d a d)(l i)(x e)(r a)
> Avant que le vent froid du temps qui vole
> Lope de Vega, sonnet, vers 1, in Anthologie bilingue de la poésie espagnole, p. 426.

Ce type de césure est extrêmement rare dans le vers français.

Le vers à ***césure classique*** se définit par le refus de ces trois types de césures. Autrement dit, dans l'alexandrin classique, on ne peut avoir de *e* muet non élidé et non compté après la sixième position (comme dans le vers épique), ni de *e* muet en sixième position compté (comme dans le vers lyrique), ni de *e* muet en septième position (comme dans le vers à l'italienne)[8].

---

[8] Dominicy [1992, p. 163] emploie le terme de césure synthétique pour les vers qui admettent des césures lyriques, épiques ou classiques, tandis que la césure est pour lui de

Reprenons le récit depuis les origines en nous appuyant sur l'article de Roubaud [1977]. La présentation que nous en faisons est extrêmement sommaire et ne rend pas compte de la richesse des recherches et des débats sur le sujet. Le vers alexandrin apparaît pour la première fois dans le « Vers del juïse », Poème du jugement, autour de 1130. On trouve également des alexandrins dans le « Pèlerinage de Charlemagne » daté de 1160. A l'époque où la rime alexandrine fait son apparition, les vers les plus utilisés sont : le décasyllabe épique en rimes assonancées employé principalement dans la chanson de geste, l'octosyllabe monorime (présent surtout dans la poésie religieuse) et l'octosyllabe à rimes plates employé dans le « roman ».

En 1170, Lambert le Tort de Chateaudun écrit l'histoire d'Alexandre le Grand en vers de douze syllabes. Ce texte marque le début de la grande période du premier alexandrin. Jusqu'au milieu du XIVe siècle, les exploits d'Alexandre rapportés en alexandrins se multiplient et circulent en Europe. C'est de cet usage que vient le nom du vers, mais il ne sera baptisé comme tel que trois siècles plus tard.

Dans ces poèmes, le vers alexandrin est composé de deux segments métriques de six syllabes autonomes. Le vers est à césure épique : il peut y avoir un *e* muet qui ne compte pas après la sixième position, quel que soit le début du segment métrique suivant.

> Oncques en icest siecle ne nasqui si bons Gris.
>
> Alexandre de Paris, *Le roman d'Alexandre*, p. 102.

Il y a une concordance très forte entre les segments métriques et la syntaxe. Les vers sont organisés en laisses ou strophes de longueur variable avec la même rime. Ces trois caractéristiques (césure épique, concordance forte entre syntaxe et métrique, laisses ou strophes monorimées) apparentent le premier alexandrin au décasyllabe épique. On trouve également des alexandrins dans la chanson de geste et dans la poésie strophique religieuse.

Pendant le dernier quart du XIIe siècle, le vers en général évolue : la rime remplace peu à peu l'assonance et le décasyllabe lyrique fait son apparition ainsi que les strophes.

---

type analytique dans le vers à césure italienne, qui admet un décalage entre la dernière voyelle du premier segment métrique et la dernière voyelle métrique du mot traversé par la césure.

Au début du XVe siècle, l'alexandrin est complètement tombé en désuétude : « le naufrage de l'alexandrin médiéval est celui de l'alexandrin épique » [1977, p. 32]. C'est curieusement à cette époque que les traités commencent à en parler et qu'il reçoit, en fait, son nom d'alexandrin. Le vers alexandrin est décrit comme étant de douze ou treize syllabes (on pourrait rajouter quatorze dans les cas de césure épique).

Ce déclin de l'alexandrin et du décasyllabe épiques s'est fait au profit de l'octosyllabe mais surtout du décasyllabe lyrique : la césure lyrique fait son entrée. Au cours du XVe siècle, l'évolution du décasyllabe se poursuit. La contradiction entre les deux types de décasyllabes, liée à l'incompatibilité entre leurs césures épiques et lyriques conduit peu à peu à un affaiblissement des deux types de césure. Dans le même temps, l'octosyllabe recule, et surtout l'octosyllabe à rimes plates. Roubaud pose comme hypothèse que ce recul de la « platitude » dans l'octosyllabe a permis que l'alexandrin et le décasyllabe s'emparent des rimes plates (très rarement attestées jusque-là, parce qu'elles étaient associées à l'octosyllabe).

De 1510 à 1530, la césure qui sera celle du décasyllabe et de l'alexandrin classiques français s'impose dans les faits. Les traités la définissent de plus en plus souvent comme norme. Cette césure n'est ni lyrique, ni épique, ni italienne. Elle impose une concordance forte entre segment métrique et syntaxique et la position avant la césure est toujours la plus marquée.

Jean Robertet, à la fin du XVe siècle, annonce le renouveau de l'alexandrin en composant des poèmes en alexandrins à rimes plates. Jean Lemaire, au début du XVIe siècle, emploie l'alexandrin dans un passage de « La Concorde des deux langages » : les vers sont à rimes plates et il s'abstient de toute césure épique. Après Lemaire, la césure se normalise, atteint l'état qu'elle conservera dans le vers classique et romantique, mais elle met du temps à s'imposer puisque c'est la « Cléopâtre captive » d'Étienne Jodelle en 1553 qui constitue le coup d'envoi du nouvel alexandrin, qui n'admet ni césure épique, ni césure lyrique. Il est à la fois un et deux segments métriques.

> Rythmiquement, l'alexandrin épique (comme le décasyllabe épique) n'est que concaténation de segments autonomes ; l'unité vers n'est assurée que par la rime (d'abord assonance) très longuement monotonement répétée. Ce qui se passe dans les quatre siècles à peu près qui séparent Lambert de Tort de Jodelle, c'est le passage à un tout autre vers où chaque segment (hémistiche) a son identité

> rythmique, mais où le vers tout entier est aussi une unité rythmique et métrique gouvernée par les mêmes lois, de même nature.
>
> [Roubaud, 1977, p. 84].

Le vers sur lequel porte notre travail est ce second alexandrin constitué à la fois d'un et de deux segments métriques : un vers de douze syllabes avec une césure aux propriétés particulières et un vers toujours relié à d'autres par la rime. Le vers classique est cependant différent de l'alexandrin de Jodelle : peu à peu s'impose la règle de l'interdiction du hiatus puis la règle des trois consonnes vient modifier le décompte des syllabes.

I.2. LE VERS FRANÇAIS : SYLLABIQUE OU SYLLABO-TONIQUE ?

S'il existe un accord pour décrire l'alexandrin comme un vers de douze syllabes avec un marquage fort en sixième et douzième positions, c'est là que finit le consensus. L'histoire de la métrique de l'alexandrin a été marquée et reste encore marquée par un vieux débat : existe-t-il ou non des « accents de vers » dans l'alexandrin ? Nous commencerons par décrire les grandes familles de versification définies par les métriciens avant de voir comment ceux-ci ont classé l'alexandrin.

I.2.1. Versification quantitative, accentuelle ou syllabique

En général, différents types de versification sont distingués : la métrique quantitative (grecque et latine), la métrique accentuelle et la métrique syllabique.

Le vers *quantitatif* est une séquence constituée d'un nombre fixe de pieds de même nature : le tétramètre est constitué de quatre pieds identiques ou équivalents, le pentamètre de cinq, l'hexamètre de six... Le pied, ou mètre, est une unité de mesure composée d'au moins deux syllabes : l'ïambe, le trochée, le spondée sont des pieds de deux syllabes, l'anapeste et le dactyle de trois, le choriambe de quatre... Dans chacun de ces pieds, les syllabes se différencient par leur longueur ou quantité (longues ou brèves) et sont agencées selon un ordre spécifique (par exemple, syllabe brève suivie

d'une longue dans le pied ïambique, deux brèves suivies d'une longue dans le pied anapestique). La syllabe ne peut être mobilisée comme unité de mesure parce qu'en dehors du fait qu'un pied peut être composé d'un nombre variable de syllabes, il existe des équivalences qui permettent parfois de considérer deux syllabes brèves comme équivalentes d'une longue. Ainsi, dans un pied ïambique (∪–), la syllabe longue peut être remplacée par deux syllabes brèves (∪∪) [Elwert, 1965, p. 14] pour donner un pied de trois syllabes (∪∪∪).

Le vers *accentuel* est constitué par la concaténation d'un nombre fixe de pieds équivalents, comme précédemment. Le pied est lui-même défini comme une séquence d'un nombre défini de syllabes. Ce qui différencie les syllabes n'est plus la quantité ou la durée, mais l'intensité ou l'accent. Le pentamètre ïambique anglais est constitué de cinq pieds de type WS (non accentué, *Weak* ; accentué, *Strong*) : WSWSWSWSWS.

When forty winters shall beseige thy brow,

When	for-	ty	win-	ters	shall	be	siege	thy	brow
W	S	W	S	W	S	W	S	W	S

Shakespeare, *Sonnet*, II.

Comme dans le vers quantitatif, la syllabe ne peut être considérée comme l'unité métrique. En effet, dans le pentamètre ïambique anglais, il arrive que la première syllabe soit absente et qu'il y ait deux syllabes surnuméraires après la dixième.

Aussi même les générativistes ont-ils renoncé à considérer que le pentamètre était constitué d'une séquence de syllabes. Bien que ce ne soit pas la même propriété qui différencie les syllabes dans la métrique quantitative et accentuelle (durée d'un côté, accent de l'autre), les mêmes termes sont employés pour désigner les pieds, par analogie : un pied ïambique est constitué d'une brève et d'une longue dans la métrique quantitative, d'une syllabe non accentuée suivie d'une syllabe accentuée dans la métrique accentuelle. Ces deux premières formes de vers sont donc assez homologues.

Le vers *syllabique*, qui caractérise les langues romanes faiblement accentuées et le japonais, est fondé sur un nombre fixe de syllabes par vers. Il en est ainsi du vers français. C'est donc de manière abusive que l'on parle de *pied* pour le vers roman. Même si l'on peut trouver dans l'organisation des syllabes des structures apparentées aux pieds, elles ne constituent pas le

squelette métrique du vers. L'unité de mesure du vers français, italien ou espagnol est la syllabe.

Les différences semblent assez irréductibles entre différents types de versifications, surtout entre les métriques quantitatives et accentuelles (où durée et accent jouent le même rôle), d'une part, et la métrique syllabique, de l'autre[9]. Même les appellations distinguent clairement les types de versification. La désignation des vers quantitatifs ou accentuels (vers antiques, vers russes et anglais, etc.) intègre la notion de pied et non pas celle de syllabe : le pentamètre, l'hexamètre sont respectivement des vers de cinq et six pieds tandis que la désignation des vers syllabiques (français, italiens et espagnols) indique un nombre de syllabes : le décasyllabe est un vers de dix syllabes, l'hendécasyllabe de onze[10]. Il paraît difficile d'avoir un dispositif universel de description de tous les vers.

I.2.2. L'accent dans le vers français ?

Cependant, les frontières ne sont pas si nettes entre types de métriques, en ce sens qu'un même vers peut par exemple recevoir une interprétation accentuelle ou syllabique. Il en est ainsi du pentamètre ïambique qui a pu être interprété comme un décasyllabe par la métrique générative. D'après Meschonnic [1982, p. 237], l'hendécasyllabe italien du XVIe pourrait être interprété comme pentamètre ïambique avec quelques irrégularités dans la réalisation. Ceci ne résiste pas à une définition précise du vers ïambique et à une observation systématique d'un grand corpus de vers. En introduisant la notion de vers syllabo-tonique, la métrique tente encore d'atténuer les

[9] Pour expliquer l'origine de ces différences, on peut recourir à l'hypothèse de Roubaud selon laquelle « la poésie [serait] mémoire de la langue ». Selon lui, « La poésie métrique (d'une langue) imite l'organisation des sons (de la langue). Le moyen de codage que choisit la langue est une mimique physique, une abstraction simplificatrice des traits les plus frappants des unités phoniques intervenant dans les énoncés » [Roubaud, 1969]. Les différences de métriques refléteraient les différences d'organisation des langues dans lesquelles s'inscrivent ces vers : à langue accentuée, vers accentué...

[10] Les appellations ne sont cependant pas de bonnes descriptions du vers dans la mesure où des vers équivalents peuvent avoir des dénominations distinctes. Par exemple, l'hendécasyllabe italien ou espagnol peut avoir 10, 11 ou 12 syllabes mais admet toujours une position accentuée en position 10. Le décasyllabe français admet 10 ou 11 syllabes (si la rime est féminine) avec la position 10 toujours marquée. Malgré les différences de noms, l'hendécasyllabe italien ou espagnol est équivalent rythmiquement au décasyllabe français.

frontières entre catégories et de rapprocher le vers syllabique du vers accentuel : le vers est défini par un nombre de syllabes et par des accents métriques sur des positions fixes.

Le fait de ne pouvoir donner une définition universelle de la versification n'a pas toujours été admis. Ainsi, Gasparov [1987] rappelle que les métriciens allemands tentaient de faire entrer l'alexandrin dans un modèle syllabo-tonique ou accentuel, de type ïambique, apparenté à celui de leurs vers. Les métriciens français s'en défendaient. Pourtant, même en France, le débat sur le caractère accentuel de l'alexandrin a été et reste animé. Il pose la question de l'existence d'accents dans le vers, et plus largement, il nous renvoie à la question de l'accent en français.

Au XVe siècle quand le vers alexandrin est nommé et défini dans les traités [Roubaud, 1977], il est présenté comme un vers de douze ou treize syllabes. Au XVIe siècle, la notion de coupe vient compléter la description syllabique. La différence avec la métrique gréco-latine était clairement soulignée au XVIIe siècle : la seule spécificité du vers français était d'être constitué d'un nombre fixe de syllabes et d'être marqué par une césure.

Ce n'est qu'au XIXe siècle que l'accent deviendra pour les métriciens un élément constitutif du vers. D'après Gouvard [1996, p. 231], l'ouvrage de Scoppa, paru en 1811, aurait joué un rôle important dans l'introduction de l'accent comme composante descriptive du vers. Pour Scoppa, un vers est constitué de la concaténation d'un nombre variable de pieds identiques. Cette théorie résiste mal à la confrontation avec des vers, mais elle a remporté un vif succès, sans doute parce qu'elle a permis de rompre l'isolement dans lequel était enfermé le vers français, puisqu'il était un des seuls vers européens à être jusque-là exclusivement défini par un nombre de syllabes et la place de la césure. Dans la recherche d'un mètre universel, la particularité du vers français était gênante. En introduisant la notion d'accent dans la définition métrique du vers français, on rapproche celui-ci des autres vers.

Dans la première moitié du XIXe siècle, Quicherat renouvelle la théorie du vers : il affirme qu'il existe dans l'alexandrin deux accents fixes en sixième et douzième positions et deux accents mobiles sur les positions 1 à 5 et 7 à 11. L'alexandrin est alors conçu comme un tétramètre à coupure variable. Cette théorie a eu un grand succès : on la retrouve chez Grammont

[*1908*, 1965], qui considère que l'alexandrin est « un vers à quatre mesures, c'est-à-dire contenant quatre éléments rythmiques, terminés chacun par un accent d'intensité ; le deuxième et le quatrième sont fixes sur la sixième et la douzième syllabes, et les deux autres sont variables dans l'intérieur d'un même hémistiche » [*1908*, 1965, p. 50]. Comme le rythme impose, selon Grammont, le retour à intervalles égaux des accents, il faut accélérer ou ralentir le débit pour donner le sentiment de l'isochronie. Les variations de débit permettent de contrecarrer la mobilité des accents dans le but de produire un édifice rythmique régulier. On retrouve la trace de ces théories chez Mazaleyrat [1974, p. 14], pour qui le sentiment du rythme est lié « au retour des accents de groupe à des intervalles proportionnels et perceptibles », ces accents de groupe étant définis en discours. Le mètre repose sur la disposition ordonnée des accents de langue, qui définissent des césures et des coupes et délimitent des groupes rythmiques spécifiques, différents de ceux que produit la prose : « un vers (ou mètre) est composé d'abord d'un système de mesures rythmiques fondé sur une série de rapports perceptibles des parties entre elles et des parties au tout » [Mazaleyrat, 1974, p. 16].

Plus récemment, dans *Dire le vers*, Jean-Claude Milner et François Regnault [1987] reprennent la théorie de Quicherat :

> - Il y a deux accents de vers obligatoires et fixes à la fin de chaque hémistiche : sixième et douzième syllabes ;
> - Il y a deux accents de vers obligatoires et mobiles, à raison d'un par hémistiche.
>
> [Milner et Regnault, 1987, p. 113].

Milner et Regnault supposent l'existence d'accents de vers, doués d'une certaine autonomie par rapport à l'accent de langue. En langue, pour Milner et Regnault, il y a un accent à la fin de chaque mot phonologique et un contre-accent de moindre intensité en début de mot phonologique. La position des deux accents mobiles du vers (un par hémistiche) se déduit de la confrontation avec les accents de langue, c'est-à-dire avec les accents de fin de mot phonologique et les contre-accents : « l'accent de vers rencontre toujours un accent de langue » [Milner et Regnault, 1987, p. 113]. Dans le vers suivant,

> Ma fortune va prendr(e) une face nouvell(e)
>
> Racine, *Andromaque*, vers 2

les accents fixes portent sur la sixième (*prendre*) et la douzième (*-velle*) positions, et les accents mobiles sur les positions 3 (*-tu-*) et 9 (*fa-*). Le

premier hémistiche est composé de deux mots phonologiques, l'accent mobile est donc situé sur la fin du premier mot (*Ma fortune*). Il aurait pu l'être sur le premier (*ma*) ou le second (*va*) contre-accent, mais hiérarchiquement, l'accent de mot phonologique l'emporte sur le contre-accent. Nous renvoyons au chapitre 7 de *Dire le vers* [Milner et Regnault, 1987] pour les détails sur l'attribution de l'accent au cas par cas. Rappelons qu'il s'agit d'un traité de diction, qui propose des règles pour *bien* dire le vers. Dans un autre article de la même époque, Milner [1987] reprend la problématique de l'accent de vers par rapport à l'accent de langue. Il pose : qu'il y a « un marquage accentuel autonome dans le vers classique français » [Milner, 1987, p. 33].

Milner démontre par l'absurde son théorème. Pour cela, il lui faut prouver la fausseté du contre-théorème, à savoir la coïncidence systématique entre accent de langue et accent de vers. En admettant le théorème de Quicherat, il met en évidence de nombreux contre-exemples et par conséquent l'autonomie du marquage accentuel dans le vers.

Reste une question : peut-on admettre le théorème de Quicherat ? La « démonstration » de ce théorème a été faite au préalable, à l'aide de plusieurs arguments indirects : par l'absence de *e* muets en sixième et douzième positions, sauf dans les cas rarissimes où il est accentué, et par l'existence d'une échelle d'évaluation des alexandrins qui rend les hémistiches constitués de deux mots phonologiques plus réguliers que les autres. Il ne s'agit pas, en fait, d'une démonstration. Comme le dit Milner lui-même : « le théorème de Quicherat est plausible » [p. 43].

A ces thèses accentuelles, on peut rapidement opposer la position de Meschonnic et celle de Cornulier.

Meschonnic [1982] propose une distinction entre mètre et rythme. Le mètre est le lieu de codifications. Le mètre alexandrin est défini par le nombre de syllabes et par la présence d'un accent en sixième et douzième positions. Les autres accents relèvent du rythme du discours et ne font pas partie du mètre :

> Comme la seule règle métrique, pour l'alexandrin régulier, concerne l'accent à la 6e et à la 12 e position, les autres accents sont rythmiques, non métriques.
>
> [Meschonnic, 1982, p. 229].

> Il n'y a pas de métrique en français, au sens où l'accent y est une notion rythmique, non métrique. (...) La métrique s'y réduit à la position et aux règles

> (prosodiques, syntagmatiques) de la césure. (...) Dans toutes les autres positions, l'accent est linguistique. Sa prédictibilité y est de l'ordre du discours, non de la métrique.
>
> [Meschonnic, 1982, p. 230].

Meschonnic reproche à Mazaleyrat d'avoir confondu mètre et rythme. Les groupes rythmiques sont des effets de discours et n'entrent pas dans la définition du mètre.

Tandis que Meschonnic conserve la notion d'accent, Cornulier défend une théorie syllabique du vers, qui refuse l'idée d'accent. Pour lui, l'alexandrin est constitué de douze syllabes avec une césure, définie par des règles syntagmatiques. Dans cette logique, le rôle de l'accent est illusoire ou secondaire :

> Il est inutile et redondant d'ajouter à la théorie de la métrique française une loi d'accentuation finale qu'on doit poser pour le français en général.
>
> [Cornulier, 1982b, p. 286].

L'effet de l'accent serait « une conséquence automatique du fait que les hémistiches et vers sont des groupes rythmiques ou sont des syntagmes » [Cornulier, cité par Meschonnic, 1982, p. 560].

Il existe, en effet, une concordance assez nette entre les syntagmes et les accents linguistiques, surtout pour les hémistiches de l'alexandrin classique où il n'y a jamais de décalage entre fin de syntagme et accent. En revanche, pour l'étude du rythme du discours à l'intérieur des hémistiches, la notion d'accent peut paraître plus immédiate que celle de fin de syntagme à cause du problème crucial des *e* muets finaux non élidés. Prenons un exemple :

> Oui, je l'aime, seigneur, et n'en fais point d'excuse ;
>
> Corneille, *Polyeucte*, vers 461.

Si l'on note les accents, il y en a un en position 3 sur **ai**me. Si l'on note les fins de syntagmes, il y en a une en position 4 sur ai**me**, puisque le *e* muet n'est pas élidé. On a un décalage d'une position et l'on admettra sans trop de difficultés que le marquage de la position 3 est plus pertinent par rapport à des questions de rythme que celui de la position 4. Dans la pratique, en général, quand un marquage de fin syntagme est utilisé, le marquage porte sur la dernière syllabe qui n'est pas occupée par un *e* muet.

Mais le rythme n'est pas l'objet d'étude de Cornulier. Rejetant la notion d'accent, il s'intéresse principalement à l'étude de phénomènes morpho-syntaxiques sur les positions métriques.

Pour Lusson et Roubaud [1974], le vers alexandrin est conçu comme la concaténation de deux segments métriques de six positions. Il existe de plus à l'intérieur des hémistiches des régularités qui constituent des pré-rythmes. Autrement dit, dans sa structure profonde, l'hémistiche admet deux pré-rythmes : l'un de type 001001 (séquence de deux groupements 001) et l'autre de type 010101 (séquence de trois groupements 01). Les hémistiches réels sont plus ou moins compatibles avec l'un de ces deux modèles. Les séquences 001001 et 010101 peuvent toutes deux être interprétées comme des mètres dans cette théorie. En effet, Lusson considère comme métrique une séquence qui au moins à un niveau s'analyse comme la concaténation de groupements identiques à un groupement générateur (ici 001 ou 01).

I.2.3. Une tentative de résolution empirique

La question de l'accent de vers dans l'alexandrin est donc le lieu d'âpres débats tout comme celle de l'accent dans la langue. L'analyse empirique des vers peut-elle aider à prendre position sur cette question ? Gasparov [1987] a tenté d'apporter des éléments de réponse à la question : l'alexandrin est-il syllabique ou syllabo-tonique ? Il propose pour cela une méthode probabiliste qu'il applique à tous types de vers russes, anglais, italiens, français... L'idée générale est de comparer le rythme de la langue à celui des vers. Si les rythmes sont très différents, on en déduira une spécificité rythmique du vers.

Gasparov rappelle que l'usage de la statistique pour l'analyse métrique date du XXe siècle et qu'il a d'abord consisté à compter les violations du mètre (dans le cas du vers ïambique, par exemple, à repérer les positions marquées métriquement non occupées par un accent de mot, ou les positions non marquées occupées par une position accentuée) et à comparer le rythme des poètes entre eux.

Ensuite, quelques métriciens ont cherché à comparer le rythme des vers avec celui de la langue, ce qui permet de différencier ce qui dans le vers provient de la langue et ce qui vient de la « volonté créatrice du poète ». Il y aurait dans le vers deux composantes rythmiques : l'une provenant de la langue, l'autre du vers.

Dans cette approche, deux types de modèles sont utilisés, soit un modèle de discours, soit un modèle de langue. Le premier modèle (speech

probability model ou modèle de discours) est construit en extrayant de corpus en prose tous les fragments qui coïncident avec les règles minimales d'un mètre donné (longueur, accent). Cet ensemble de « vers » constitue le « rythme spontané du discours du mètre étudié » [Gasparov, 1987, p. 323]. Et si l'on compare les vers attestés avec ce modèle, les différences sont interprétées comme des particularités du rythme des vers en question. Gasparov considère que cette approche est limitée parce que le rythme des vers est comparé à celui du discours et non pas à celui de la langue en général.

Les métriciens russes, et en particulier Gasparov, préfèrent recourir à un modèle de langue. La démarche adoptée est la suivante :

1. A partir d'un corpus de prose, on identifie les mots phonologiques (mots pleins accompagnés de leurs enclitiques et proclitiques), puis on calcule la proportion des différents types rythmiques de mots phonologiques : les monosyllabes, dissyllabes oxytons[11], dissyllabes paroxytons, trisyllabes oxytons...
2. On établit la liste de toutes les combinaisons de mots phonologiques qui respectent une longueur de vers donnée.
3. On calcule la probabilité d'une combinaison donnée en multipliant les probabilités d'apparition de chaque mot phonologique (le modèle de langue implique l'indépendance rythmique des mots phonologiques).
4. On calcule la probabilité de chacune des combinaisons. Le total est inférieur à 1 et donne la probabilité d'avoir un mètre donné en prose.
5. On compare la fréquence théorique d'une combinaison (hors combinaisons non métriques) avec la fréquence d'occurrence dans des textes de poètes. Si la différence est significative, il y a un écart par rapport au rythme naturel de la langue.
6. On peut aussi calculer le profil accentuel de l'ensemble des combinaisons, et cette approche est plus fiable. Comme il y a globalement beaucoup plus d'accents dans le vers que dans le modèle, on procède à un redressement.

Gasparov a appliqué sa méthode à l'alexandrin, en considérant comme unité métrique l'hémistiche.

[11] Oxyton : accentué sur la dernière syllabe ; paroxyton : sur l'avant-dernière.

Le vocabulaire rythmique du français a été construit à partir d'échantillons de 1 000 mots tirés des préfaces de Racine à ses tragédies, du *Malade Imaginaire* de Molière, de *Candide* de Voltaire et du *Colonel Chabert* de Balzac, soit en tout 4 278 mots phonologiques. Le corpus de vers est constitué des 600 à 1 000 premiers vers des sonnets de Du Bellay[12], de *Mithridate*, des *Bucoliques* et des *Élégies* de Chénier et des *Fleurs du Mal* de Baudelaire.

Gasparov montre que le vers français, du moins dans son échantillon, ne s'éloigne en rien du modèle syllabique. Pour lui, « la structure accentuelle de chaque hémistiche est entièrement déterminée par le vocabulaire rythmique de la langue française et rien d'autre » [Gasparov, 1987, p. 334].

Il n'y a pas de tendance syllabo-tonique dans l'alexandrin. Ce résultat vient de la confrontation de la séquence des poids des vers attestés avec la séquence des poids du modèle de langue. Effectivement l'allure de la courbe du vers ressemble à celle du modèle, même si Gasparov n'a pas fait de test statistique. Il dit que l'écart est non significatif sans en donner les preuves.

Si l'on s'en tient à la figure globale, le vers français ne serait pas syllabo-tonique. Mais lorsqu'il compare les séquences d'hémistiches, comme il le fait sans en tirer toutes les conclusions, il montre une nette préférence pour deux organisations rythmiques (001001 et 010101) par rapport à des organisations « mixtes » (*cf.* Tableau 1 ci-dessous).

	modèle de langue	mille vers de *Mithridate*
Premier hémistiche		
Ïambique	45	52
Anapestique	35	35
Mixte	20	13
Second hémistiche		
Ïambique	48	43
Anapestique	37	47
Mixte	15	10

Ïambique : 010001, 000101, 100101, 010101, 110001, 010011
Anapestique : 001001, 101001, 001011
Mixte : 000001, 100001, 000011, 100011, 011001, 001101, 000111

Tableau 1. Répartition des types d'hémistiches [Gasparov, 1987, table 19, p. 356]

[12] Le texte de l'article indique Rémi Belleau, mais c'est le nom de Du Bellay qui apparaît dans toutes les annexes.

La limite que Gasparov reconnaît à son approche tient aux choix d'accentuation en prose et dans le vers, qu'il faudra affiner.

Gasparov, en particulier, ne semble pas avoir pris en compte les problèmes de frontière de mots phonologiques dans le vers, qui renvoient à la question des *e* de fins de mots phonologiques. Ceux-ci comptent ou non selon l'environnement qui les suit. Les combinaisons de mots phonologiques de langue pour arriver à six syllabes varient selon leur agencement... Par exemple dans le vers suivant, *dans Athènes* est un trisyllabe en fin de vers, mais il serait un mot phonologique de quatre syllabes s'il était placé devant un mot à initiale consonantique.

Un poison que Médée apporta dans Athènes

Racine, *Phèdre*, vers 1638.

Enfin et surtout, alors qu'il prétend avoir prouvé que l'alexandrin n'était pas syllabo-tonique, il présente des tableaux qui montrent que la répartition des accents linguistiques dans les hémistiches n'est pas conforme au modèle de langue, et que cette répartition privilégie des figures « métriques »[13], au sens que donne Lusson à ce terme : concaténation de groupements de même structure (deux fois 001 ou trois fois 01).

Admettons que le vers n'est que syllabique, avec deux positions marquées en 6 et 12. Gasparov a cependant montré involontairement qu'il existe des répartitions spécifiques des accents sur les hémistiches, différentes du modèle de langue.

De ce rapide parcours historique et théorique, retenons que :

- le mètre alexandrin est constitué de douze positions métriques ;
- le vers est à la fois constitué d'un segment métrique de douze positions et de deux de six positions. La fin de ces deux segments est marquée par un accent ;
- le vers n'est pas isolé, il fait partie d'une structure d'un niveau supérieur, marquée notamment par la rime ;
- les autres phénomènes accentuels observés sont des effets du rythme, et ne relèvent pas de la métrique ;

[13] Nous ne pouvons établir de test statistique, car Gasparov n'indique pas les effectifs de « vers », construits à partir des mots phonologiques de langue.

- pourtant le rythme du vers semble plus régulier que celui de la prose (dans des contextes avec un nombre de syllabe identique).

I.3. LA METRIQUE GENERATIVE AMERICAINE ET SES SUITES

La métrique générative s'est principalement intéressée au pentamètre ïambique anglais. Elle n'a donc qu'un lointain rapport avec l'objet qui nous occupe, l'alexandrin classique. Cependant, la théorie a une visée universelle et doit pouvoir s'appliquer aussi à d'autres vers[14]. De plus, quelles que soient les critiques qui ont pu lui être adressées, elle a eu le mérite de renouveler considérablement la description du vers et de poser les premiers jalons nécessaires à la mécanisation de l'analyse du vers. Il importe donc de mentionner ses apports.

I.3.1. Les apports de la théorie générative

Halle et Keyser [1966 et 1971], qui sont les deux fondateurs de la métrique générative, projettent sur le vers le modèle de la phonologie et de la syntaxe génératives : il existe, selon ces auteurs, une structure métrique profonde et une structure de surface qui correspond à la séquence du vers. L'objet de la métrique est alors double : l'analyse des structures métriques profondes et l'étude des règles de transformation qui permettent de passer de la structure profonde à la structure de surface, autrement dit du mètre au

[14] J. Roubaud, dans un article publié dans *Poétique* en 1971, applique la théorie au décasyllabe des troubadours et au vers d'Arte Mayor espagnol. Après avoir présenté les concepts et les apports de la théorie générative, il propose un modèle abstrait du décasyllabe épique : « un segment métrique de longueur dix contenant deux segments métriques élémentaires ayant au plus une position commune » [p. 373]. Puis il distingue seize types de figures accentuelles, selon la position des accents : l'hypothèse est que la répartition des accents est un constituant de la structure accentuelle profonde. Pour le vers d'Arte Mayor, un des plus énigmatiques de la métrique espagnole, il propose aussi une structure profonde, en déduit toutes les structures de surface et les règles de passage de l'une à l'autre. Si le modèle permet de rendre compte d'une très grande partie des cas, il ne les couvre pas tous et mérite donc pour Roubaud d'être affiné. Beaumatin, dans un article récent [1999], réexamine les propositions de Roubaud qui avaient jusque-là été peu discutées par la critique.

vers[15]. La métrique générative n'a jamais prétendu rendre compte ni de la production, ni de la réception des vers. Son ambition est essentiellement descriptive : elle cherche à donner une définition formelle du vers qui permette de marquer au mieux la frontière entre vers et non-vers et de mesurer la complexité du vers. En effet, dès leurs premiers travaux sur le poète anglais Chaucer, Halle et Keyser [1966] cherchent à élaborer un indice de complexité des vers qui augmente avec le nombre et l'intensité des violations du modèle de base.

Proposer une définition formelle du vers ne paraît pas en forte contradiction avec l'objet habituel de la métrique. Pourtant l'irruption de la métrique générative a sérieusement modifié le champ de l'étude métrique. Halle et Keyser, dans la tradition chomskienne du « faire du passé table rase », proposent un modèle du vers ïambique qui rompt avec toute la tradition métrique. Traditionnellement, le pentamètre ïambique, dont la paternité est attribuée à Chaucer, est décrit comme une séquence de cinq pieds, chaque pied étant composé d'une syllabe non accentuée suivie d'une syllabe accentuée. Les poètes s'étant accordé de nombreuses licences par rapport à ce modèle, il a fallu constituer des listes d'exceptions autorisées. Les générativistes considèrent alors que la description est trop lâche puisqu'elle ne permet plus de distinguer un pentamètre ïambique d'une ligne de prose, et ils proposent leur propre modèle constitué d'un ensemble de règles ou de principes à partir desquels on pourra dire si telle séquence est oui ou non un vers[16]. La première définition qu'ils proposent du pentamètre ïambique est la suivante :

---

[15] « The study of meter must be therefore composed of two separate parts, namely, the study of abstract patterns and the study of the correspondence rules which enable a given string of words to be viewed as an instance of a particular abstract pattern. » [Halle et Keyser, 1971, p. 140].

[16] « Rather then look upon headless lines, lines with an inverted first foot, etc. as somehow deviant... we propose below a set of principles or rules which by their nature yield a large variety of metrical patterns, in the same way that rules of syntax yield a large variety of syntactic patterns. With respect to these rules, there will be one or two possible judgments. Either a line is metrical by virtue of conformity to the rules, or else a line is unmetrical by virtue of nonconformity to the rules. » [Halle et Keyser, 1966, p. 371-372].

> Principe 1 : le pentamètre ïambique est constitué de dix positions auxquelles peuvent être attachées une ou deux positions extramétriques ;
> Principe 2 : une position est normalement occupée par une syllabe, mais, sous certaines conditions, elle peut être occupée par plus d'une syllabe ou par aucune ;
> Principe 3 : un accent maximal ne peut occuper que les positions paires à l'intérieur du vers, mais il n'est pas nécessaire que toutes les positions paires soient occupées par un accent maximal.
>
> [Halle et Keyser, 1966].

Halle et Keyser abandonnent la notion de pied au profit de la syllabe et proposent le principe le plus controversé : le *Stress Maximum Principle* (SMP), qui dit qu'un accent maximal (c'est-à-dire supérieur à l'accent des deux syllabes adjacentes) ne peut apparaître que sur les positions paires. S (*Strong*) désigne les positions marquées ou accentuées du vers, W (*Weak*) les positions non marquées. Tel est le modèle du pentamètre ïambique que Halle et Keyser proposent au début des années 70 :

WSWSWSWSWS

et les règles de transformation sont les suivantes :

> (14a) Chaque S et W de la structure métrique correspond à une seule syllabe du vers.
> (14b) L'accent d'une syllabe correspondant à un S donné doit être plus fort que l'accent des syllabes correspondant au W qui précède et suit le S en question.
>
> [Halle, 1972, p. 111].

La proposition 14b est affinée pour pouvoir mesurer des degrés de complexité.

La définition du pentamètre ïambique paraît bien trop naïve aux métriciens déjà en place et bien des vers contredisent les principes proposés. Les générativistes ne cesseront d'enrichir leur définition du pentamètre, rétabliront la notion de pied, jusqu'à proposer une définition plus adéquate, mais beaucoup plus complexe, et plus proche de celle de leurs prédécesseurs. Ainsi Kiparsky [1975] met à jour un nouveau principe à la place du SMP : le *Monosyllabic Word Constraint* (MWC), qui n'autorise des accents en positions faibles que s'ils appartiennent à des monosyllabes. En fait, le MWC impose des contraintes sur les polysyllabes : les accents des polysyllabes doivent coïncider avec les positions accentuées du mètre.

On notera que ni les uns ni les autres dans leurs modèles de description du pentamètre ïambique n'ont porté une attention particulière à la question de la césure.

Pour Youmans [1986 et 1989], les principes métriques de la métrique générative doivent être interprétés comme des normes statistiques plutôt que comme des règles génératives. Cette attitude plus souple paraît nécessaire dans la mesure où jusqu'à présent, quel que soit le modèle proposé, des occurrences de vers ont à chaque fois contredit le modèle. La tâche de l'analyse métrique est alors d'identifier le prototype et de mesurer la distance de chaque vers à ce prototype[17], ce qui est finalement assez proche de la mesure de la complexité, l'évaluation du nombre de violations des règles. Autrement dit, la compatibilité d'un vers avec le modèle métrique est évaluée [*cf.* Lusson et Roubaud, 1974], et plus la compatibilité est faible, plus le vers peut être considéré comme complexe. La logique des ensembles flous serait mieux adaptée que les règles génératives axiomatiques pour décrire le pentamètre [Youmans, 1986, p. 400].

Les apports de la théorie générative au sens large sont nombreux. Nous retiendrons principalement ceux-ci :

- La distinction entre le mètre (structure profonde ou abstraite) et le vers (réalisation concrète de ce mètre), d'une part, et la notion de règles de transformation qui permettent de passer du mètre au vers, d'autre part, sont particulièrement pertinentes en vue de la modélisation. Comme le soulignait Roubaud [1971, p. 368], cette distinction entre structure profonde et structure de surface pourrait avec réserve être associée à celle qu'avait déjà faite Jakobson [*1960*, 1963, p. 229 et suivantes] entre *verse design* (modèle de vers) et *verse instance* (exemple ou occurrence de vers), cette dernière paraissant cependant plus pertinente. Complétée par les règles de transformation, cette distinction offre un cadre opérationnel pour la construction d'un outil d'analyse systématique du vers.
- La distinction entre l'accent métrique et l'accent linguistique, et le lien profond qui existe entre les deux méritent d'être retenus. Il existe des syllabes accentuables, celles qui correspondent à l'accent primaire des mots pleins, et des syllabes qui ne le sont pas. « Une syllabe *s* est porteuse d'accent métrique si elle est accentuable et si la syllabe *s'* qui la

[17] « I argue, therefore, that generative metrical rules such as the MWC are statistically normative principles rather than categorical constraints. Under this view, the task of metrical theorists is to establish a metrical prototype and determines degrees of deviation from this prototype rather then to define a precise boundary between metrical and unmetrical lines. » [Youmans, 1989, p. 5-6].

précède dans le vers, aussi bien que la syllabe *s*'' qui la suit, sont non accentuables ou bien n'appartiennent pas aux mêmes constituants syntaxiques » [Roubaud, 1971, p. 370]. Les accents métriques se déduisent des accents linguistiques, mais il n'y a pas toujours parfaite concordance entre les deux.

- La notion d'indice de métricalité peut être intéressante. Elle est dangereuse, puisqu'elle est fondée sur la notion d'écart par rapport à une norme, tant critiquée en stylistique, mais elle permet, dans la formulation « floue » de Youmans, de mesurer une distance par rapport à un prototype et donc d'évaluer les jeux avec le mètre.

I.3.2. L'approche connexionniste

L'approche de Malcolm Hayward se distingue nettement de celle de la métrique générative, même s'il cherche à montrer la compatibilité des théories. Hayward [1991] propose un modèle connexionniste d'analyse du vers, qui selon lui répond à trois insuffisances de la métrique générative : pas de prise en compte de l'interprétation, ni de l'incidence d'une position sur ses voisines, et notation qui ne se prête pas au traitement quantitatif. En plus de la structure syntaxique [Halle et Keyser, 1971] et de l'accent lexical [Kiparsky, 1977], son modèle prend aussi en compte l'interprétation du lecteur. Ensuite, la figure métrico-rythmique du vers produite par son modèle est le résultat d'interactions entre les différentes positions du vers. Enfin, concernant la notation, en adoptant des modalités numériques pour chaque variable, il se donne les moyens de faire des traitements quantitatifs sur des échantillons conséquents.

En quoi consiste ce modèle connexionniste ? Le vers y est représenté par une séquence d'unités qui correspondent aux syllabes métriques, par exemple dix unités pour le pentamètre anglais. L'objectif du modèle est de définir la probabilité d'être accentuée de chaque position métrique.

Les unités métriques sont connectées à cinq autres types d'unités représentant : l'intonation, l'accent lexical, les figures prosodiques, les caractéristiques propositionnelles, et les choix d'interprétation. Ces types d'unités peuvent être décrits comme des types de marquages tels que définis par Lusson et Roubaud [Lusson 1973 ; Roubaud *1978*, 1988] (*cf. infra*), où les poids sont compris entre 0 et 1.

Le premier type d'unité qui représente l'intonation correspond à un marquage de saillance de certaines syllabes. Le deuxième type d'unité, accent lexical, correspond à un marquage d'accent de mot. Le troisième, figure prosodique, correspond à un marquage de figures phonétiques : rime, allitération, assonance... Le quatrième revient à marquer les frontières de syntagmes. Enfin, le cinquième marquage correspond à des choix interprétatifs ; il offre par exemple un moyen de souligner un mot qui entre dans l'isotopie dominante du poème. Ce dernier marquage est clairement le plus subjectif de tous et correspond à un choix de lecture. Ces marquages sont faits « à la main » et sont donnés en entrée au modèle.

Par ailleurs, puisqu'il s'agit d'un modèle connexionniste, et c'est bien là l'originalité du modèle, il y a des « connexions » entre les différentes unités. Premièrement, des connexions négatives entre les unités métriques voisines : si une position métrique voit sa probabilité d'être accentuée augmenter, les positions adjacentes verront leur probabilité d'être accentuées diminuer. Ceci tend bien entendu à ïambiciser le vers. Le modèle part de l'hypothèse que le vers est ïambique, il introduit même un biais pour accentuer les positions paires. Mais il y a aussi des connexions entre les différents types de marquages : ainsi les marquages d'interprétation, de proposition et d'accent lexical ont tous trois des connexions positives avec le marquage d'intonation, qui lui agit directement sur le marquage métrique. Les connexions et leurs valeurs sont définies *a priori* dans le modèle.

Avant d'exécuter le programme informatique, les unités métriques ont une probabilité d'être accentuées nulle. Une fois que les valeurs correspondant aux cinq marquages sur les dix positions sont saisies, le programme est appliqué une trentaine de fois (nombre d'itérations nécessaire pour que les valeurs se stabilisent) et l'on obtient la séquence des poids de chaque vers.

Hayward [1996a] a appliqué son modèle à 100 vers d'une dizaine de poètes (neuf anglais et un américain ; de la Renaissance à l'époque victorienne), soit un corpus de 1 000 pentamètres ïambiques. Il recherche l'activation moyenne (la probabilité moyenne d'être accentuée) de chaque position et son écart-type pour chacun des dix poètes, ce qui lui permet de montrer, par un test statistique, que chacun des dix poètes a un mode d'accentuation particulier du vers. Ensuite, il mesure l'écart d'activation entre les deux syllabes de chaque pied, ce qui permet de mettre en évidence une manière de faire propre à chaque poète. Plus les écarts sont élevés, plus

le vers est ïambique. Enfin il analyse l'activation de chaque position selon les époques et montre ainsi que le vers victorien est le plus irrégulier.

Le modèle connexionniste propose une façon de combiner différents types de marquages linguistiques (lexicaux, morpho-syntaxiques et prosodiques) avec des marquages métriques (biais pour accentuer les positions paires et connexions négatives entre positions voisines) pour construire une mélodie des poids synthétique. Ce modèle combine des règles métriques à des règles linguistiques pour construire la figure du vers et propose une résolution élégante et paramétrable de la combinaison des marquages.

L'hypothèse ïambique est introduite d'emblée dans le modèle. Quelle serait la figure du pentamètre ïambique sans cette hypothèse ?

Hayward [1996b] cherche à montrer que son modèle peut être utile pour répondre à une des grandes questions de la métrique générative : la « métricalité » du vers. Quand peut-on dire qu'un vers est métrique ou bien qu'il ne l'est pas ? Il reprend des vers classiques longuement analysés dans la tradition générative, qu'il analyse avec son modèle. Si la séquence des probabilités d'accent produite peut être utilisée pour dire le vers, sans engendrer la violation de règles linguistiques, alors elle peut être considérée comme métrique, sinon elle est non métrique.

Reprenons le premier exemple cité par Hayward [1996b], repris de [Halle et Keyser, 1971] :

Ode to the West Wind, by Percy Bysshe Shelley

Comme *Per* occupe une position faible alors qu'il porte l'accent lexical, ce « vers » n'est pas métrique pour Halle et Keyser. Le modèle connexionniste conduit à un résultat cohérent avec la théorie générative, puisqu'il produit une mélodie impossible à dire car elle contredit les règles de l'accent en anglais : en effet, elle donne un poids accentuel plus fort à *-cy* qu'à *Per-* alors que l'accent lexical est porté par la première syllabe.

Ode (0,9) to (0,6) the (0,1) West (0,9) Wind (0,8), by (0,6) Per- (0,4) cy (0,7) Bysshe (0) Shel- (0,9) ley

Le modèle connexionniste est cohérent avec le modèle génératif et, de plus, il permet de transformer en syllabe accentuée toute syllabe entourée de deux syllabes non accentuées. Mais la séquence des poids n'est-elle pas trop tributaire des marquages donnés en entrée au modèle ?

L'introduction de l'interprétation comme marquage ne nous paraît pas être l'apport le plus important de Hayward, car ce marquage est trop soumis à la subjectivité du lecteur. Nous retiendrons de l'approche connexionniste le choix d'un système de description mathématique des positions par un ensemble de marquages (ce qui est déjà fait depuis les années 70 au Centre de Poétique Comparée) mais surtout la combinaison et l'interaction des marquages entre eux qui permet de construire une figure rythmique du vers globale qui intègre toutes les dimensions.

I.4. APPLICATION DE LA THEORIE DU RYTHME DE PIERRE LUSSON A L'ALEXANDRIN

Nous avons évoqué quelques-uns des apports pour l'étude du vers de la théorie générative. En France, le courant qui s'est constitué autour du cercle Polivanov et du Centre de Poétique Comparée a été marqué par la théorie générative du vers, en particulier par les écrits de Halle et Keyser [1966 et 1971].

Le cercle Polivanov a été créé en 1969. L'objet d'étude de ce cercle est la poétique, c'est-à-dire « l'étude des propriétés formelles et sémantiques des textes (et autres manifestations) de ce qu'on entend généralement par "poésie" » [Cercle Polivanov, 1969]. On lit entre autres comme thèmes de recherche :

- exploration typologique des rapports entre forme poétique (au sens large) et mécanismes linguistiques.
- problèmes de formalisation.
- contrôle des thèses élaborées en les soumettant à l'épreuve d'objets concrets très divers.

[Cercle Polivanov, 1969].

Ce cercle est complété par la création du Centre de Poétique Comparée et de la revue *Cahiers de Poétique Comparée* (dirigée par Pierre Lusson, Léon Robel et Jacques Roubaud).

Ce cercle s'est intéressé à la métrique générative américaine et a développé sa propre « théorie du rythme ». Les méthodes d'investigation formelle développées ont été appliquées à des vers et à des formes poétiques très diverses : l'alexandrin ordinaire [Lusson et Roubaud, 1974 ; Roubaud,

1978, 1988, 1986 et 1988], le sonnet en vers rapportés[18] [Lusson et Roubaud, 1980 et 1981], le sonnet [Roubaud, 1990], le pentamètre ïambique anglais [Roubaud, 1983], le pentamètre ïambique russe [Prijovíc et Sola, 1983], et même à des formes musicales ou mixtes : analyse des rapports texte/musique dans *L'Alceste* de Lulli [Lusson 82]... Ces méthodes furent également appliquées pour l'analyse formelle de traductions. Par exemple, la figure rythmique des huit traductions du sonnet en vers rapportés de Domenico Veniero a été étudiée par [Lusson et Roubaud, 1983]. Enfin, des lectures rythmiques de poèmes spécifiques ont été faites tout au long de ces années : elles consistent à faire émerger le squelette rythmique des poèmes et à mettre en relation les phénomènes métrico-rythmiques avec des phénomènes lexicaux et sémantiques.

L'application à l'étude du vers, et en particulier à celle de l'alexandrin, de la théorie du rythme développée par Lusson [1973, 1998][19] est présentée dans l'annexe I rédigée par Lusson.

Dans ce cadre, le modèle métrique de l'alexandrin à rimes plates présente les caractéristiques suivantes. Il est, comme le prévoit la théorie, constitué de plusieurs niveaux parfaitement hiérarchisés (*cf.* Figure 2 et Figure 3) [Roubaud, *1978*, 1988, p. 94-95 pour les niveaux 4 et 5 ; 1986, p. 64 pour les niveaux 0 à 3]. Chaque niveau est défini par un type d'événement élémentaire spécifique.

---

[18] Voici le premier quatrain d'un sonnet en vers rapportés de Jodelle. Trois figures, qui correspondent au monde hault, moyen et bas, sont traitées en parallèle sur chaque vers.

Des astres, des forest et d'Achéron l'honneur
Diane au monde hault, moyen et bas préside,
Et ses chevaux, ses chiens, ses Euménides guide,
Pour esclairer, chasser, donner mort et honneur.

Étienne Jodelle

On pourra également se reporter à la lecture sémantique que Rastier [1989] propose de ce sonnet.

[19] Une première esquisse de la théorie a été publiée par Lusson en 1973 : « Notes préliminaires sur le rythme ». Cette théorie est également exposée dans « Une méthode d'analyse des rapports texte/musique : application d'une théorie générale du rythme » [Lusson, 1998]. Une présentation sommaire en est faite par Roubaud dans *La Vieillesse d'Alexandre* [*1978*, 1988] et dans « T.R.A.M(M,m) » [1991].

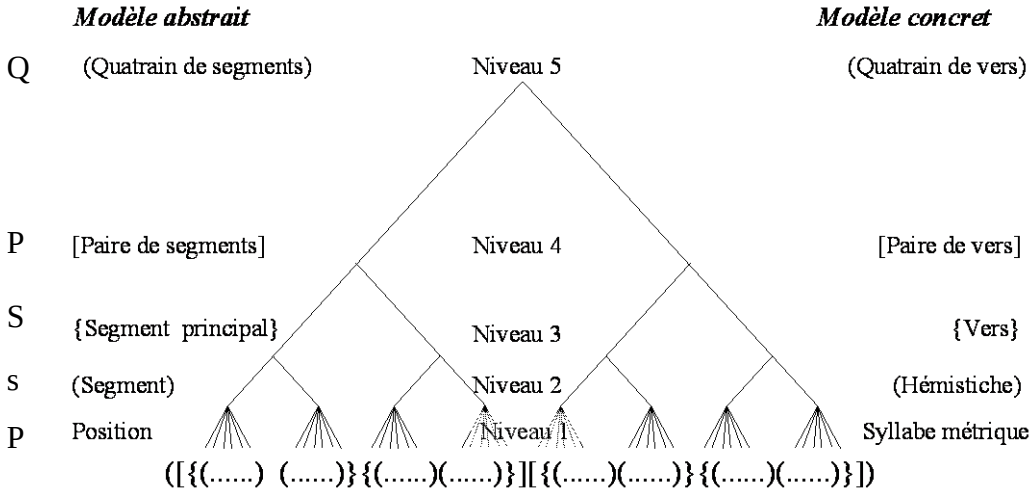


Figure 2. Le modèle de l'alexandrin classique à rimes plates alternées

Niveau 5	Q (Quatrain) = S S S S L'événement élémentaire est la séquence de quatre segments principaux (S)
Niveau 4	P (Paire) = (S S) L'événement élémentaire est la séquence de deux segments principaux (S)
Niveau 3	S (Segment principal) L'événement élémentaire est le segment principal (S)
Niveau 2	S = (s s) L'événement élémentaire est le segment (s)
Niveau 1	S = ((p p p p p p)(p p p p p p)) L'événement élémentaire est la position (p)
Niveau 0	S = $(((nvn)_p(nvn)_p(nvn)_p(nvn)_p(nvn)_p(nvn)_p)_s((nvn)_p(nvn)_p(nvn)_p(nvn)_p(nvn)_p(nvn)_p)_s)_{Sn}$ Les événements élémentaires sont de deux types n ou v

Figure 3. Niveaux métriques abstraits de l'alexandrin à rimes plates

Au niveau 1 du modèle, le vers est composé de douze positions (la position est l'événement élémentaire à ce niveau). Ces positions sont regroupées au niveau 2 en deux séquences de six, qui constituent deux segments métriques, les hémistiches. L'hémistiche coïncide en général avec un syntagme ou groupe de souffle et sa dernière position est *toujours* très marquée (surtout par rapport à la précédente). Les deux hémistiches sont regroupés à un niveau supérieur (niveau 3) pour former le vers, comme événement élémentaire.

Ensuite, dans la mesure où nous sommes dans un contexte d'alexandrins à rimes plates avec alternance en genre (*cf.* chapitre II sur le corpus), nous pouvons introduire deux niveaux métriques supplémentaires. Le niveau 4

est défini par la paire de vers rimant ensemble (son existence est assurée par la rime) et le niveau 5 par le quatrain constitué de deux rimes plates féminines suivies de deux rimes plates masculines ou l'inverse (en fonction du genre du premier vers).

Ces différents niveaux d'organisation métrique de l'alexandrin sont parfaitement emboîtés les uns dans les autres et constituent l'organisation hiérarchique la plus simple. Pour les rimes embrassées ou alternées, la hiérarchisation serait plus complexe puisqu'elle impliquerait des intrications.

Rappelons que ce modèle est exclusivement métrique. Par conséquent, le niveau qui se situe entre celui de la position métrique et celui de l'hémistiche n'est pas pris en compte puisqu'il est essentiellement rythmique. Nous explorerons les particularités rythmiques de ce niveau intermédiaire entre la position et l'hémistiche.

La Figure 3 présente la partie abstraite du modèle. On y trouve un niveau métrique supplémentaire par rapport à la Figure 2 qui décrit les constituants de la position métrique [Roubaud, 1986]. Une position métrique est occupée par une voyelle métrique, notée v, entourée de non-voyelles (consonnes, semi-voyelles, vide...), notées n. Le cœur de la position métrique est la voyelle, puisque les autres éléments sont facultatifs.

Ce modèle de l'alexandrin est construit à partir d'un système de marquage spécifique. Considérons les positions métriques de l'alexandrin comme événements élémentaires, et définissons le système de marquage suivant appliqué à 48 positions, soit quatre vers à rimes plates :

- la dernière position d'un hémistiche est marquée (marquage 2 : m2) ;
- la dernière position d'un vers est marquée (m3) ;
- la dernière position d'une paire de vers rimant ensemble est marquée (m4) ;
- la dernière position de quatre alexandrins à rimes plates avec alternance en genre est marquée (m5).

On obtient les séquences de 0 et de 1 suivantes :

m2 : 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1
m3 : 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
m4 : 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
m5 : 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

Si l'on combine ces marquages en les additionnant, on obtient la « mélodie des poids » suivante :

0 0 0 0 0 1 0 0 0 0 0 2 0 0 0 0 0 1 0 0 0 0 0 3 0 0 0 0 0 1 0 0 0 0 0 2 0 0 0 0 0 1 0 0 0 0 0 4

dont on peut extraire le squelette rythmique du vers. Il suffit de placer une parenthèse fermante après le ou les événement(s) élémentaire(s) de poids maximal, puis de réitérer la procédure à l'intérieur des groupements ainsi définis et enfin de compléter les parenthésages par des parenthèses ouvrantes. Le squelette rythmique des alexandrins est alors :

((((000001)(000001))((000001)(000001)))((((000001)(000001))((000001)(000001))))

Etant donné ces particularités, ce squelette peut être considéré comme le modèle métrique de l'alexandrin.

D'un point de vue pratique, maintenant que nous avons le modèle métrique du vers, comment peut-on identifier les positions métriques à l'intérieur du vers de manière systématique (et automatique) ?

La question du repérage des positions métriques et la question du dénombrement qui en découle concernent les niveaux 0 et 1 du modèle. Comment établir la correspondance entre la séquence des graphèmes ou phonèmes (selon que l'on privilégie l'œil ou l'oreille) et les positions abstraites ? Un bon lecteur ou auditeur (ayant longuement fréquenté les vers) fait intuitivement une bonne séparation des segments, un lecteur ou auditeur moins chevronné y parviendra également en tâtonnant pour atteindre le chiffre douze (s'il s'agit d'alexandrins...). Mais l'exercice qui consiste à faire opérer ce même découpage en positions métriques par une machine oblige à formaliser les règles ou les principes à l'œuvre et révèle une série de difficultés. Un de nos objectifs premiers sera donc la définition des règles de transformation qui permettent de passer de la séquence du vers au modèle abstrait.

Nous venons de décrire rapidement la structure hiérarchisée du modèle métrique de l'alexandrin et le principe à mettre en œuvre pour établir la correspondance entre le vers attesté et le modèle. Comme on pouvait s'y attendre, la théorie du rythme ne se contente pas d'avoir comme seule

ambition l'analyse du *mètre*. Autrement dit, elle ne se limite pas à l'exploration de la structure et de l'articulation des différents niveaux hiérarchisés telle que nous venons de l'examiner : elle vise aussi l'analyse du *rythme*. En effet, une fois posé le modèle précédent, qui a le mérite d'intégrer le vers dans une structure plus complexe que ne le font la plupart des modèles du vers, on a épuisé, ou plutôt on a dit l'essentiel de ce qui constitue la nature du modèle métrique de l'alexandrin classique. C'est maintenant le rythme, la variation à l'intérieur du cadre métrique, que nous allons explorer.

Pour aller plus loin, il nous faut examiner d'autres formes qui se dessinent dans la séquence des vers, trouver d'autres régularités. Par exemple, à l'intérieur du niveau métrique que constitue l'hémistiche, Roubaud et Lusson [Lusson et Roubaud, 1974 ; Roubaud, *1978*, 1988] ont montré que l'alexandrin privilégie les deux groupements rythmiques suivants : les figures 001001 et 010101. Ils ont, par commodité de langage, désigné celles-ci comme figures anapestiques et ïambiques (sans que jamais il n'y ait eu méprise sur l'utilisation de ces termes, contrairement à ce que laisse entendre Meschonnic [1982, p. 201 et p. 230 et 231]).

Pour l'exploration de ces structures rythmiques, il nous faut recourir à la notion de marquage définie dans la théorie du rythme. Une fois que sont identifiées les positions, l'analyse de la structure métrico-rythmique repose sur des systèmes de description associés aux positions. Autrement dit, quelles sont les caractéristiques de chaque position ? Quels critères de marquage retenir pour faire apparaître le mieux possible la structure rythmique du vers ? L'approche de Lusson consiste à multiplier les points de vue, c'est-à-dire les types de marquages, à les combiner pour faire ressortir avec le plus de netteté possible la structure métrico-rythmique. Les descripteurs pertinents sont ceux qui permettent de contraster au mieux les positions.

Ceux-ci peuvent être de n'importe quelle nature. Ainsi, pour le vers, nous pourrions imaginer à partir d'enregistrement de vers de considérer la hauteur, la durée et l'intensité. La combinaison de ces trois critères, qui sont censés se renforcer dans la théorie de l'accent, permettraient d'identifier les accents de vers comme des maxima locaux. Mais nous pouvons aussi opter pour des descripteurs graphiques (les signes de ponctuation par exemple), pour des marquages linguistiques (syntaxiques, lexicaux, prosodiques) ou pour des marquages métriques.

Nous avons fait le choix d'un certain nombre de marquages ou étiquetages, qui relèvent tous de caractéristiques du discours écrit : marquages phonétiques, lexicaux, morpho-syntaxiques et prosodiques. Il y a ici implicitement l'hypothèse que le rythme est un effet de l'organisation de la langue dans le discours et non pas un effet métrique : nous n'introduisons pas de marquage métrique. C'est sur la séquence des positions métriques que nous avons choisi de projeter ces marquages, que nous aurons l'occasion de détailler au chapitre V.

Les marquages linguistiques ne définissent pas un système hiérarchisé comme pour le modèle métrique de l'alexandrin vu précédemment : nous traiterons donc chacun de ces marquages de manière indépendante.

La représentation adoptée pour chaque marquage peut être décomposée sur un mode binaire, représenté de manière simplifié par un « 1 » si le critère est réalisé ou un « 0 » s'il ne l'est pas. Toutes les positions marquées « 1 » sont identiques au regard d'un critère donné (ainsi que toutes les positions « 0 », par simple symétrie), et opposées aux positions « 0 » (respectivement « 1 »). Malgré sa rusticité, ce système de notation permet de décrire n'importe quel critère, quelle que soit sa complexité, pourvu que le nombre de dimensions notées 0/1 soit suffisant. Nous aurons autant de séquences de 0/1 que nous avons choisi de marquages.

Pour clore ce chapitre, l'alexandrin sur lequel nous travaillons n'est donc pas l'alexandrin du XIIe siècle, employé principalement pour décrire les exploits d'Alexandre et marqué par sa césure épique, mais un vers qui renaît au XVIe siècle et dont le type de césure est nouveau, inspiré par celui du décasyllabe de l'époque.

Nous avons une représentation hiérarchisée de son organisation métrique et un ensemble de concepts utiles pour la mise en œuvre d'un outil d'analyse systématique du vers. Nous commencerons par décrire le corpus (chapitre II), puis nous aborderons la question de la rime (chapitre III) qui concerne les niveaux 4 et 5 du modèle métrique avant d'en venir à l'identification de positions métriques (chapitre IV), aux marquages associés (chapitre V), à l'outil construit (chapitre VI), à l'analyse de la structure du vers (chapitre VII) et à l'existence de liens entre la structure rythmique et le sens (chapitre VIII).

II. LE CORPUS

L'étude du vers s'appuie nécessairement sur l'examen de matériel empirique. Mais le plus souvent, celui-ci était limité par la capacité d'observation d'un seul être humain. Aujourd'hui, grâce au développement de la numérisation, de plus en plus de textes, y compris en vers[20], sont saisis sur support numérique et les techniques d'analyse « automatique » des textes ne cessent de croître en efficacité. Assistés par l'ordinateur, nous avons les moyens d'examiner systématiquement des corpus de textes de plus en plus vastes. Il s'agit de vers écrits et non pas de vers dits, limite que nous avons évoquée en introduction et sur laquelle nous reviendrons dans la section consacrée aux limites du métromètre (*cf.* chapitre VI.2).

Le travail ici présenté se veut avant tout empirique, dans la mesure où ses principales conclusions sur l'alexandrin classique sont tirées de l'examen systématique d'un corpus de près de 80 000 vers, constitué des œuvres théâtrales complètes de Corneille et Racine, parmi lesquelles on compte plus de 77 000 alexandrins. Ce corpus a été saisi sur support informatique par l'INaLF, d'après l'édition de Ch. Marty-Laveaux [1862] pour Corneille et de P. Mesnard [1885] pour Racine, qui toutes deux présentent les textes dans une orthographe modernisée. Le corpus a été mis en forme par la Bibliothèque Nationale et a été confié au Centre de Poétique Comparée. Jacques Roubaud nous a donné ces textes dès le début de notre entreprise en 1993. Tous ces vers ont été analysés par le métromètre.

De quelles pièces est constitué notre corpus ? Comment ces textes ont-ils été préparés ? Nous présentons ci-dessous en détail l'anatomie du corpus CORRAC. Ce corpus a servi de base à la construction du métromètre, notre outil d'analyse automatisée du vers, et à l'exploration empirique de

[20] Cependant, dans les textes numérisés la part accordée aux textes en vers est relativement faible. Henri Béhar [1994] a montré que c'était en particulier le cas pour la base Frantext. Il s'est pour cela appuyé sur la Banque de Données de l'Histoire Littéraire (BDHL), répertoire des œuvres littéraires réalisé par Henri Béhar et son équipe à Paris III.

l'alexandrin classique. Il a aussi servi de référence pour l'étude des variations et de l'évolution de l'alexandrin. En effet, d'autres corpus ont permis de donner un aperçu de l'évolution du vers (*cf.* chapitre VII.6).

II.1. PIECES DE CORNEILLE ET RACINE

Les tableaux ci-dessous présentent l'ensemble des pièces écrites par Corneille et Racine, leur date de première publication et leur genre. Pour le genre des pièces de Corneille, nous indiquons la dernière désignation donnée par l'auteur lui-même, mais aussi la classification proposée et commentée par Charles Muller dans son ouvrage de référence : *Le Vocabulaire du théâtre de Pierre Corneille : Étude statistique lexicale* [*1967*, 1992] ; pour les pièces de Racine, nous avons repris la classification donnée par Racine et les commentaires de Charles Bernet [1983] qui a fait, sur Racine, un travail similaire à celui de Ch. Muller. Ce sont finalement les classifications de Muller et Bernet que nous utilisons par la suite.

CORNEILLE			
Pièce	**Date**	**Genre d'après l'édition de Ch. Marty-Laveaux**	**Genre d'après Ch. Muller**
Mélite	1629	comédie	comédie
Clitandre	1632	tragédie	Pièces diverses (tragi-comédie, puis tragédie)*[21]
La Veuve	1633	comédie	comédie
La Galerie du palais	1634	comédie	comédie
La Suivante	1634	comédie	comédie
La Comédie des Tuileries, 3eacte	1635	comédie	
La Place royale	1635	comédie	comédie
Médée	1635	tragédie	tragédie
L'Illusion comique	1636	comédie	comédie
Le Cid	1636	tragédie	tragédie (tragi-comédie, puis tragédie)*
Horace	1640	tragédie	tragédie
Cinna	1640	tragédie	tragédie
Polyeucte	1640	tragédie	tragédie
Pompée	1641	tragédie	tragédie
Le Menteur	1642	comédie	comédie
La suite du Menteur	1643	comédie	comédie
Rodogune	1644	tragédie	tragédie
Théodore	1645	tragédie chrétienne	tragédie
Héraclius	1647	tragédie	tragédie
Andromède	1650	tragédie	Pièces diverses (pièce à machine[22])*
Don Sanche d'Aragon	1650	comédie héroïque[23]	Pièces diverses
Nicomède	1651	tragédie	tragédie
Pertharite, roi des lombards	1652	tragédie	tragédie
Œdipe	1659	tragédie	tragédie
La Toison d'or	1660	tragédie	Pièces diverses (pièce à machine)*
Sertorius	1662	tragédie	tragédie
Sophonisbe	1663	tragédie	tragédie
Othon	1664	tragédie	tragédie
Agésilas	1666	tragédie	tragédie
Attila	1667	tragédie	tragédie
Tite et Bérénice	1670	comédie héroïque	Pièces diverses
Psyché	1671	tragédie-ballet	Non retenue comme œuvre de Corneille
Pulchérie	1672	comédie héroïque	Pièces diverses
Suréna	1674	tragédie	tragédie

Tableau 2. Les pièces de Corneille

[21] L'appartenance à un genre donné est douteuse, a été discutée.

[22] Pièces à machine : « grands spectacles accompagnés de musiques, de chants et de danses » [Devaux, 1993, p. 6].

[23] Seule la fin heureuse distingue la comédie héroïque de la tragédie.

RACINE			
Pièce	**Date**	**Genre d'après l'édition de P. Mesnard**	**Genre d'après Ch. Bernet**
La Thébaïde ou les frères ennemis	1664	tragédie	tragédie profane
Alexandre le Grand	1665	tragédie	tragédie profane
Andromaque	1667	tragédie	tragédie profane
Les Plaideurs	1668	comédie	comédie
Britannicus	1669	tragédie	tragédie profane
Bérénice	1670	tragédie	tragédie profane
Bajazet	1672	tragédie	tragédie profane
Mithridate	1673	tragédie	tragédie profane
Iphigénie	1674	tragédie	tragédie profane
Phèdre	1677	tragédie	tragédie profane
Esther	1689	tragédie tirée de l'écriture sainte	tragédie sacrée
Athalie	1691	tragédie tirée de l'écriture sainte	tragédie sacrée

Tableau 3. Les pièces de Racine

Les éditions électroniques de Corneille et Racine dont nous disposons utilisent une orthographe moderne. Cette question de l'orthographe mérite qu'on s'y attarde. Pour l'étude du vers, l'idéal est de disposer des textes dans leur orthographe originelle, ne serait-ce que pour mieux apprécier la qualité graphique de la rime. Or, les outils d'analyse automatisée de la langue sont en général conçus pour l'analyse de textes d'aujourd'hui et rencontrent d'importantes difficultés face à des textes du XVIIe siècle. Nous voilà confrontés à un choix pénible : d'un côté, l'objet d'étude - le vers - nous pousse à travailler sur l'orthographe d'époque ; de l'autre, les outils mis en œuvre nous incitent à privilégier des textes dans une orthographe modernisée.

Comment concilier ces deux injonctions contradictoires ? Une solution aurait consisté à construire des outils de linguistique informatique adaptés à l'écriture du XVIIe siècle. Il aurait alors fallu disposer des versions informatiques des œuvres de Corneille et de Racine d'après leur édition d'origine. En effet, l'édition électronique de Frantext a été réalisée dans une orthographe moderne. Il faudrait par exemple saisir pour Racine la remarquable édition de Pierre Mélèse aux éditions Richelieu [1951]. Cette solution n'a pas été retenue : fabriquer un métromètre adapté à l'orthographe du XVIIe siècle posait aussi la question de son utilisation pour les textes modernes. Nous avons donc utilisé des corpus dans une orthographe modernisée.

Ce corpus des pièces de Corneille et Racine a servi à mettre au point le métromètre. Les interactions entre l'outil et le corpus au fil des tests ont

évidemment permis d'améliorer l'outil mais aussi de corriger le corpus et d'améliorer sa présentation. De nombreuses erreurs (*cf. infra*) ont été repérées au cours des tests : certaines provenaient du métromètre, d'autres de l'édition. Le repérage de ces erreurs a permis d'aboutir à une version du métromètre assez satisfaisante sur le corpus CORRAC et à une édition des œuvres de Corneille et Racine bien moins entachée d'erreurs.

Précisons enfin que l'objet de l'étude n'est pas ici la poésie, ni d'ailleurs le théâtre, mais le vers d'un point de vue formel. Sans doute le vers de poésie n'a-t-il pas exactement les mêmes caractéristiques que le vers de théâtre, aussi construisons-nous une image du vers classique de théâtre.

II.2. PREPARATION DU CORPUS

La préparation du corpus revient à créer une nouvelle édition harmonisée des pièces de Corneille et Racine. A supposer que l'on ait environ vingt-cinq vers par page, notre corpus tiendrait sur trois mille pages, soit treize volumes de 250 pages. On comprend alors que la préparation, la correction et l'harmonisation du corpus aient constitué une partie importante de notre entreprise.

Le corpus a été balisé dans une double perspective : permettre au métromètre de n'analyser que les vers dans le corpus et faire des traitements statistiques sur les résultats de ces analyses. Pour que le métromètre n'analyse pas n'importe quelle séquence de caractères, il a fallu clairement distinguer les vers de toutes les lignes de texte qui n'étaient pas en vers. Par ailleurs, nous souhaitions conserver les didascalies[24] de chaque pièce (numéro d'acte, de scène, nom des personnages, indications de mise en scène...) en vue d'éventuels traitements stylistiques. Le but était de conserver dans la base statistique toute l'information disponible dans l'édition électronique.

Le format de balises, directement accessible par le métromètre, est préservé par l'outil et facilitera la constitution d'une base de données très

[24] Les spécialistes de théâtre emploient le terme de *didascalies* pour désigner tous les éléments du texte qui ne sont pas prononcés par les personnages, et non pas seulement pour désigner les indications scéniques.

complète : pour chaque vers, sont enregistrées toutes les informations nécessaires pour le situer (pièce, acte, scène, personnage...)[25].

La mise en forme adoptée pour les textes respecte donc des contraintes rigoureuses qui n'ont pu dans le cadre de ce travail être que partiellement exploitées, mais qui faciliteront la poursuite des travaux et en particulier le test d'hypothèses plus fines. Le respect de ce format permettra d'analyser aisément d'autres corpus, et d'enrichir notre base de vers.

II.2.1. Mise en forme du paratexte

Comme le rappelle Anne Ubersfeld dans *Lire le théâtre* [1977, p. 20], un texte de théâtre « se compose de deux parties distinctes mais indissociables, le *dialogue* et les *didascalies* ». La distinction entre dialogue et didascalies renvoie à la question du sujet de l'énonciation : le dialogue est pris en charge par les personnages, les didascalies par l'auteur. Mais, dans les pièces de Corneille et Racine, cette distinction recouvre celle du vers et de la prose. Le dialogue est évidemment en vers, les didascalies toujours en prose. En effet, dans une pièce classique dite en vers, aucune parole prononcée n'échappe à la forme versifiée. Le métromètre n'étant pas destiné au théâtre, mais au vers en général, il vaut sans doute mieux reprendre l'opposition entre texte et paratexte (titres et sous-titres, préfaces, post-faces, avant-propos, épigraphes, notes…), proposée par Genette dans *Palimpsestes* [1982, p. 10], qui a une portée plus générale que la distinction entre dialogue et didascalies.

Le codage du texte doit distinguer clairement les vers de tout ce qui n'est pas vers, le dialogue des didascalies, le texte du paratexte. Ces deux grandes catégories ont chacune donné lieu à des codages spécifiques.

Les noms des pièces, leur auteur, les numéros d'acte, de scène, les noms de personnages, de lieux, les indications scéniques définissent « le contexte de la communication ; elles déterminent donc une *pragmatique*, c'est-à-dire les conditions concrètes de l'usage de la parole » [Ubersfeld, 1977, p. 21] : qui parle, d'où, à qui, comment ?

---

[25] Etant donné la taille du corpus et la sophistication des outils utilisés dans le métromètre, tous les traitements ont été réalisés sous UNIX dans un environnement de station de travail Sun. Pour la préparation du corpus, nous avons utilisé un éditeur de texte perfectionné, Emacs, qui permet d'automatiser en partie les tâches de balisage.

Nous avons fait précéder d'au moins une étoile (*) toutes les lignes de texte qui ne correspondent pas à des vers : les informations structurelles sur les pièces (numéros d'acte, de scène, de vers, noms de personnages...), les didascalies et les fragments de discours en prose. Nous avons donc harmonisé tous les textes avec nos règles de codage qui figurent dans le Tableau 4.

Pièces de théâtre
**** *{auteur} *{nom de l'œuvre} *{genre} *{date}
**ACTE_{n° de l'acte}
*SCENE_{n° de scène}
***{nom de personnage}
***_{didascalies}
*{n° de vers, avant chaque vers multiple de 5}

Tableau 4. Règles de codage

Ce codage a un objectif trivial : éviter que le métromètre n'analyse comme des vers les indications para-textuelles. Tout ce qui est précédé par au moins une étoile ne sera pas traité, mais codé comme autant d'informations sur le contexte d'énonciation du vers. Ainsi, tous les vers qui viennent après le nom d'un personnage seront reconnus comme étant pris en charge par ce personnage. Ce codage nous permet par la suite de donner les références des vers d'une manière assez précise, de comparer les auteurs, les textes, éventuellement les différents actes voire les scènes d'une pièce...

Suite à toutes ces opérations de mise en forme ou de formatage, les pièces de théâtre se présentent comme une suite de lignes codées. Le numéro de vers apparaît tous les cinq vers. Voici le début de *Mélite* de Corneille :

**** *Corneille *Mélite *comédie *1629
**ACTE_I
*SCENE_PREMIERE.
***ÉRASTE.
Je te l'avoue, ami, mon mal est incurable ;
Je n'y sais qu'un remède, et j'en suis incapable :
(...)
*5
(...)

L'INaLF est en train de mettre au point un nouveau CD-ROM, DISCOTEXT 2, consacré au théâtre classique [Martin, 1997, p. 39]. Les pièces seront présentées en hypertexte, ce qui permettra de distinguer aisément les vers des didascalies et noms de personnages.

II.2.2. Reconstitution et correction des vers

Outre le paratexte, les vers eux-mêmes ont donné lieu a des traitements spécifiques.

Nous avons tout d'abord reconstitué les vers fragmentés pris en charge par plusieurs personnages. Dans le théâtre classique, il arrive que des changements de tour de parole interviennent au milieu d'un vers. Ce travail a été fait d'une part pour préparer le travail sur la rime et d'autre part pour analyser les vers avec le métromètre.

Pour le métromètre, il doit y avoir coïncidence entre le vers et la ligne, et cela conformément à l'étymologie du mot *vers* (latin v*ersus)* : « sillon, ligne ». En effet, même lorsqu'il est découpé, le vers respecte les contraintes métriques. Ainsi, dans le vers suivant d'*Alexandre le Grand*, le *e* muet de *prince*, qui serait élidé s'il était en fin de vers, compte pour une voyelle métrique :

***PORUS.
(...) Pourquoi par tant de soins voulez-vous épargner
Un prince...
***AXIANE.
C'est pour vous que je le veux gagner.
Racine, *Alexandre le Grand*, vers 299.

Si nous n'avions reconstitué les vers, *Un prince* aurait été analysé comme un vers de deux syllabes, alors qu'il n'est qu'un morceau à trois positions métriques d'un alexandrin. Prenons un autre exemple :

***ŒNONE.
Pour qui ?
***PHEDRE.
*260
Tu vas ouïr le comble des horreurs.
J'aime... à ce nom fatal, je tremble, je frissonne.
J'aime...
***ŒNONE.
Qui ?
***PHEDRE.
Tu connais ce fils de l'amazone,
Ce prince si longtemps par moi-même opprimé ?

Racine, *Phèdre*, vers 260-263.

Ces vers fragmentés ont été reconstitués, comme dans les vers suivant, en conservant la trace du découpage du vers par l'insertion du signe « / » [26].

***ŒNONE.
Pour qui ? / Tu vas ouïr le comble des horreurs.
***PHEDRE.
*** Tu vas ouïr le comble des horreurs.
*260
J'aime... à ce nom fatal, je tremble, je frissonne.
J'aime... / Qui ? / Tu connais ce fils de l'amazone,
***ŒNONE.
*** Qui ?
*** PHEDRE.
*** Tu connais ce fils de l'amazone,
Ce prince si longtemps par moi-même opprimé ?

Racine, *Phèdre*, vers 260-263.

La reconstitution des vers permet d'identifier plus simplement les mots-rimes.

Chez Racine, 4 % des vers sont dits par plusieurs personnages. Etant donné la longueur de ce travail de ce travail de reconstitution des vers, il n'a été fait que sur les pièces de Racine. Sur les œuvres de Corneille, il a été fait pour l'étude du nombre de vers et de la rime, à partir des sorties du métromètre, au lieu d'être fait en amont.

[26] Cette reconstitution entraîne quelques incohérences dans l'attribution du personnage. Tout le « vers » sera attribué à celui qui l'a entamé. Les segments de vers non initiaux sont conservés, mais précédés de « *** », ce qui les marque comme des commentaires.

Outre ce travail de reconstitution des vers, de nombreuses corrections manuelles ont été opérées. Elles ont comme premier objectif de faciliter l'analyse syntaxique. Par exemple, beaucoup de *à* ne portaient pas l'accent. Le second objectif de ces corrections est d'assurer la cohérence métrique des vers. L'examen des alexandrins analysés par le métromètre comme ayant moins ou plus de 12 positions a permis de repérer des erreurs d'édition. On a recensé plus d'une trentaine d'erreurs : elles sont négligeables face à la quantité de vers, mais elles témoignent de la fiabilité du métromètre. Toute nouvelle édition de Corneille et Racine devrait passer le test du métromètre avant d'être publiée.

Voici à titre d'exemple quelques erreurs d'édition que nous avons repérées, qui ont un impact direct sur le décompte :

(13 positions) Aussi fais-je encore plus, je lui donne un époux. => encor
Corneille, *Sertorius*, vers 979.

(11 positions) Tout ce que peut l'effort de cher conquérant, => de ce cher
Corneille, *Othon*, vers 1795.

(11 positions) Ce solyman jeta les yeux sur Roxane => Roxalane
Racine, *Bajazet*, vers 466.

(12 positions : 7-5) Polynice, Etéocle, Jocaste, Antigone, =>Iocaste
Racine, *La Thébaïde*, vers 1509.

Cette dernière erreur apparaît dans les éditions les plus courantes de Racine. Si *Iocaste* commençait par un J, on serait face à un dodécasyllabe sans césure en sixième position, tout à fait en contradiction avec la nature de la césure dans le vers classique, qui n'admet césure ni épique, ni lyrique, ni italienne.

(13 positions) À venir prodiguer sa vie sur un théâtre => sa voix
Racine, *Britannicus*, vers 1475.

Le dernier exemple, avant correction, présente une situation (*e* muet final après voyelle devant un mot commençant par une consonne) interdite dans le vers classique : en effet, au XVIIe, ce *e* muet ne se prononce déjà plus en langue, mais les règles du vers tendent à en faire une syllabe métrique qui serait de plus en hiatus avec le *i* de *vie* : pour éviter cette contradiction entre langue et vers, ce type de situation a été proscrit par Malherbe.

Enfin, des corrections ont été apportées pour assurer la cohérence de la transcription avec les contraintes de la rime. Ainsi, une rime sans écho dans un contexte d'alexandrins à rimes plates indiquait qu'un vers n'avait pas été saisi, deux mots-rimes identiques qu'un même vers avait été saisi deux fois.

Le travail sur la rime nous a permis de déceler un nombre élevé d'erreurs. Beaucoup sont des choix de poètes, plus ou moins autorisés, qui ont été gommés par les éditeurs, comme dans les exemples suivants :

> Ils ne sont pas venus, nos deux rois ? Qu'on leur dise => die
> Qu'ils se font trop attendre, et qu'Attila s'ennuie ;
>
> Corneille, *Attila*, vers 1-2.

> Semble s'être assemblé contre nous par hasard : => hasar
> Je veux dire la brigue et l'éloquence. Car
>
> Racine, *Les Plaideurs*, vers 729-730.

> Elle va revenir ; elle vient, je la vois : => voi
> Du moins, pour son honneur, Rodrigue, cache-toi.
>
> Corneille, *Le Cid*, vers 772.

Parfois ce sont de véritables erreurs d'édition, comme dans les vers ci-dessous :

> Si dans les premiers jours que vous m'avez vu maître => naître
> Votre feu mal éteint avait voulu parfaites, => paraître
>
> Corneille, *Tite et Bérénice*, vers 499-500.

> Mais je suis seul encor, les amis de mon pères => père
>
> Racine, *Britannicus*, vers 323.

Chacune des étapes de l'analyse du vers (le mètre et la rime) a permis de repérer des erreurs d'édition et de les corriger. Mais elle a également mis en évidence quelques variations entre poètes. Par exemple, le verbe *hésiter* n'a pas le même statut métrique chez Corneille et Racine, ce qui reflète des états différents de la langue. Alors que chez Racine ce mot n'a pas de *h* aspiré, comme c'est le cas aujourd'hui en langue :

> Elle flotte, elle hésite ; en un mot, elle est femme
>
> Racine, *Athalie*, vers 876.

il en admet un chez Corneille :

> Et bien que sur le choix il semble hésiter
>
> Corneille, *Attila*, vers 459.

Dans une version initiale du *Menteur*, on trouvait également :

> Ne hésiter jamais, et rougir encor moins.
>
> Corneille, *Le Menteur*, vers 936.

Dans une version ultérieure, *hésiter* a été remplacé dans ce vers par *se brouiller*.

Autre exemple, *Harpie* a un *h* non aspiré chez Corneille, alors qu'en langue le *h* est aspiré (les dictionnaires n'ont cependant pas tous le même point de vue sur la question).

> Est celui qu'une harpie en fuyant laissa choir
>
> Corneille, *Médée*, vers 988.

Nous avons insisté un peu longuement sur les erreurs d'éditions, alors qu'elles sont finalement assez peu nombreuses, relativement au nombre de vers analysés (près de 80 000). Les corpus électroniques de la BNF, qui proviennent des éditions de l'INaLF, sont d'une qualité remarquable, surtout quand on les compare aux autres éditions électroniques que l'on peut trouver sur le réseau. Il est sans doute dommage que ces corpus soient si difficiles d'accès.

Après tous ces prétraitements, on dispose d'un ensemble de textes harmonisés prêts pour le métromètre. Quels types de vers trouve-t-on dans notre corpus : y a-t-il une forte hétérométrie, ou au contraire l'hégémonie d'un vers ?

II.3. PREDOMINANCE DE L'ALEXANDRIN CHEZ CORNEILLE ET RACINE

Toute pièce de Corneille et Racine est écrite intégralement en vers et ne laisse dans les dialogues aucune place à la prose. Qu'elles soient écrites en alexandrins n'étonnera personne. Cependant, elles offrent une place, certes mineure, à d'autres mètres. Au fil de cette extrême régularité alexandrine, l'apparition d'autres mètres est frappante tant pour l'auditeur que pour le lecteur et mérite une attention particulière.

La répartition métrique sur l'ensemble du corpus apparaît dans le Tableau 5. Les calculs ont été faits sur les vers reconstitués pour Racine, comme pour Corneille.

Nombre de positions métriques	Effectifs	%
3	21	0,0
4	26	0,0
5	36	0,0
6	190	0,2
7	111	0,1
8	1987	2,5
9	4	0,0
10	360	0,5
12	77186	96,6
Total	79921	100

Tableau 5. Répartition des vers de CORRAC selon le nombre de positions métriques

On a donc 96,6 % d'alexandrins sur le corpus, majorité tout à fait écrasante. Viennent ensuite, loin derrière, l'octosyllabe et le décasyllabe. On notera que les vers pairs représentent 99,9 % des vers, ce qui laisse une place infime aux vers impairs.

Les différences ne sont pas très sensibles entre Corneille et Racine. Racine, plus classique, accorde une place légèrement plus importante à l'alexandrin, et c'est principalement par la moindre utilisation de l'octosyllabe, le vers privilégié des stances, qu'il se distingue de Corneille (*cf.* Tableau 6).

	Corneille		Racine	
Nombre de positions métriques	**Effectifs**	**%**	**Effectifs**	**%**
3	21	*0,0*	0	*0,0*
4	15	*0,0*	11	*0,1*
5	33	*0,1*	3	*0,0*
6	138	*0,2*	52	*0,3*
7	60	*0,1*	51	*0,3*
8	1788	*2,9*	199	*1,1*
9	4	*0,0*	0	*0,0*
10	244	*0,4*	116	*0,6*
12	58679	*96,2*	18507	*97,7*
Total	60982	*100,0*	18939	*100,0*

Tableau 6. Diversité métrique chez Corneille et Racine

C'est principalement à cause de la variation métrique dans *Psyché* et *Agésilas* que la part des octosyllabes est plus forte chez Corneille que chez Racine (*cf.* Tableau 7).

Vingt-cinq pièces ont été exclusivement écrites en alexandrins, soit plus de la moitié. La pièce dans laquelle la variation métrique est la plus importante est *Psyché*, puisque tous les mètres répertoriés y sont représentés. *Agésilas*, souvent considérée comme une pièce en vers libres (*cf.* préface des œuvres complètes de Corneille, collection « L'Intégrale », Le Seuil), est finalement assez régulière en ne faisant alterner que des alexandrins et des octosyllabes (sauf trois décasyllabes). Dans les deux dernières pièces de Racine, les chœurs emploient d'autres mètres que l'alexandrin. Hormis dans *Psyché*, l'alexandrin est majoritaire dans toutes les pièces.

Dans les pièces largement dominées par l'alexandrin à rimes plates, les passages qui ne sont pas en alexandrins à rimes plates correspondent principalement à des stances ou à du discours rapporté. Pour la localisation précise de ces passages hétérométriques chez Corneille et Racine, on se reportera au Tableau 7 (les fonds grisés signalent les pièces ne présentant aucune hétérométrie).

CORNEILLE												
Nbre de positions	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**Total**	Vers hors alexandrins à rimes plates	Types (S. : stances)
Mélite									1823	1823		
Clitandre									1624	1624		
La Veuve	0	0	0	16	0	45	0	0	1922	1983	v. 393-442 v. 1093-1135	S. Philiste S. Clarice
La Galerie du palais	0	0	0	7	0	0	0	0	1791	1798	v. 983-1010	S. Célidée
La Suivante	0	0	0	0	0	35	0	0	1665	1700	v. 373-402 v. 1021-1040 v. 1661-1700	S. Florame S. Daphnis S. Amarante (alexandrins en rimes croisées)
La Comédie des Tuileries, 3e acte									400	400		
La Place royale	0	0	0	0	0	40	0	0	1489	1529	v. 142-178 v. 343-352 v. 721-732 v. 1157-1166 v. 1490-1529	S. Cléandre Lettres S. Angélique Billet S. Alidor (alexandrins en rimes croisées)
Médée	0	0	0	0	0	24	0	0	1604	1628	v. 1161-1208	S. Aegée
L'Illusion comique									1692	1692		
Le Cid	0	0	0	12	0	18	0	18	1792	1840	v. 291-350 v. 1565-1596	S. D. Rodrigue S. L'Infante
Horace									1782	1782		
Cinna									1780	1780		
Polyeucte	0	0	0	0	0	25	0	0	1788	1813	v. 1105-1160	S. Polyeucte
Pompée									1812	1812		
Le Menteur									1810	1810		
La suite du Menteur	0	0	0	0	0	27	0	0	1881	1908	v. 917-933 v. 939-948	Rêverie de Dorante
Rodogune	0	0	0	0	0	4	0	0	1840	1844	v. 1643-1646 v. 1663	Dernières paroles de Séleucus Reprise
Théodore									1882	1882		
Héraclius	0	0	0	0	0	44	0	0	1872	1916	v. 595-600 v. 1511-1553	S. Héraclius Billet de Maurice
Andromède	0	3	0	31	0	130	0	75	1528	1767	trop de variations	
Don Sanche d'Aragon									1830	1830		
Nicomède									1854	1854		
Pertharite, roi des lombards									1854	1854		
Œdipe	0	0	0	0	0	48	0	0	1962	2010	v. 605-610 v. 779-828	L'ombre S. Dircé
La Toison d'or	0	0	0	12	0	98	0	16	2109	2235	trop de var. v. 1480-1519	S. Médée

CORNEILLE												
Nbre de positions	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**Total**	Vers hors alexandrins à rimes plates	Types (S. : stances)
Sertorius									1920	1920		
Sophonisbe	0	0	0	1	0	4	0	0	1817	1822	v. 1591-1598	Billet Massinisse
Othon									1832	1832		
Agésilas	0	0	0	0	0	477	0	3	1642	2122	trop de variations	
Attila									1788	1788		
Tite et Bérénice									1774	1774		
Psyché	21	12	33	61	60	768	4	132	1023	2114	trop de variations	
Pulchérie									1758	1758		
Suréna									1738	1738		
RACINE												
Nbre de positions	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**Total**	Vers hors alexandrins à rimes plates	Types (S. : stances)
La Thébaïde	0	0	0	3	0	12	0	4	1497	1516	v. 393-396 v. 1203-1232	Oracle S. Antigone
Alexandre le Grand									1548	1548		
Andromaque									1648	1648		
Les Plaideurs									884	884	v. 809-810	2 vers en latin
Britannicus									1768	1768		
Bérénice									1506	1506		
Bajazet	0	0	0	1	0	1	0	5	1742	1749	v. 1135-1144 v. 1185-1192 v. 1267-1269	Lettre de Bajazet Lettre d'Amurat Reprise de la lettre de Bajazet
Mithridate									1698	1698		
Iphigénie	0	0	0	0	0	2	0	0	1794	1796	v. 57-62	Oracle
Phèdre									1654	1654		
Esther	0	9	3	33	33	96	0	70	1112	1356	v. 115-121 v. 132-154 v. 293-372 v. 713-825 v. 934-1015 v. 1200-1286	Chœur id. id. id. id. id.
Athalie	0	2	0	15	18	88	0	37	1656	1816	v. 311-371 v. 751-848 v. 1135-1175 v. 1187-1236 v. 1463-1509	Chœur id. id. id. id.
Total	21	26	36	190	111	1987	4	360	77186	79921		

Tableau 7. Diversité métrique selon les pièces

Pour pouvoir appliquer des méthodes de linguistique informatique, il faut que les corpus soient « propres » et bien balisés. C'est la condition première pour avoir des résultats satisfaisants. Le balisage porte sur les différentes catégories de texte (vers, prose, paratexte) pour que l'outil n'analyse pas une didascalie comme si c'était un vers ; qu'il retienne que tel vers apparaît dans telle scène, est prononcé par tel personnage… Les vers doivent être corrigés pour qu'il n'y ait pas de confusion, lorsque l'on explore les phénomènes rares, entre le choix fait par un auteur de transgresser ou de jouer avec les règles et une erreur d'édition. Les vers doivent être reconstitués, parce que même lorsqu'ils sont prononcés par plusieurs personnages, ils respectent la structure métrique de l'alexandrin. Mais cette reconstitution doit être signalée. Enfin, il nous a paru utile de signaler où se trouvaient les passages qui ne sont pas en alexandrins à rimes plates. Même s'il n'y en a que 3 %, il vaut mieux savoir les localiser quand on décide par exemple d'apparier un vers avec le suivant pour étudier la rime…

Mais la préparation du corpus n'est pas seulement une opération d'édition informatique. C'est aussi et surtout une plongée en profondeur dans le corpus, qui permet d'avoir un contact intime et prolongé avec sa matière textuelle informatisée. Cette connaissance en profondeur malgré son aspect laborieux, nous a paru utile en ce qu'elle donne une certaine intuition de la nature des problèmes rencontrés par les programmes informatiques. C'est enfin et surtout un inévitable coût à payer pour pouvoir appliquer des méthodes de linguistique informatique, que nous engageons ceux qui s'aventurent dans cette voie à ne pas sous-estimer dans leur planning.

III. LA RIME

Le vers alexandrin n'apparaît quasiment jamais seul[27]. Il ressemble à ceux qui l'entourent par le nombre de syllabes, par son organisation en deux segments métriques et par sa figure métrico-rythmique, comme nous le verrons dans le chapitre VII. Mais surtout il est lié à l'un ou à plusieurs d'entre eux par une terminaison commune, la rime.

Jamais isolé donc, l'alexandrin se voit toujours intégré dans un *groupement de niveau supérieur* défini par la rime. Dans le roman d'Alexandre, ce niveau supérieur est la *laisse*, groupement d'un nombre variable d'alexandrins qui ont tous la même rime. Dans la versification classique, plus précisément dans les alexandrins à rimes plates de Corneille et Racine, le groupement des vers en fonction de la rime se fait à deux niveaux : le premier niveau correspond au *couple* de vers défini par la rime, le second au *quatrain*[28] de vers défini par l'alternance en genre de la rime.

Une rime est dite féminine[29] si la dernière voyelle métrique du vers est suivie d'un *e* muet, masculine sinon. Cette distinction existait déjà en ancien français : une terminaison féminine ne pouvait jamais rimer avec une terminaison masculine. En revanche, l'alternance en genre des rimes, qui consiste à « différencier par le genre chaque nouvelle rime de la rime précédente » [Elwert, 1965, p. 83], n'est systématiquement appliquée qu'à

---

[27] Le poème d'Apollinaire, *Chantre*, constitué d'un unique alexandrin, est l'exemple le plus fameux de vers sans rime. Ce vers, qui a donné lieu à d'innombrables gloses, supporte une simple lecture métrique : le vers dit sa propre solitude.

Et l'unique cordeau des trompettes marines.

[28] Nous parlons de quatrain dans le sens où il y a un niveau qui regroupe les paires de vers avec une structure d'alternance régulière. Notons qu'Aroui [communication personnelle] conteste l'existence de ce niveau en raison de « la possibilité d'un nombre impair de distiques (dans une pièce) ».

[29] D'après Elwert [1965, p. 83], la dénomination rimes féminines, masculines daterait du XVe siècle. Elle est donnée par analogie avec le genre des mots : le genre féminin des adjectifs (*petit*/*petite*) est souvent marqué par un e muet.

partir du milieu du XVIe siècle ; elle est érigée en règle par Malherbe. Ainsi, les deux strophes suivantes (premières strophes de sonnets) ont-elles la même structure de rime *abba*, mais la première ne respecte pas l'alternance en genre.

O longs désirs ! O espérances vaines,	a	f
Tristes soupirs et larmes coutumieres	b	f
A engendre de moy maintes rivieres,	b	f
Dont mes deus yeus sont sources et fontaines :	a	f

Louise Labé, sonnet II.

Amour me tue, et si je ne veux dire	a	f
Le plaisant mal que ce m'est de mourir,	b	m
Tant j'ay grand peur qu'on vueille secourir	b	m
Le doux tourment pour lequel je soupire	a	f

Pierre Ronsard, *Le premier livre des amours*, sonnet XLV.

Voici un exemple d'enchaînement de rimes d'alexandrins à rimes plates :

	rime	genre
Dans un mois, dans un an, comment souffrirons-nous,	a	m
Seigneur, que tant de mers me séparent de vous ?	a	m
Que le jour recommence, et que le jour finisse,	b	f
Sans que jamais Titus puisse voir Bérénice,	b	f
Sans que de tout le jour je puisse voir Titus ?	c	m
Mais quelle est mon erreur, et que de soins perdus !	c	m
L'ingrat, de mon départ consolé par avance,	d	f
Daignera-t-il compter les jours de mon absence ?	d	f
...		

Racine, *Bérénice*, vers 1113-1117[30].

Cette règle de l'alternance, introduite tardivement, sera respectée jusqu'au XIXe siècle[31]. Dans *Les Fleurs du Mal*, on ne trouve que deux poèmes (*Ciel Brouillé* et *A une mendiante rousse*) qui ne respectent pas l'alternance (toutes les rimes sont masculines) :

[30] Quand l'édition n'est pas précisée, il s'agit des éditions électroniques d'après les éditions de Ch. Marty-Laveaux pour Corneille et de P. Mesnard pour Racine.

[31] Roubaud a cependant trouvé dans *La Célimène* de Rotrou, 5 strophes en rimes féminines (acte II, scène II) : « Un jour Filène / Au bord de la Seine, / Pour soulager son amoureuse peine, / Disoit voyant couler cette belle onde : / C'est ainsi que nos jours passent au monde ».

On dirait ton regard d'une vapeur couvert ;	a	m
Ton œil mystérieux (est-il bleu, gris ou vert ?)	a	m
Alternativement tendre, rêveur, cruel,	b	m
Réfléchit l'indolence et la pâleur du ciel. (...)	b	m

Baudelaire, *Les Fleurs du Mal*, L, Ciel brouillé.

Avant d'aller plus avant, ces exemples nous montrent qu'il faut, pour l'analyse de la rime descendre au niveau de la graphie des terminaisons des mots porteurs de la rime (et donc se déconnecter du niveau lexical) : le *genre de la rime n'a rien à voir avec le genre des mots*. *Dire* et *soupire* sont des constituants de rimes féminines, mais en tant que verbes ils n'ont pas de genre lexical. *Bérénice* et *finisse* correspondent tous deux à des rimes féminines ; *Bérénice* est un nom féminin, *finisse* n'a pas de genre. Ensuite, pour le nombre comme pour le genre, ce qui est vrai au niveau de la rime est indépendant de ce qui l'est au niveau lexical. Quand nous distinguerons les terminaisons plurielles des terminaisons singulières, ces termes ne seront employés que pour désigner la terminaison du mot et non le nombre du mot en tant que tel. Ainsi, *perdus* est bien un adjectif pluriel, mais *Titus* est un nom singulier ; pourtant ils riment ensemble car la terminaison des deux est en *us*. Enfin, dans le vers, rien n'impose que les mots rimant ensemble soient de la même catégorie syntaxique (rimes grammaticales ou catégorielles). Les règles de la rime, comme nous le verrons, concernent les terminaisons graphiques et phoniques et non pas des instances morpho-lexicales, même si celles-ci peuvent intervenir à un autre niveau.

On connaît les structures que la rime permet d'établir entre les vers : laisses, strophes, distiques, quatrains... Nous cherchons à identifier ce qui définit l'équivalence des fins de vers et l'alternance en genre sur le corpus CORRAC[32], autrement dit quels sont empiriquement les éléments nécessaires pour que la structure en paires d'alexandrins à rimes plates alternées soit reconnue ? Comme on va le voir, cette étude débouche sur la question de la nature même de la rime.

De nombreux travaux ont été consacrés à la rime. Tous les traités lui accordent une place importante et de nombreux dictionnaires de rimes ont

[32] Nous considérons, peut-être un peu hâtivement, que Corneille et Racine traitent la rime de la même manière et de manière constante tout au long de leur période d'écriture.

accompagné l'histoire du vers[33]. Notre approche de la rime est entièrement empirique et prétend simplement observer comment, sur un corpus donné, les vers riment ensemble. L'intérêt pour la pratique effective de la rime dans les vers est assez modérée. Premièrement, la rime a quelque chose d'évident. En réception, c'est l'identité phonique principalement qui permet l'identification des vers rimant ensemble. La reconnaissance de la rime en réception semble aller de soi pour tout lecteur (même chez les enfants). En production, c'est sans doute un peu moins le cas, comme l'atteste en particulier l'existence de dictionnaires de rimes. Deuxièmement, une analyse manuelle du contenu de la rime est difficile à réaliser : étant donnée la très grande variété des terminaisons possibles, il faut de très vastes corpus pour avoir un nombre suffisant d'occurrences de rimes similaires pour tirer des conclusions solides sur la rime. Par exemple, sur 9 000 vers du *Roman de Perceval*, nous n'avons compté que 76 occurrences de rimes en *ir*.

Pourtant, la nature de la rime débouche sur la question de la forme, et de la similitude des formes. Quelles conditions doivent être réunies pour que l'on puisse dire qu'un vers est lié à un autre par l'équivalence de sa terminaison ? A partir de quand peut-on dire que deux vers riment ensemble ? Cette équivalence que constitue la rime ne va pas de soi. Nous allons voir qu'elle fait intervenir des traits de langue divers (graphiques et phonétiques), avec une extension variable (nous le verrons à propos du rôle de la consonne d'appui). Le principe d'identité n'est pas trivial, car il ne s'agit pas d'identité absolue, mais d'une grande similitude graphico-phonétique. Le système est beaucoup moins figé que celui des positions métriques. Tout se passe comme si l'identité était admise dès lors que l'on reconnaît une masse critique de traits en commun, qui font intervenir des dimensions graphique, phonique, morpho-syntaxique et fréquentielle.

L'analyse empirique de la rime a donc une double visée. L'une est d'identifier ce qui crée l'identité entre fins de vers pour, à plus long terme, proposer un système automatique de notation des rimes qui permettrait de travailler sur la combinatoire des rimes d'une manière systématique (pas sur des alexandrins à rimes plates, bien sûr, mais sur des modes d'organisation

[33] On pourra se reporter par exemple aux livres de Grammont [*1908*, 1965, p. 33-41], d'Elwert [1965, p. 75-111], de Mazaleyrat [1974, p. 181-214], au *Dictionnaire des rimes* de Martinon [1905] et au traité de *Versification française* qui le précède. Billy [1984] recense différentes tentatives de nomenclatures de rimes. Morin [1993] propose une lecture du *Dictionnaire des rimes* de Lanoue.

strophiques plus complexes). L'autre, plus théorique, est de mieux comprendre la dialectique du même et du différent à l'œuvre dans la rime.

Nous avons procédé à un examen des rimes chez Corneille et Racine, qui nous permet de décrire en partie quelques phénomènes de la rime. Cette étude entièrement empirique, fondée sur l'examen d'un corpus cependant limité, donnera une vision exhaustive de la pratique de deux auteurs classiques. On pourra poser certaines hypothèses sur le fonctionnement de la rime à l'époque classique, qui devront être vérifiées sur d'autres corpus.

Après un bref examen du point de vue des métriciens sur la rime, nous décrirons la méthode utilisée pour construire le « RIMARIUM », c'est-à-dire la liste des couples de mots rimant ensemble dans CORRAC organisée par type de séquence finale. Ensuite, nous essaierons de formaliser quelques faits concernant la rime. La nature de la rime soulève deux questions. L'une porte sur sa nature : est-elle graphique, phonique, ou fait-elle intervenir les deux dimensions ? L'autre concerne son extension : quels sont les éléments (phonèmes ou graphèmes) nécessaires pour qu'il y ait rime et quels sont les facteurs qui définissent son extension ?

III.1. LA RIME SELON QUELQUES THEORICIENS

Dans cette section, nous examinerons le point de vue de quelques métriciens sur la rime : autour de la fonction de la rime, de sa nature sonore, du système d'équivalences et d'oppositions et de la nomenclature des rimes. Nous ne prétendons pas rendre compte de tous les travaux qui ont été faits dans le domaine, mais donner un bref aperçu de quelques positions[34].

III.1.1. La fonction de la rime

Trois fonctions peuvent être associées à la rime. Premièrement, comme nous l'avons vu dans le modèle métrique de l'alexandrin classique présenté dans le chapitre I, la rime organise les vers dans des structures de niveau

[34] Pour avoir une revue plus complète sur la rime nous renvoyons à Grammont [*1908*, 1965, p. 33-41], à Elwert [1965, p. 75-111], à Mazaleyrat [1974, p. 181-214], à Billy [1984], Cornulier [1995, p. 127-128 et 192-200].

supérieur. Dans le cas de l'alexandrin classique à rimes plates, la rime permet de regrouper les vers par paires qui riment ensemble, et par quatrain avec alternance en genre des rimes. Dans d'autres contextes de vers, la rime aura pour fonction de structurer des laisses, des strophes...La rime permet « d'assurer, tant qu'elle[s] se reproduit[sent], l'unité des séries où elle[s] constitue[nt] la fin de chaque vers » [Mazaleyrat, 1974, p. 75], elle contribue à la « constitution de superstructures » [Cornulier, 1982a, p. 508].

Deuxièmement, la tradition métrique française accorde à la rime une fonction démarcative : elle considère que la rime marque la fin du vers. Cette fonction démarcative serait d'autant plus forte pour le vers français que la mesure n'y est pas suffisamment marquée. Le vers français, défini par un nombre de syllabes et l'existence d'une césure, ne serait pas identifiable en tant que tel, et la rime viendrait compenser la faible structuration métrique (l'absence de pieds) Nous ne donnons qu'un exemple de cette position :

> [La rime] sert à accentuer et à marquer la fin des vers, pour délimiter nettement les vers, qui par ailleurs ne sont marqués que par le nombre de syllabes. Elle est d'autant plus nécessaire que la différence entre syllabes toniques et atones faiblit, et que disparaît l'élément rythmique.
>
> [Elwert, 1965, p. 76].

Nous renvoyons pour plus de détails aux références citées par Cornulier [1982a] dans un article où il réfute cette fonction de la rime, en considérant que la première fonction (constitution de superstructures incluant les vers) est suffisante pour justifier la rime. Remarquons au passage que pour Mazaleyrat [1974], la rime cumule les deux fonctions : elle assure la liaison entre vers mais elle peut aussi contribuer à perfectionner le vers.

Enfin Roubaud [*1978*, 1988, p. 93-94] considère que la rime « "désambiguise" l'appartenance des segments métriques à un vers ou à un autre », elle contribue à assurer l'unité du vers en tant que tel, en distinguant le second hémistiche du premier[35] : sans la rime, les alexandrins seraient plus facilement perçus comme des hexasyllabes[36]. La fonction de

---

[35] Évidemment, bien d'autres phénomènes qui apparaissent dans l'environnement de la césure et que nous aurons l'occasion d'examiner en détail assurent l'unité du vers alexandrin classique.

[36] Cornulier [1982a] rattache la fonction de désambiguisation des segments métriques proposée par Roubaud à la fonction démarcative. Nous préférons les distinguer car le fonction démarcative s'appuie sur la notion de mesure du vers, de rythme interne, tandis que la désambiguisation relève d'un point de vue exclusivement métrique.

désambiguisation des segments métrique n'empêche pas que la rime ait aussi pour fonction chez Roubaud de créer dans l'alexandrin à rimes plates une « macro-organisation à quarante-huit positions » [p. 94], les paires de rimes alternées.

Nous ne nous intéressons ici qu'à la fonction première de la rime, comme organisatrice de structures intégrant des vers. Par rapport au modèle métrique de l'alexandrin (*cf.* Figure 2), nous nous situons aux niveaux 4 et 5, là où se définissent les paires de vers rimant ensemble et les quatrains de vers avec alternance en genre des rimes.

III.1.2. La rime : une identité sonore ?

Voici quelques définitions de la rime :

> Homophonie de la voyelle finale et des éléments sonores qui la suivent.
>
> [Martinon, *1905*, 1962, p. 35].

> Homophonie non seulement de la dernière voyelle accentuée, mais en même temps de tout ce qui suivait cette voyelle.
>
> [Grammont, *1908*, 1965, p.34].

> Par rime on entend l'identité acoustique de la syllabe ou des syllabes finales des deux mots, identité qui comprend la dernière voyelle tonique et les sons qui la suivent.
>
> [Elwert, 1965, p. 75].

> La rime se définit par l'homophonie, entre deux ou plusieurs vers, de leur dernière voyelle tonique ainsi que de tous les phonèmes qui, éventuellement, la suivent.
>
> [Mazaleyrat, 1974, p. 184].

Ces quelques définitions indiquent une identité (homophonie) entre les fins de vers. Shapiro [1974, p. 506] propose de remplacer la notion d'identité phonique pour la rime par celle d'équivalence phonologique :

> La rime est la récurrence régulière, dans des positions se correspondant séquentiellement (syntagmatiquement), de sons équivalents phonologiquement (paradigmatiquement), leur équivalence s'exprimant en termes de valeurs de marque identiques.
>
> [Shapiro, 1974, p. 506].

Ces définitions semblent indiquer que cette équivalence commence à la dernière voyelle métrique et qu'elle englobe toute la suite : elles définissent l'existence d'un seuil unique dans toutes les situations. Contrairement à l'assonance, la rime porte sur la voyelle métrique et sur tout le segment qui suit.

Ces mêmes définitions insistent sur le caractère sonore de ces équivalences, bien que la question de la rime pour l'œil ait suscité de nombreuses contributions. Certains rejettent l'idée même que la graphie puisse jouer un rôle dans la rime. Ainsi, Martinon, puis Grammont, affirment catégoriquement :

> La rime étant faite pour l'oreille, il en résulte que des mots qui ont la même orthographe et une prononciation différente ne sauraient rimer ensemble. La rime *mer* et *aimer*, dite normande, a pu être autorisée autrefois par la prononciation.
>
> [Martinon, *1905*, 1962, p. 38].

> N'étant faite que pour l'oreille, la rime n'a pas à tenir compte de l'orthographe.
>
> [Grammont, *1908*, 1965, p. 39].

Dans une note, Martinon atténue légèrement son propos en reconnaissant, à propos de la distinction entre rimes féminines et masculines, qu'elle a perdu sa valeur phonique pour ne conserver que sa valeur graphique :

> Cette définition était d'autant plus fondée à l'origine que l'*e* muet se prononçait encore légèrement. (...) Pourtant la distinction classique s'est maintenue. C'est que, pour les vers comme pour l'orthographe, les mots ont une physionomie propre, à laquelle on est accoutumé, et qui en fait pour ainsi dire partie intégrante : on les voit en même temps qu'on les entend, et les adversaires les plus acharnés des conventions de la rime acceptent celles-là sans se douter que c'en est une.
>
> [Martinon, 1905, p. 35].

Elwert analyse quelques contraintes imposées par la rime pour l'œil, tout en considérant que « tenir compte pour la pureté de la rime des consonnes amuïes est un archaïsme » [Elwert, 1965, p. 99].

Mazaleyrat refuse de réduire le débat rime pour l'œil / rime pour l'oreille à des questions historiques, la rime pour l'œil n'étant que la trace d'une diction ancienne. Il redonne consistance au caractère graphique de la rime.

> Le principe de la rime pour l'œil ne se réduit pas à une illusion historique : c'est un élément effectif de l'esthétique du vers français traditionnel de forme classico-romantique.
>
> [Mazaleyrat, 1974, p. 197].

Le débat rime pour l'œil / rime pour l'oreille tient en grande partie aux incertitudes que nous pouvons avoir aujourd'hui sur la prononciation du français au XVIIe siècle : le débat n'est pas clos en dépit des avancées de la linguistique. Tandis qu'Elwert [1965, p. 56] considère que les consonnes finales se sont amuïes dès le XIVe siècle, Morin [1993, p. 112] que les

consonnes [r] et [s] sont « probablement réduites et en voie de disparition », Cornulier [1995, p. 199] considère que le *s* final n'avait pas disparu à l'époque de Malherbe. A nos incertitudes actuelles liées à la distance historique, il faut sans doute ajouter les variations de la prononciation selon les régions et l'origine sociale des locuteurs.

En ce qui concerne le *e* muet final, il semble s'être amuï entre le XVIe et le XVIIe siècles [Aroui, 1996]. L'alternance entre rime masculine et rime féminine, marquée graphiquement par un *e*, acquiert alors une dimension véritablement graphique. Ainsi, trouve-t-on dans *Andromaque :*

> Ne vous souvient-il plus, seigneur, quel fut Hector ?
> Nos peuples affaiblis s'en souviennent encor.
>
> Racine, *Andromaque*, vers 155-156.

> Je vous offre mon bras. Puis-je espérer encore
> Que vous accepterez un cœur qui vous adore ?
>
> Racine, *Andromaque*, vers 293-294.

Si la rime n'était que phonique, *Hector* pourrait bien rimer avec *encore*.

Il y a de nombreux cas où l'identité phonique n'autorise cependant pas la rime. La rime est-elle donc graphique ? Une identité graphique peut parfois autoriser la rime même si la prononciation est différente, comme c'est le cas pour la rime dite normande qui continue de faire rimer au XVIIe siècle *cher* avec *chercher*, alors que la prononciation de *chercher* a changé : dans le langage courant, le -r final s'est amuï sauf dans les monosyllabes [Elwert, 1965, p. 104]. Mais la graphie ne suffit pas, à moins d'introduire des règles d'équivalences entre graphèmes. Un examen un peu attentif montre qu'il faut, comme le propose Roubaud, considérer que la rime, et plus généralement le vers, est « écrit-oral » :

> Cela veut dire que le vers, qu'il soit déclamé ou lu, s'adresse toujours à un œil-oreille, même quand la page n'est pas visible dans certains de ses emplois, comme la récitation ; même si la voix semble muette dans d'autres (lecture).
>
> [Roubaud, 1986, p. 50].

Aroui [1996] pose aussi que l'équivalence phonétique et l'équivalence graphique sont constitutives de la rime prototypique. Nous essaierons, dans la description, d'identifier et de prendre en compte à la fois les dimensions phonétiques et graphiques de la rime.

III.1.3. Principes d'identité et d'opposition

Chez les divers théoriciens, les commentaires sur la rime sont extrêmement normatifs et s'organisent en deux catégories : à faire / à ne pas faire. Ils sont présentés d'une manière synthétique par Mazaleyrat en deux catégories : principes d'identité et principes d'opposition [1974, p. 181-214].

Les principes d'identité sont les suivants : l'homophonie et parfois même l'homographie des rimes, le respect des genres des terminaisons et l'alternance en genre. Les principes d'opposition que Mazaleyrat identifie sont : l'opposition lexicale (condamnation du mot rimant avec lui-même, des rimes sur des mots lexicalement apparentés), l'opposition grammaticale (condamnation de l'identité des marques de catégories, des mots de même série grammaticale, des désinences de même type), opposition sémantique (autorise l'identité des mots si les sens sont différents) et l'opposition de volume (qui privilégie l'assemblage de mots-rimes de longueur différente[37]). Ces principes d'opposition sont historiquement marqués et conviennent à la rime post-médiévale.

Morin propose une définition concise de ces deux catégories de principes :

> Deux terminaisons sont permises à la rime si elle obéissent à une double contrainte d'identité/distinction : identité phonique et distinction lexicale.
>
> [Morin, 1993, p. 108].

Autrement dit, la rime impose simultanément la ressemblance des signifiants (identité graphico-phonétique des terminaisons) et la distinction des signifiés (interdiction de l'identité des mots, des morphèmes grammaticaux et du sens)[38].

III.1.4. Les nomenclatures de rimes

Les quelques traités que nous avons observés proposent chacun leur nomenclature des rimes en essayant d'affiner les catégories et les frontières entre catégories. Ces traités du XXe siècle tentent de rendre compte du

---

[37] Ceci mériterait d'être démontré par un test sur la longueur des mots-rimes.

[38] Pour ce qui est de la catégorie syntaxique des mots-rimes, il ne s'agit pas d'une variable sur laquelle s'impose une norme : on compte 57 % de rimes de même catégorie sur CORRAC.

fonctionnement du vers classique. La nomenclature qui revient le plus souvent distingue : la rime pauvre, la rime suffisante et la rime riche. Nous renvoyons à l'article de Billy [1984] qui recense les principales nomenclatures de rimes proposées par la tradition métrique, celles de Grammont, de Suberville, Lubarsh, Suchier et Töbler. La qualification des rimes et la mise en place de nomenclatures remonterait à la Pléiade.

Le nombre de phonèmes communs est un des critères utilisé le plus souvent pour catégoriser la richesse des rimes.

Chez Elwert, la rime est pauvre si « l'harmonie des sons porte uniquement sur les voyelles accentuées » [1965, p. 84]. Il considère que Grammont a tort dans ce cas d'employer le terme d'assonance. En effet, pour Grammont [*1908*, 1965, p. 37], la rime de la voyelle nue doit être proscrite car elle se confond avec l'assonance. Dans les vers de Corneille et Racine, comme nous le verrons, les cas de rimes pauvres ne sont pas négligeables. La rime est suffisante si « l'harmonie porte sur la voyelle tonique et sur la consonne qui la suit » [Elwert, 1965, p. 85]. La rime est riche s'il y a « répétition non seulement de la même voyelle accentuée, mais aussi d'une même consonne précédente, dite consonne d'appui » [p. 85].

Pour Grammont, la catégorie de rime riche ne se définit pas dans l'absolu mais par rapport à la catégorie de rime suffisante :

> On appelle rimes *riches* celles qui présentent l'homophonie d'un élément de plus que ceux qui sont indispensables aux rimes *suffisantes.*
>
> [Grammont, *1908*, 1965, p. 38].

Malheureusement les exemples donnés montrent une méconnaissance de ce que sont les éléments indispensables à la rime. Ainsi Grammont considère que « *banni* et *fini* ne riment pas richement (...) mais *bannir* et *finir* riment richement (...) » [*1908*, 1965, p. 38]. Or chez Corneille et Racine, la consonne d'appui est pour les terminaisons en *i* comme pour les terminaisons en *ir* quasiment obligatoire : dans les deux cas, il s'agit de rimes appartenant à la même catégorie de rime suffisante.

Le nombre de phonèmes communs, absolu comme relatif, ne peut suffire à définir les catégories de rimes. Morin [1993, p. 112] en a l'intuition quand il cherche à identifier le « seuil », c'est-à-dire le point de départ de « l'identité phonique minimale des terminaisons », en particulier les cas où la consonne d'appui (« seuil prétonique consonantique ») est requise pour qu'il y ait rime. Pour lui, les voyelles orales brèves seules ou suivies de *r* ou *s* nécessitent la consonne d'appui. La longueur de la voyelle influe donc sur

le seuil : plus la voyelle est brève, plus la rime s'étend à gauche de la voyelle. Si cette distinction de longueur des voyelles est bien établie au XVIe siècle, elle ne l'est plus au XVIIe siècle. Par exemple, pour Morin, *brebis* et *esprits* (*i* bref) nécessitent un seuil prétonique, contrairement à *précis* (*i* long). Mais dans CORRAC, *précis* comme *esprits* riment tous deux avec *fils*. Nous retiendrons de cette position la notion de seuil de la rime variable selon les contextes.

Billy [1984, p. 69-72], après avoir exposé quelques uns des systèmes de nomenclature de rimes, s'appuie sur celle de Tobler [*1885*, 1972], pour proposer la nomenclature suivante :

a) Le segment phonologique commun (SPC) a une extension inférieure à celle des mots concernés :

1. la rime suffisante qui comprend la dernière voyelle accentuée et les phonèmes suivants
2. la rime riche : l'identité finale s'étend à la consonne d'appui ou à la seule liquide précédant la voyelle
3. la rime léonine : l'identité remonte jusqu'à l'avant-dernière voyelle métrique

b) Le SPC a une extension identique aux mots-rimes concernés :

4. la rime homonyme : où la rime coïncide avec les deux mots-rimes (*pas* / *pas*)
5. la rime équivoque : qui porte sur plusieurs mots (*abondance* / *a bon danse*)

c) Le SPC a une extension identique à celle d'un mot ou d'une suite de mots constitutif(s) d'une occurrence rimique, mais inférieure à celle du mot ou de la suite de mots constitutif(s) d'une occurrence proche.

6. la rime paronyme : *amis* / *mis*
7. la rime identique : *il a droit* / *sera droit*
8. la rime dérivée : *complaint* / *son plaint*

Si cette nomenclature est adaptée à la description de la rime médiévale, elle trouve ses limites à l'époque classique où pour Billy, la notion de rime riche, telle qu'elle est décrite dans les nomenclatures, perd son sens :

> Le classicisme a systématiquement favorisé, dans la foulée de la Pléiade, un équilibre compensatoire entre *fréquence* et *richesse* des rimes (entre autres contraintes) dont un modèle théorique ne peut rendre compte sans une complexité qui le rendrait inexploitable dans le cadre d'une nomenclature.
>
> [Billy, 1984, p. 74].

Pour la rime post-médiévale,

> La caractérologie des types [de rime] revêt un caractère hétéroclite, des critères graphiques, parfois grammaticaux, etc. intervenant à côté de critères phoniques et statistiques qui prévalent et constituent le complexe central à travers lequel est

> recherchée cette valeur moyenne (...) où la pauvreté phonétique est rachetée par la rareté et inversement.
>
> [Billy, 1984, p. 75].

Billy propose une définition dynamique de la rime post-médiévale, « par un "centre de gravité" autour duquel s'équilibre interactivement un ensemble de contraintes inextricables dont l'application modifie constamment la physionomie de ses applications. » [Billy, 1984, p. 75].

Notre travail constitue une tentative de mise en œuvre du programme proposé par Billy, bien que nous l'ayons fait avant d'avoir eu connaissance de son article.

Il nous semble que pour pouvoir proposer une nomenclature des rimes, il faut avant tout avoir évalué ce qu'était *pour chaque terminaison* le contenu minimal indispensable pour qu'il y ait rime. C'est par rapport à ce contenu minimal qu'éventuellement pourra être évaluée la richesse de la rime. Nous conserverons l'idée qu'une rime riche fait intervenir au moins un phonème de plus que la rime suffisante, mais nous montrerons que cette distinction n'est applicable que dans un contexte *local*, pour un type de terminaison donné. Ainsi, selon les voyelles, l'extension de la rime pour qu'il y ait rime riche pourra être différente.

Nous posons que la rime suffisante est définie par la situation majoritaire pour un type de terminaison donné, la rime riche, respectivement pauvre, par l'intervention d'une unité (lettre ou phonème) de plus, respectivement de moins, dans l'identité par rapport à la rime suffisante. Autrement dit, si dans un corpus, *devant* rime avec *grand*, *vivant*, *avant* et *levant*, la rime sera pauvre dans le premier cas, suffisante dans les deux suivants (situation majoritaire : deux cas sur quatre) et riche dans le quatrième : le *e* muet à gauche du *v* intervenant dans la rime. Cet exemple est caricatural et ne permet pas, en raison du nombre limité de cas, de tirer la moindre conclusion. Cette définition locale de la richesse de la rime implique naturellement un examen systématique d'un très grand corpus de vers.

Nous avons donc le sentiment que les manuels, en proposant des nomenclatures en trois ou quatre catégories, n'ont pas été confrontés à une quantité suffisante de vers. Un modèle de la rime sera forcément bien plus complexe car l'extension de la rime varie selon les environnements et selon différents types de critères.

Nous chercherons donc à identifier plus précisément les éléments indispensables pour qu'il y ait rime chez Corneille et Racine en fonction des contextes locaux, et en particulier à identifier le statut pas toujours clair de la consonne d'appui.

III.2. EXAMEN EMPIRIQUE DE LA RIME

On appelle mot-rime[39] le mot qui supporte la séquence qui permet d'établir un écho avec un autre vers. C'est la nature de cet écho, oral-écrit, qui nous occupe.

L'objectif est de constituer des groupes, ou plus précisément des réseaux de mots-rimes sur la base d'attestations réelles. L'examen de ces groupes permettra d'abstraire des rimèmes, c'est-à-dire les segments graphico-phonétiques associés à chaque groupe (certains groupes ayant le même rimème pourront ainsi être rapprochés). Ensuite, cet examen permettra de définir précisément les conditions nécessaires pour que deux vers puissent rimer ensemble.

Pour cela, il a fallu apparier les vers rimant ensemble, construire les groupes de mots-rimes, examiner les groupes de mots-rimes à la lumière de la structure du réseau et inférer quelques résultats à partir de l'examen empirique.

Une des limites majeures de ce travail tient au fait que nous avons travaillé sur un corpus dans une orthographe modernisée (les éditions d'époque de Corneille et Racine ne sont pas disponibles sur support informatique). Nous avons régulièrement eu recours à des éditions imprimées dans l'orthographe d'époque pour vérifier les résultats, mais il aurait été plus satisfaisant de travailler directement sur les éditions d'origine, étant donné l'importance de la dimension graphique.

[39] Ce terme apparaît pour la première fois dans la *Description du projet* [Roubaud, 1979, p. 48].

III.2.1. Identification des mots-rimes et de leur fréquence

Notre hypothèse de travail est que la rime ne peut être portée que par le dernier mot du vers, qu'elle ne peut s'étendre au mot précédent[40]. Cette hypothèse n'est quasiment jamais contredite dans le vers de Corneille et Racine. L'identification du mot-rime et de sa fréquence ne présente *a priori* aucune difficulté. Il s'agit de repérer le dernier mot de chaque vers et ensuite de compter le nombre d'occurrences. Trois difficultés sont cependant apparues pour repérer automatiquement les mots-rimes.

L'identification du mot-rime est complexe quand le vers est prononcé par plusieurs personnages, puisqu'il n'y a plus correspondance entre fin de ligne et fin de vers. C'est d'ailleurs principalement dans le but d'étudier la rime que nous avons reconstitué tous les vers découpés dans le corpus, comme cela a été présenté dans le chapitre II.

Texte de départ
ATALIDE.
Bajazet !
ACOMAT.
Que dis-tu ?
OSMIN.
Bajazet est sans vie.
L'ignoriez-vous ?
ATALIDE.
Ô ciel !
OSMIN.
Son amante en furie,

[40] Cette hypothèse n'est acceptable que pour le vers classique où les cas d'extension de la rime au-delà du dernier mot (à gauche) sont très rares. Prendre le dernier mot du vers comme unité porteuse de la rime en général est insuffisant, comme l'atteste la pratique de la rime équivoque (*cf.* nomenclature de Billy [1984]). Il existe bien des exemples de vers, en particulier au Moyen-Age où la rime porte sur plusieurs mots : *Preservez moy de faire j**amais ce** / Vierge portant sans rompure encourir / Le sacrement qu'on celebre a l**a messe**.* (Ballade de François Villon, in J. Roubaud [1998, p. 130]) ; ou bien, *En ce vergier ou le pied et l'on**gle ay** / La guerre avoit defloré maint **bouton** / Sus ung fauveau fus je le descen**glay** / Chascun le fiert et hors le de**boute on*** (Ballade de Jean Molinet, in J. Roubaud [1998, p. 196]).

Vers reconstitués
Bajazet ! / Que dis-tu ? / Bajazet est sans vie.
L'ignoriez-vous ? / Ô ciel ! / Son amante en furie,

Racine, *Bajazet*, vers 1693-1694.

Ensuite, l'édition modernisée respecte avec des degrés de précision divers les contraintes de la rime et ne prend pas toujours la peine d'introduire les variantes orthographiques qu'impose la rime. Ainsi *dise*, selon le contexte de la rime qui lui fait écho, doit s'écrire *die* ou *dise*....

Et toutefois, madame, il faut que je vous die
Qu'un trône est plus pénible à quitter que la vie :

Racine, *La Thébaïde*, vers 731-732.

Rien ne peut-il, seigneur, changer votre entreprise ?
Vous ne répondez point.
ANTIOCHUS.
Que veux-tu que je dise ?

Racine, *Bérénice*, vers 123-124.

Dans de nombreux cas, ces contraintes ne sont pas respectées dans l'édition électronique.

C'est ainsi également que dans l'édition des *Plaideurs*, on trouve le mot *hasard* orthographié *hasard*, alors qu'il doit être orthographié *hazar*, variation orthographique que s'était autorisée Racine, et qui est bien reproduite dans l'édition de Pierre Mélèse. Cette licence orthographique redouble ici une licence métrique, la présence d'une conjonction de coordination en douzième position (*cf.* chapitre VII) :

Semble s'être assemblé contre nous par hasar :
Je veux dire la brigue et l'éloquence. Car
D'un côté le crédit du défunt m'épouvante ;
Et de l'autre côté l'éloquence éclatante
De maître Petit Jean m'éblouit.

Racine, *Les Plaideurs*, vers 730-731.

C'est dans le cadre de l'analyse de la rime que nous avons corrigé beaucoup d'erreurs d'édition.

Errante, et sans dessein, je cours dans ce Palois.=> Palais[41]
Ah ! ne puis-je savoir si j'aime, ou si je hais ?

Racine, *Andromaque*, vers 1395-1396.

Une fois ces questions réglées, nous avons cherché à vérifier que la rime ne porte effectivement que sur le dernier mot du vers, ce qui sous-entend que la rime ne peut porter sur plus d'un mot.

Le dernier mot de chaque vers a été repéré systématiquement. En traitement automatique, le mot est une chaîne de caractères comprise entre deux séparateurs (le point, la virgule, l'espace). Généralement, l'apostrophe et le tiret sont également considérés comme des séparateurs. Ainsi dans *aujourd'hui*, le mot-rime sera *hui*, dans *l'air*, ce sera *air*...

Cette option est apparue inadéquate pour le pronom personnel *je*. Dans les vers suivants, *je* ne définit pas la rime qui porte aussi sur le mot précédent, le *e* muet marque l'alternance en genre des rimes, mais ne constitue jamais le noyau vocalique de la rime elle-même.

L'oserai-je, grands dieux, Ou plutôt le pourrai-je,
L'amour, dès qu'il le veut, se fait un privilège,

Corneille, *Suréna*, vers 145-146.

PAULINE
De grâce, permettez…
FELIX
Laissez-nous seuls, vous dis-je :
Votre douleur m'offense autant qu'elle m'afflige.

Corneille, *Polyeucte*, vers 989-990.

Pour le pronom *je*, nous avons donc considéré comme mot-rime l'ensemble verbe-*je*. C'est la première exception à notre hypothèse que la rime ne porte que sur le dernier mot du vers. Les autres pronoms personnels postposés peuvent au contraire être le noyau de rime.

En dehors du pronom *je*, nous n'avons rencontré qu'un cas de rime s'étendant au-delà du dernier mot. Le problème ne se pose de fait que pour les mots monosyllabiques à initiale vocalique : pour ceux-là, on pouvait se demander si le mot précédant intervenait dans la rime. Les mots concernés étaient dans notre corpus : *il*, *on*, *or*, *es*, *air*. Pour les quatre premiers mots, la rime ne dépasse pas le mot, comme le montre l'exemple suivant :

---

[41] La graphie *oi* était prononcée /ɛ/ au XVIIe siècle. Nous aurions pu garder cette graphie, mais dans l'édition originale *palais* est orthographié avec *ai*.

Je te fais plus parfait mille fois que tu n'es :
Ton feu ne peut aller au point où je le mets ;

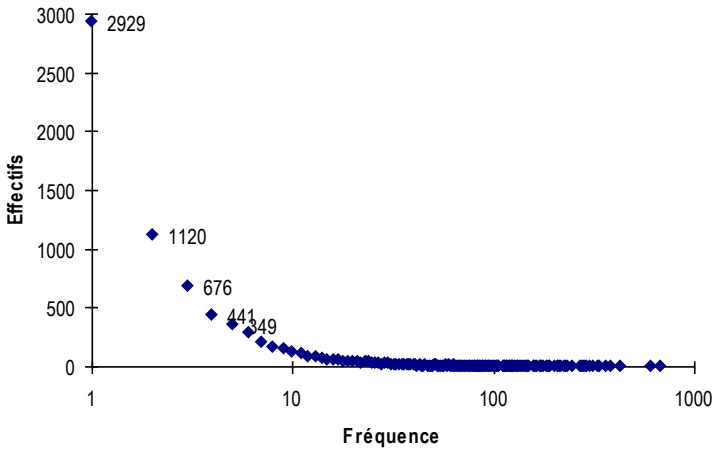
Corneille, *La Place Royale*, vers 115-116.

Pour les vers suivants, nous ne savions pas *a priori* si *air* était suffisant pour constituer la rime ou si l'appui du *l* était indispensable. Dans le doute, on a considéré *air* comme mot-rime :

La foi que j'en reçus s'en est allée en l'air,
Et ces traits de sa plume osent encor parler,

Corneille, *Mélite*, vers 913-914.

Au total, on a recensé 8 330 mots-rimes différents sur près de 80 000 vers que comprend le corpus normalisé, en comptant alexandrins et autres vers, soit un nombre moyen d'occurrences de 10,4. La courbe de répartition des fréquences de ces mots-rimes apparaît dans la Figure 4.



Clef de lecture : 2 929 mots-rimes dans le corpus CORRAC, ont une fréquence de 1, ce qui correspond à 37% des occurrences de mots-rimes, 1 120 ont une fréquence de 2...

Figure 4. Courbe de répartition des mots-rimes

Très peu de mots-rimes sont très fréquents, mais même le plus fréquent ne parvient pas à atteindre 1% des occurrences. Ils sont présentés dans le Tableau 8. En cela, la répartition des fréquences des mots-rimes a la même allure que celle des mots dans les corpus de textes. En effet, la distribution des fréquences dans un corpus suit la loi de Zipf [1934], d'après le nom de l'auteur qui, le premier, a mis en évidence les caractéristiques de la distribution des fréquences dans les corpus de textes : plus la fréquence d'occurrence augmente, plus le nombre de mots ayant cette fréquence

diminue. Dans un corpus, on a beaucoup de mots très peu fréquents, et peu de mots très fréquents. Les mots les plus fréquents correspondent en général à des mots-outils. Alors que nous avons vu que les mots-outils étaient particulièrement rares à la rime, et en tout cas qu'ils acquéraient une valeur substantive (*cf.* chapitre VII), quelques uns apparaissent parmi les mots-rimes les plus fréquents : *vous*, *moi*, *lui*, *elle*.

Mot	**Fréquence**	**% occurrences**	**Mot**	**Fréquence**	**% occurrences**
vous	685	0,98	empire	230	0,33
moi	610	0,87	courage	227	0,32
pas	428	0,61	jour	218	0,31
vie	384	0,55	époux	213	0,30
père	363	0,52	doux	211	0,30
même	338	0,48	flamme	209	0,30
yeux	334	0,48	aime	203	0,29
amour	316	0,45	colère	192	0,27
foi	301	0,43	crime	192	0,27
âme	291	0,41	dire	191	0,27
lui	288	0,41	courroux	189	0,27
roi	283	0,40	rien	183	0,26
cœur	274	0,39	mort	178	0,25
gloire	273	0,39	haine	176	0,25
elle	246	0,35	maître	176	0,25
frère	236	0,34	rois	138	0,20

Clef de lecture : le mot *vous* apparaît 685 fois à la rime sur CORRAC, il représente à peine 1 % des occurrences de mots-rimes.

Tableau 8. Les mot-rimes les plus fréquents sur CORRAC

Les couples de mots-rimes les plus fréquents sont :

Mot-rime 1	**Mot-rime2**	**Freq**
foi	moi	131
doux	vous	110
jour	amour	106
aujourd'hui	lui	104
roi	moi	88
courroux	vous	85
âme	flamme	85
aime	même	81
envie	vie	80
moi	roi	79
même	aime	79
autre	vôtre	78
époux	vous	72

III.2.2. Constitution des paires de vers rimant ensemble

Afin de pouvoir constituer des réseaux de mots *pouvant* rimer ensemble, il faut au préalable repérer quels sont les vers qui riment ensemble. *Manuellement*, bien que très longue, cette investigation ne pose pas de difficultés que ce soit pour des rimes plates ou même pour des combinaisons plus complexes, comme les rimes alternées ou embrassées. En effet, dans le vers classique, l'écho de la rime - la terminaison identique la plus proche -, ne se trouve jamais à plus de trois vers et il n'est séparé au maximum que par une autre rime. Cela tient à deux phénomènes. Tout d'abord, le nombre de formules de rimes est limité par le respect de l'alternance en genre des rimes. Ainsi, les formules du type *abcabc* n'apparaissent pas, car elles contredisent la règle d'alternance : *c* et *a*, rimes différentes sont juxtaposées alors qu'elles sont de même genre. Ensuite, le vers classique français manifeste une nette préférence pour les formules de rimes où l'écho n'est séparé de la rime que par un autre timbre[42]. Si le repérage *manuel* des vers rimant ensemble est simple, en revanche, le repérage *automatisé* est loin d'être trivial, tant que l'on n'a pas déterminé les règles d'équivalence graphique et phonique de chaque type de rime.

Notre corpus offre un cadre simplifié pour étudier la rime puisque qu'il est essentiellement constitué de vers à rimes plates avec alternance de rimes féminines et masculines.

```
aa bb cc dd...
ff mm ff mm...
```

Il suffit donc d'apparier le premier vers avec le second, le troisième avec le quatrième...

Dans la pratique, plusieurs éléments s'opposent à ce simple appariement. Tout d'abord les changements de tours de parole qui interviennent à l'intérieur d'un vers : nous avons résolu cette difficulté en reconstituant les vers.

---

[42] C'est ce que Cornulier [1995, p. 267] désigne comme Règle des deux couleurs : « Une terminaison de vers n'est jamais séparée de son plus proche écho par plus d'un type de terminaisons ». Cette règle fonctionne pour le vers où les dispositions des rimes suivent les trois modèles (rimes plates, embrassées, alternées).

Ensuite, nous avons vu dans la présentation du corpus (*cf.* chapitre II) qu'il existait une certaine variation métrique dans les pièces étudiées et que la majorité d'entre elles n'étaient pas *exclusivement* en alexandrins. Or la variation métrique s'accompagne le plus souvent d'une variation dans la disposition des rimes. On trouve de plus quelques passages en alexandrins avec des rimes alternées[43]. Ainsi, si au fil d'un passage en rimes plates interviennent quelques rimes embrassées, l'appariement avec le vers suivant est perturbé. Si au lieu d'une rime sur deux vers, la rime porte sur trois, tout le système est décalé.

Un corpus de travail a été construit en mettant « à plat » à la main tous les vers des formules de rimes alternées ou embrassées dans l'unique but de construire les groupes de mots-rimes. Ce travail a été fait pour toutes les pièces, y compris pour *Andromède*, la *Toison d'or* et *Agésilas* où rimes alternées et embrassées sont fréquentes. Seule *Psyché* a échappé à cette réorganisation (par épuisement et parce que cette pièce n'entre pas toujours dans le corpus des œuvres de Corneille). La mise à plat a été plus difficile pour les pièces où domine très largement l'alexandrin à rimes plates. Il fallait procéder à un examen attentif pour repérer toutes les variations dans la disposition des rimes : la localisation des zones sans rimes plates est indiquée dans le Tableau 7 du chapitre II.

La réorganisation du corpus en rimes plates, faite à la main, était entachée de quelques erreurs. Pour améliorer l'appariement des vers, des critères de contrôle supplémentaires ont été introduits.

Dans un premier temps, les vers n'ont été appariés qu'à condition que les phonèmes vocaliques finals transcrits par le métromètre (transcription sur la base de la prononciation actuelle) soient identiques. Ceci a fait apparaître de nombreuses erreurs d'appariement pour différentes raisons :

Premièrement, deux voyelles dont la transcription phonétique est différente peuvent rimer ensemble. Chez Corneille et Racine, *flamme* rime avec *âme* (*a* antérieur / *a* postérieur), *parole* avec *rôle* (*o* ouvert / *o* fermé), mais aussi *arracher* avec *cher* (rime dite normande), *sûr* avec *possesseur* (rime dite gasconne)... Voici quelques exemples :

---

[43] *Cf.* Stances d'Alidor dans *La place royale,* vers 1490-1529.

Mais ce flatteur espoir qu'il rejette en mon âme
N'est qu'un doux imposteur qu'autorise ma flamme,

Corneille, *Mélite*, vers 17-18.

Mais prends garde surtout à bien jouer ton rôle :
Remarque sa couleur, son maintien, sa parole ;

Corneille, *Mélite*, vers 575-576.

Cherche en ce désespoir qui t'en veuille arracher
Quant à moi ton trépas me coûterait trop cher.

Corneille, *Mélite*, vers 857-858.

Je veille, déloyal : ne crois plus m'aveugler ;
Au milieu de la nuit je ne vois que trop clair :

Corneille, *L'Illusion comique*, vers 1384-1385.

Je veux prendre un moyen et plus court et plus seur (sûr)
Et sans aucun péril t'en rendre possesseur.

Corneille, *La Place royale*, vers 703-704.

De sa possession je me tiens aussi seur
Que tu te peux tenir de celle de ma sœur.

Corneille, *Mélite*, vers 797-798.

A vrai dire sur la question du *a*, nous ne pouvons dire si le trait antérieur/postérieur est distinctif pour la rime. L'évolution de l'orthographe rend difficile l'accès à ce que pouvait être la prononciation de l'époque. Dans l'édition de P. Mélèse, on trouve :

L'interest du public agit peu sur mon ame,
Et l'amour du païs nous cache une autre flâme

Racine, *La Thébaïde*, vers 279-280.

Quoy ma grandeur seroit l'ouvrage d'une femme ?
D'un éclat si honteux je rougirois dans l'ame

Racine, *La Thébaïde*, vers 1123-1124.

Si l'on considère que l'accent circonflexe marque un changement de timbre de la voyelle (recul du point d'articulation), il y a une inversion entre l'orthographe de Racine et la nôtre, puisque chez lui le *a* de *ame* est antérieur, celui de *flâme* postérieur. L'orthographe est instable et n'est pas « phonographique » : les sonorités ne nous sont plus accessibles. Nous n'avons qu'une présomption sur l'équivalence entre *a* antérieur et postérieur pour la rime. Notre seule certitude est que la distinction entre les deux timbres de *a* telle qu'elle fonctionne aujourd'hui n'est pas pertinente pour le vers classique.

Deuxièmement, deux voyelles phonétiques similaires ne riment pas ensemble si les mots-rimes ne sont pas de même genre et nombre et s'ils n'admettent pas des séquences équivalentes après la voyelle. Or, en ne

s'appuyant que sur la voyelle métrique pour les appariements, on risque d'associer des vers ayant la même voyelle finale mais pas la même séquence finale.

Troisièmement, en général, la rime fait intervenir la semi-voyelle qui précède la voyelle de la rime, en particulier, on ne peut dissocier la voyelle de la semi-voyelle pour *oi* et *oin* et rarement pour *ui*. En atteste le fait qu'il n'y a pas de diérèse pour ces groupements vocaliques (*cf.* chapitre IV.1). Les autres semi-voyelles associées aux voyelles hautes *i*, *u*, *ou* n'interviennent pas systématiquement dans la rime. Or, la transcription phonétique de la dernière voyelle n'inclut pas forcément les semi-voyelles. C'est ce travail sur la rime qui nous a d'ailleurs conduit à affiner la définition d'une voyelle métrique (*cf.* chapitre V), en associant dans les trois cas *oi*, *oin* et *ui* la semi-voyelle à la voyelle métrique.

Nous avons renoncé à utiliser la voyelle comme critère de contrôle pour adopter un autre critère plus sûr : le genre de la rime. Un vers a été apparié avec le suivant à condition que la rime soit de même genre.

Pour cela, il a fallu caractériser de manière systématique le genre de chacun des mots-rimes.

Pour qu'une rime soit féminine, il faut au moins que le mot s'achève par *e*, *es* ou *ent*. Une terminaison graphique en *e* ou en *es* indique toujours une rime féminine, sauf pour la forme *es* (deuxième personne du présent) de l'auxiliaire *être*. Pour la terminaison *ent*, la rime est toujours féminine si, dans la transcription phonétique, la douzième position n'est pas occupée par [ɑ̃]. Lorsqu'elle est occupée par [ɑ̃], et que l'on a des séquences du type *anCent* ou *enCent* la rime est toujours féminine. Dans tous les autres cas, on a une rime masculine.

pièce	**vers**	**voyelle 12**	**terminaison**	**genre**
Mélite	Au cœur de cette belle aucun embrasement	ɑ̃	ent	m
Mélite	Ces âmes du commun n'ont pour but que l'argent	ɑ̃	ent	m
Mélite	Du moins les compliments, dont peut-être ils se jouent	u	ent	f
Mélite	Sont des civilités qu'en l'âme ils désavouent	u	ent	f
Le Cid	Ils abordent sans peur, ils ancrent, ils descendent	ɑ̃	ent	f
Le Cid	Et courent se livrer aux mains qui les attendent	ɑ̃	ent	f

***Tableau 9. Genre des terminaisons en* ent**

Dans 86 % des cas sur CORRAC, la séquence *ent* en fin de vers correspond au phonème [ɑ̃] et indique une rime masculine.

Une fois que le genre de la rime a été affecté à chaque vers, on constitue un fichier de paires rimées en appariant chaque vers avec le suivant sous la seule contrainte que le genre de la rime soit le même. Cet appariement a été refait chaque fois que nous découvrions une erreur d'édition, une erreur dans l'élimination des vers autres que les alexandrins à rimes plates ou dans le genre de la rime.

III.2.3. Constitution de groupes et de réseaux de mots-rimes

Une observation *rapide* de la rime semble montrer que si le mot A rime avec B, B peut aussi rimer avec A ;

Ô ciel, qui malgré moi me renvoyez au jour. A
Ô sort, qui redonniez l'espoir à mon amour, B
Corneille, *Polyeucte*, vers 423-424.

Mon vrai crime est ma gloire, et non pas mon amour : B
Je l'ai dit, avec elle il croîtra chaque jour ; A
Corneille, *Suréna*, vers 1651-1652.

et que si A rime avec B et B rime avec C, A peut rimer avec C.

C'est... je l'aurais nommé mille fois en un jour : A
Que ma mémoire ici me fait un mauvais tour ! B
Corneille, *La Veuve*, vers 1425-1426.

Sitôt que j'aurai su quel est ce mauvais tour. B
PHILANDRE.
Sache donc qu'il ne vient sinon de trop d'amour. C
Corneille, *Mélite*, vers 255-256.

Tu serais incivil de la voir chaque jour A
Et ne lui pas tenir quelques propos d'amour ; C
Corneille, *Mélite*, vers 47-48.

Pour la construction des ensemble de mots-rimes, nous posons comme hypothèse que la relation « rime avec » est une relation symétrique et transitive. Si *hasard* rime avec *regard*, l'inverse est possible. Si *hasard* rime avec *regard*, qu'ailleurs *regard* rime avec *départ* alors *hasard* rime potentiellement avec *départ*. En revanche, la relation n'est pas réflexive. Dans le vers classique, un mot ne rime jamais avec lui-même, ce qui a été érigé en règle ; et si par hasard on trouve le même signifiant à la rime, il correspond forcément à deux signifiés différents. Ainsi, *pas* et *point*

adverbes riment respectivement avec les noms *pas* et *point* et ne sont admis que parce qu'il s'agit d'homonymes dont les sens et les fonctions grammaticales sont très éloignés.

> BRITANNICUS.
> (...) Ah ! S'il vous avait dit, ma princesse, à quel point...
> JUNIE.
> Mais Narcisse, seigneur, ne vous trahit-il point ?
>
> Racine, *Britannicus*, vers 1533-1534.

Partant de cette hypothèse, si quelque part le mot *hasard* rime avec *regard*, qu'ailleurs *regard* rime avec *départ* et qu'enfin, *départ* rime avec *tard*, on pourra conclure que ces mots {*hasard, regard, départ, tard*} constituent ce que nous appellerons un groupe de mots-rimes.

Nous avons donc entrepris de construire tous les groupes de mots-rimes à partir de nos couples de vers rimant ensemble. Partons du mot-rime le plus fréquent *vous*. Il sera le cœur du premier groupe de mots-rimes. On recherche tous les couples de vers où il apparaît à la rime. Alors, tous les mots avec lequel il rime sont affectés à ce même groupe. Ensuite, on recommence l'opération pour chacun des mots trouvés, jusqu'à épuisement et puis l'on passe au second groupe défini par le second mot le plus fréquent et cela jusqu'à épuisement des mots-rimes[44].

En d'autres termes, notre travail a consisté à rechercher les composantes connexes dans un graphe où les nœuds sont les mots-rimes et les arrêtes correspondent à la relation « rime avec ».

> Une composante connexe est une partie du graphe telle qu'il y ait un chemin entre deux nœuds quelconques. Une clique est un graphe où chaque nœud est relié à tous les autres par une arrête.
>
> [Fabre et Habert, 1998].

Chaque composante connexe correspond à un groupe de mots-rimes.

Comme l'examen des résultats des groupes de mots-rimes révélait régulièrement des erreurs typographiques ou des erreurs d'appariement des

---

[44] Heureusement, une fois que les mots-rimes d'un couple sont classés, ils sont éliminés du corpus parcouru. Le programme écrit avec SAS, qui n'aurait sans doute jamais fonctionné sans l'écriture élégante proposée par Julie Micheau, est particulièrement long, il tourne plus de huit heures sur une station de travail SPARC 5. Comme nous le fait remarquer François Yvon, il existe sans doute des algorithmes connus et performants pour construire des classes d'équivalence de ce type.

vers rimant ensemble, les opérations ont été recommencées un grand nombre de fois.

Au final, on obtient 867 groupes de mots-rimes distincts, dont 415 sont des couples de mots qui n'apparaissent qu'une fois dans le corpus. Chacun de ces groupes de mots-rimes a été examiné manuellement et nous avons abstrait de la liste des mots pouvant rimer ensemble le rimème, c'est-à-dire la séquence graphico-phonétique commune à tous les mots du groupe. Le rimème est une séquence de phonèmes et de graphèmes unique pour un type de rime spécifique, qui comprend : la consonne d'appui (ou un tiret s'il n'y en a pas), le noyau vocalique de la rime, la consonne ou le groupe consonantique qui suit (un tiret s'il n'y en a pas) et la terminaison qui contrairement aux trois premiers éléments est exprimée par des graphèmes et non par des phonèmes. Il y a moins de rimèmes (540) que de groupes de mots-rimes (812) en raison du mode de construction empirique des groupes de mots-rimes : certains groupes n'ont pas été regroupés simplement parce qu'il manquait un couple de mots-rimes pour faire le pont entre les deux. Voici à titre d'exemple le premier groupe de mots-rimes construit :

Groupe de mots-rimes	**Rimème**	**Nb mots V**	**Nb occ N**	**Mot (fréquence à la rime sur CORRAC)**
1	- u - [s]	18	1862	clous(1) coups(116) courroux(189) cous(1) dessous(5) doux(212) dégoûts(1) filous(2) fous(1) genoux(40) jaloux(161) loups(1) nous(145) résous(5) tous(86) vous(685) époux(211)

Clef de lecture : le premier chiffre correspond au numéro du groupe. Il est suivi par la transcription graphico-phonétique de la terminaison commune (rimème), puis par le nombre de mots dans le groupe (V), et par le nombre d'occurrences (N). Viennent ensuite tous les mots-rimes du groupe ainsi que leur fréquence.

Figure 5. Exemple de groupe de mots-rimes

L'ensemble des groupes de mots-rimes est présenté dans l'annexe XI. Voici à titre d'exemple, les premiers groupes de mots-rimes correspondant à la voyelle /*u*/.

Groupe de mots-rimes	Rimème	Nb mots V	Nb occ N	Mot (fréquence à la rime sur CORRAC)
256	- u - [e]	8	28	avoue(5) boue(5) dénoue(1) désavoue(5) joue(6) loue(3) roue(2) voue(1)
399	- u - [ent]	3	4	désavouent(2) jouent(1) renouent(1)
1	- u - [s]	18	1862	clous(1) coups(116) courroux(189) cous(1) dessous(5) doux(212) dégoûts(1) filous(2) fous(1) genoux(40) jaloux(161) loups(1) nous(145) résous(5) tous(86) vous(685) époux(211)
269	- u b l [e]	3	14	double(3) redouble(4) trouble(7)
626	- u b l [ent]	2	2	redoublent(1) troublent(1)
127	- u d r [e]	3	64	foudre(31) poudre(14) résoudre(19)
129	- u d/t [-]	5	58	bout(24) debout(1) partout(1) résout(5) tout(27)
343	- u l [e]	3	6	foule(3) roule(1) écoule(2)
277	- u p [-]	2	12	beaucoup(6) coup(6)
546	- u p [e]	2	4	coupe(2) troupe(2)

Clef de lecture : le premier chiffre correspond au numéro du groupe. Il est suivi par la transcription graphico-phonétique de la terminaison commune (rimème), puis par le nombre de mots dans le groupe (V), et par le nombre d'occurrences (N). Viennent ensuite tous les mots-rimes du groupe ainsi que leur fréquence.

Tableau 10. Extrait de la liste des groupes de mots-rimes

Nous avons complété la description des groupes de mots-rimes par le RIMARIUM (*cf.* Annexe XI) qui présente pour chaque groupe tous les couples de mots-rimes attestés avec leur fréquence. Ci-dessous nous avons extrait du RIMARIUM le contenu des deux premiers groupes de mots-rimes correspondant à /u/. La situation de la première occurrence du couple sur CORRAC est indiquée entre parenthèses.

u

- u - [e] **(v=8 n=14)**

—#256: - u - [e] (v=8 n=14)—

2 .....avoue-joue.....(C, Attila, IV, III)
2 ..... joue-boue..... (C, Polyeucte, IV, III)
2 .....loue-désavoue.....(C, Mélite, I, IV)
1 .....avoue-boue..... (C, Don Sanche d'Aragon, I, I)
1 .....avoue-voue.....(R, Andromaque, IV, V)
1 .....boue-avoue.....(C, Pompée, I, III)
1 ..... boue-roue.....(C, Clitandre, III, I)
1 .....désavoue-dénoue..... (C, Le menteur, IV, I)
1 .....désavoue-loue.....(C, Suréna, V, III)
1 .....joue-désavoue..... (C, La suite du menteur, III, XI)
1 ..... joue-roue..... (C, L'illusion comique, V, V)

- u - [ent] **(v=3 n=2)**

—#399: - u - [ent] (v=3 n=2)—

2 ..... désavouent-renouent.....(C, Horace, IV, IV)
1 ..... jouent-désavouent..... (C, Mélite, II, VII)

Clef de lecture : le groupe de mots-rimes 256 est caractérisé par le rimème - u - [e], il regroupe 8 mots-rimes différents, qui apparaissent dans 14 couples de mots-rimes. Le couple *avoue-joue* apparaît 2 fois dans CORRAC et la première occurrence se situe dans *Attila*, acte IV, scène III.

Tableau 11. Extrait du RIMARIUM

Les groupes de mots-rimes sous forme de liste de mots assortis de leur fréquence d'occurrence et le RIMARIUM constituent le matériau de base pour explorer la rime. Nous avons également eu recours à un outil de représentation des données du RIMARIUM, pour faciliter le travail d'exploration.

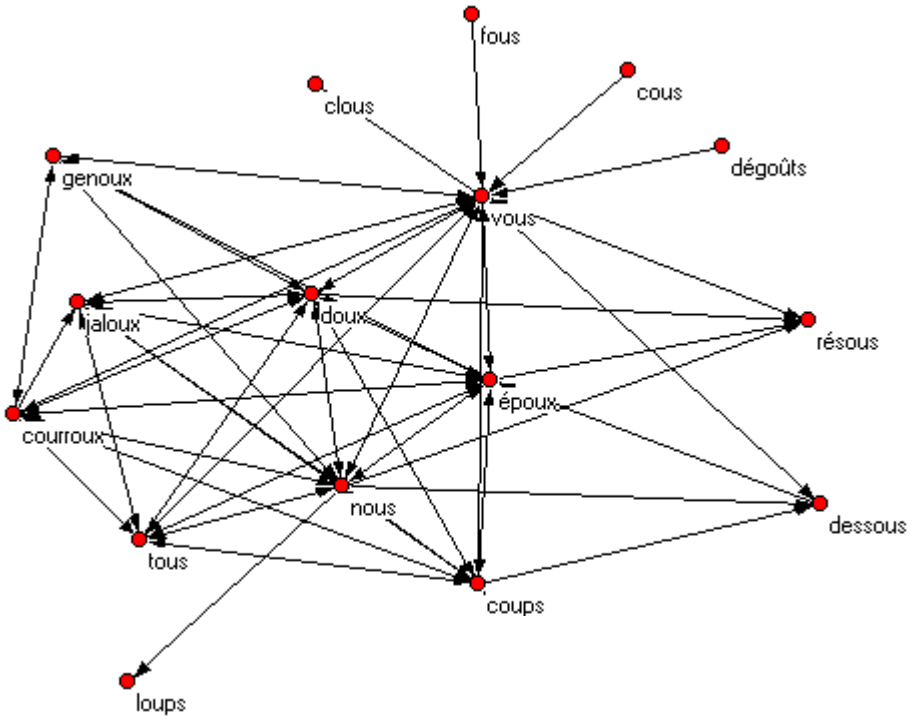
En effet, il nous fallait pouvoir représenter l'intensité des relations « rime avec » à l'intérieur d'un groupe de mots-rimes. Par exemple, il existe un groupe avec des mots-rimes en *on* et en *ion*. Or, les rimes *on*/*ion* sont très rares. Il nous fallait une forme de représentation qui nous montre que l'on a d'un côté les rimes en *on*, de l'autre les rimes en *ion* et qu'entre les deux les liens sont très rares.

Les représentations sous forme de réseaux correspondent mieux à nos besoins, puisqu'elles donnent une représentation des relations de rime attestées et de leur densité à l'intérieur de chacun des groupes de mots-rimes. En fait la notion de classe d'équivalence est assez inadaptée à notre

problème : la propriété de transitivité n'est pas toujours respectée comme nous le verrons. Nous avons longtemps cherché un outil qui puisse permettre de représenter graphiquement les réseaux en tenant compte de la densité des relations. Un certain nombre de produits sont en accès libre sur le réseau Internet. Nous avons utilisé le logiciel Pajek, développé par Vladimir Batagelj et Andrej Mrvar à l'Université de Ljubjana[45].

La Figure 6 montre le réseau des mots rimes en *ous(x)*. Ce type de graphe donne les relations attestées et la position des points les uns par rapport aux autres est liée à la densité des liens. Mais la fréquence des couples n'apparaît pas sur les graphes, il faut les compléter par le RIMARIUM. Par exemple, pour répondre à des questions du type : est-ce que les relations entre les mots sémantiquement et/ou syntaxiquement proches sont moins fréquentes que les autres ? il faut se reporter au RIMARIUM. Ainsi, entre *courroux* et *jaloux*, sur le graphe il n'y a qu'une flèche qui correspond à deux paires de vers, qu'il faudrait comparer à la fréquence de tous les autres couples avant de pouvoir dire que le vers évite de redoubler identité phonique et proximité sémantique. *A contrario*, les couples de mots-rimes *vous-nous* et *nous-vous* sont très fréquents (4 % environ), alors que la proximité morphologique est si grande que les métriciens recommandaient d'éviter ces associations.

[45] Nos remerciements à Zbigniew Smoreda, qui travaille entre autre sur les réseaux de sociabilité téléphonique, d'avoir trouvé et de nous avoir communiqué ce logiciel (Adresse : http : //vlado.mat.uni-lj.si/pub/networks/pajek).



Clef de lecture : *fous* a rimé avec *vous*, mais *vous* n'a pas rimé avec *fous* chez Corneille et Racine

***Figure 6. Réseau de mots-rimes* ous**[46]

Munis de la liste des groupes de mots-rimes, du RIMARIUM et des graphes de réseaux, nous sommes en mesure d'explorer certains phénomènes propres à la rime.

III.3. QUELQUES ELEMENTS SUR LE FONCTIONNEMENT DE LA RIME DANS LE VERS CLASSIQUE

L'examen des groupes de mots-rimes que nous avons constitués et du RIMARIUM nous ont permis de dégager certains éléments sur le fonctionnement de la rime. Nous commencerons par exposer les principes d'équivalences qui sont admis pour la rime : équivalences graphiques ou phoniques pour la voyelle cœur de rime, pour l'environnement droit de la

[46] Il aurait été plus intéressant de pouvoir indiquer sur ce type de graphe la fréquence de chaque mot-rime, et la fréquence de chaque couple sur l'arête, comme nous le suggérait Benoît Habert. Actuellement, le logiciel ne nous le permet pas.

voyelle et pour la terminaison. L'examen des groupes de mots-rimes, et le retour aux éditions originales, nous ont permis de mieux rendre compte de la dimension graphique de la rime.

III.3.1. Les principes d'équivalence et d'opposition

Proposons un modèle de description de la rime. Considérons que la terminaison du vers se décompose en trois éléments V, C, T organisés dans des séquences VCT. Les trois conditions suivantes sont nécessaires, mais pas suffisantes, pour que deux vers riment ensemble : a) les voyelles noyaux[47] V de la dernière position métrique de chaque vers sont *équivalentes*, b) les groupes consonantiques[48] C qui suivent cette voyelle (environnement droit) sont *équivalents* et c) les terminaisons T des deux mots-rimes sont *équivalentes*. Nous reprenons le terme d'équivalence, qui nous paraît plus approprié que celui d'identité, à Shapiro [1974, p. 506].

Ce terme d'équivalence correspond mieux à la réalité que celui d'identité, mais comme nous l'avons déjà dit, ces équivalences ne sont pas uniquement d'ordre phonologique, mais également d'ordre graphique, c'est pourquoi nous ferons intervenir dans la description parfois des graphèmes, parfois des phonèmes. Pour V, C et T, ce sont des principes d'équivalences phonétiques et/ou graphiques qui sont à l'œuvre et que nous allons maintenant décrire.

La terminaison (T)

Précisons pour commencer ce que nous appelons terminaison. Une terminaison admet deux genres : elles est féminine si le mot-rime finit par *e*,

---

[47] Dans l'hendécasyllabe espagnol, la rime implique au moins la voyelle de la dixième position - marquée par un accent tonique - et tout ce qui la suit soit : aucune, une ou deux syllabes. La rime commence au moins à la dernière voyelle métrique marquée et implique toujours l'environnement droit, ce qui veut dire qu'un mot accentué sur une syllabe donnée ne peut rimer qu'avec un mot accentué sur la même syllabe : ainsi, *árbol* (accentué sur l'avant dernière syllabe) ne pourrait jamais rimer avec *caracol* (accentué sur la dernière).

De ese modo, no es cordura / Querer curar la pasión, / Cuando los remedios son / Muerte, mudanza y locura (Miguel de Cervantes, in Anthologie bilingue de la poésie espagnole, p. 345).

[48] Les groupes consonantiques peuvent être vides, dans ce cas-là, il faut qu'ils soient absents dans les deux cas : la ressemblance est dans l'absence.

es, *ent* (les *e* correspondant à des *e* muets) et elle est masculine sinon. Elle admet trois nombres : *singulier*, *pluriel 1* si le mot finit par *s*, *z* ou *x* et *pluriel 2*, si le mot finit par *ent*. Il ne s'agit pas du genre et du nombre des *mots* mais du genre et du nombre de la *terminaison* par analogie avec les mots.

Pour chaque voyelle suivie d'un groupe consonantique, les mots-rimes se répartissent dans cinq ensembles selon le genre de la terminaison et le nombre. Chaque terminaison admet une seule forme, y compris le vide. La seule exception vient du pluriel 1. Dans les éditions originales, un seul type d'équivalence graphique est introduit entre *s* et *x* :

> Mais si de nos devins l'oracle n'est point faux,
> Sa grandeur doit atteindre aux degrés les plus hauts ;
>
> Corneille, *Attila*, vers 92-93.

> Rome vous voit, Madame, avec des yeux jaloux,
> La rigueur de ses Loix m'épouvante pour vous.
>
> Racine, *Bérénice*, vers 293-294, éd. P. Mélèse.

L'orthographe moderne introduit une équivalence supplémentaire entre *z* et *s*, qui n'existe pas dans l'orthographe du XVIIe siècle où le participe passé pluriel comme la seconde personne du pluriel finissent en *ez* :

> Quel temps choisissez-vous ? Ah de grace ! Arrestez.
> C'est trop pour un Ingrat prodiguer vos bontez.
>
> Racine, *Bérénice*, vers 617-618, éd. P. Mélèse.

Bien entendu les cinq possibilités ne sont pas toujours attestées dans notre corpus, et certaines sont impossibles d'emblée dans la langue.

nombre/genre de la rime[49]	masculin	féminin
singulier	vide	e
pluriel	s/z/x	es ent

Dans le vers classique, jamais un mot féminin n'a pu rimer avec un mot masculin, même si le *e* final n'était plus audible ; ni une terminaison singulière avec une terminaison plurielle. En cela, la langue française marquait son attachement à ses origines (comme dans le maintien de certaines consonnes latentes, qui étaient des traces de l'origine latine [Cerquiglini, 1995]) : ne peuvent rimer ensemble des mots qui se

[49] Il s'agit du genre et du nombre de la rime et non pas du genre et du nombre grammatical.

prononcent de la même façon, mais qui dans un état antérieur de la langue étaient distincts dans leurs sonorités. Ces principes apparaissent dans tous les manuels.

Pour V et C, il nous reste à définir les principes d'équivalences. Nous n'avons pas précisé, volontairement, si la voyelle V et le groupe consonantique C étaient définis par des graphèmes ou des phonèmes ; volontairement car les deux catégories interviennent pour les définir, prouvant une fois encore, que la rime est faite pour un œil-oreille. Un œil-oreille qui aurait une longue mémoire de la langue.

Les voyelles (V)

Nous avons examiné, à l'intérieur de chaque groupe de mots-rimes, les différentes graphies admises comme équivalentes, puisque mises en correspondance par la rime. Le problème de l'orthographe moderne de notre corpus se pose ici une fois encore, puisque nous n'avons pas pu vérifier toutes les correspondances dans les éditions originales. De même, les équivalences phonétiques correspondent davantage à une diction d'aujourd'hui qu'à une diction d'époque. Le Tableau 12 est la synthèse de cet examen empirique.

	équivalences graphiques	équivalences phoniques
1	a, â	[ɑ], [a]
2	an, en	[ɑ̃]
3	ai, ei, ê, *e* + consonne double	[ɛ]
4	é, *e* suivi de {r, z}	[e], [ɛ]
5	i, y, ï	[i]
6	ain, ein, in	[ɛ̃]
7	ien, yen	[j ɛ̃]
8	o	[ɔ]
9	eau, au[50]	[o]
10	ô (os), au	[o]
11	on	[ɔ̃]
12	ou	[u]
13	u	[y]
14	un	[œ̃]
15	eu	[ø], [œ]
16	oi	[w ɑ] ([wɛ]ou [ɛ] au XVIIe siècle)
17	oin	[w ɛ̃]
18	ui	[ɥ i]

Tableau 12. Voyelles cœur de rime

Le noyau vocalique V qui constitue le cœur de la rime peut correspondre à un ou deux phonèmes (évidemment à un ou plusieurs graphèmes). Par exemple, [w ɑ], [w ɛ̃], [ɥ i] sont des noyaux vocaliques au même titre que [e], [ɛ] : ils sont constitués de deux phonèmes qui forment une unité insécable. Les groupes *oi* et *oin* constituent une entité vocalique qui intègre la semi-voyelle, ce sont des cas de « diphtongue de forte cohésion » [Klein, 1991], puisque l'équivalence porte toujours sur le groupe vocalique entier : *oi* ne peut rimer qu'avec *oi* et *oin* qu'avec *oin.* La prononciation a évolué (en passant de [wɛ] ou [ɛ] à [wa]) en même temps que la pratique de la rime puisque l'on trouve fréquemment dans la chanson française actuelle des cas de rimes *oi*/*a* :

[50] Il n'y a dans CORRAC que deux uniques cas d'équivalence eau/au.

J'arrivais au moment précis, juste un peu tard,
Il saignait cinq kilos d'Héro sur le comptoir.[51]

Bernard Lavilliers, *Nuit d'amour*, « Night bird ».

En revanche, le groupe *ui* constitue une diphtongue plus instable. En fin de vers avec une terminaison nulle, la rime porte sur l'ensemble (glissante + voyelle), mais si la terminaison existe, la glissante n'intervient plus systématiquement : ainsi, *fuie* rime avec *vie*, *suivre* avec *vivre*.

Il apparaît clairement que pour les voyelles, les équivalences portent tantôt sur l'écrit, tantôt sur l'oral : ainsi, quatre graphies sont admises pour [ɛ̃] : *ai, ei, ê, e* (suivi d'une double consonne). Les couples suivants sont attestés :

chaînes-peines	5
Athènes-peines	2
peines-siennes	1

Deux réalisations orales sont admises pour la graphie *er* : [e] ou [ɛ], même si les traités recommandaient d'estomper au maximum la différence dans la prononciation[52] :

Je crains qu'il ne vous soit malaisé d'empêcher
Qu'un gendre valeureux ne vous coûte bien cher.

Corneille, *Médée*, vers 1095-1096.

Un même son peut donner lieu à plusieurs groupes de mots-rimes si la graphie est différente. Ainsi, un groupement vocalique *eau* ne rime quasiment jamais avec un groupement *au* et encore moins avec *ô* et pourtant dans les deux cas il s'agit de *o* fermés, [o].

Les deux *uniques exceptions* sont :

---

[51] Ces « vers » n'ont de l'alexandrin que le nombre de positions : rime uniquement pour l'oreille, pas de césures à l'hémistiche.

[52] La question de la prononciation différente ou non des rimes normandes au XVIIe siècle n'est pas vraiment tranchée (*cf.* Aroui, 1996]).

Mes fils, que j'ai perdus pour m'élever au trône,
Tant d'autres malheureux dont j'ai causé les maux,
Font déjà dans mon cœur l'office des bourreaux.

Racine, *La Thébaïde*, vers 1511-1512.

L'impétueuse ardeur de ces transports nouveaux
À son sang prisonnier ouvre tous les canaux ;

Corneille, *Attila*, vers 1755-1756.

Nous aurions pu exclure de ce tableau les voyelles nasales, mais la nasalisation autorise des rimes supplémentaires, par exemple entre *an* et *en*, alors que non nasalisées, ces voyelles ne peuvent pas rimer ensemble.

Il l'épouse. Ai-je donc attendu ce moment
Pour me venir encor déclarer son amant ?

Racine, *Bérénice*, vers 29-30.

On notera qu'à la rime, il y a équivalence entre le *a* antérieur et le *a* postérieur, mais pas entre le *o* fermé et ouvert. De fait, plutôt que la prononciation, c'est la différence de graphie *eau* versus *o* qui justifie l'absence d'équivalence.

Nous n'avons pas distingué les prononciations [ø], [œ] car cette différence est déterminée par la ou les consonnes qui suivent.

Le groupe consonantique (C)

Les groupes consonantiques graphiques attestés après la voyelle sont les suivants : *ø, b, bl, br, c, ch, cl, ct, d, dr, f, g, gl, gn, gu,* (*i*)*l, l, lt, m, mb, mbl, mbr, mp, n, nc, nch, nd, ndr, ng, nt, ntr, p, q(u), r, rb, rc, rd, rg, rm, rq(u), rs, rt, rv, s, st, str, sq(u), t, tr, v, vr,* (où ø représente le graphème vide).

Pour le groupe consonantique, le principe d'équivalence phonique est dominant.

Pour les consonnes sonores, non latentes, les équivalences graphiques sont simples : *ss*, c et s sont équivalents dans certains contextes, s et z dans d'autres, *ph* est équivalent à *f*, C*h* à C (avec C différent de *p* et de *c*), et C à CC.

Pour les consonnes aujourd'hui latentes, l'équivalence phonique ne joue pas évidemment, mais les équivalences graphiques sont plus complexes :

• *t* et *d*, s'ils sont latents, sont équivalents :

Ma présence la chasse, et son muet départ
A presque devancé son dédaigneux regard.

Corneille, *La Place Royale*, vers 111-112.

• *c* et *g* également,

Je vois que votre honneur demande tout mon sang,
Que tout le mien consiste à vous percer le flanc,

Corneille, *Horace*, vers 469-470.

• *m* et *n* aussi

J'ai choisi dès longtemps pour digne compagnon,
L'héritier de mon sceptre, et surtout de mon nom.

Racine, *Mithridate*, vers 617-618.

Nous n'introduisons pas ici l'équivalence entre *s*, *x* et *z* car elle concerne ce que nous avons appelé la terminaison.

Quand, après la voyelle, on a un groupe consonantique latent avec une terminaison de la rime plurielle (*s*, *z* ou *x*), l'identité de la consonne latente n'est pas nécessaire.

Par exemple, pour les rimes en /a r/, alors qu'au singulier, on a deux groupes de mots-rimes, l'un constitué des mots finissant en *ar*, l'autre en *ard* ou en *art*, au « pluriel », il n'y a qu'un seul groupe, avec des fins en *ars*, *ards*, *arts*.

Groupe de mots-rimes	Rimème	Nb mots V	Nb occ N	Mot (fréquence à la rime sur CORRAC)
85	- a r d/t [-]	13	140	art(7) départ(13) fard(6) gaillard(1) hasard(35) part(43) poignard(5) regard(6) rempart(2) renard(1) tard(17) vieillard(4) étendard(1)
245	- a r [-]	3	18	Aspar(5) César(9) char(4)
156	- a r [s]	12	80	Césars(2) Mars(8) chars(1) césars(5) dards(2) hasards(9) pars(1) parts(18) regards(21) remparts(5) épars(3) étendards(5)

Clef de lecture : le premier chiffre correspond au numéro du groupe. Il est suivi par la transcription graphico-phonétique de la terminaison commune (rimème), puis par le nombre de mots dans le groupe (V), et par le nombre d'occurrences (N). Viennent ensuite tous les mots-rimes du groupe ainsi que leur fréquence.

Ainsi, *César* qui au singulier ne rime qu'avec *char* ou *Aspar*, rime au pluriel (*Césars*) avec *hasards*, *regards* et aurait pu rimer avec *Mars*[53].

> Il semblait triomphant, et tel que sur son char
> Victorieux dans Rome entre notre César.
>
> Corneille, *Polyeucte*, vers 227-228.

> C'est m'exposer en butte aux fureurs de Marcelle,
> Au pouvoir de son frère, au courroux des Césars,
> Et pour un vain effort courir mille hasards.
>
> Corneille, *Théodore*, vers 1741-1742.

On notera qu'au XVIIe siècle la consonne latente finale du radical avait tendance à disparaître quand la terminaison était au pluriel : ainsi, on trouve dans la graphie de l'époque *momens*, *sermens* au pluriel, et *moment*, *serment* au singulier. Cette variante graphique entre le pluriel et le singulier peut aussi nous indiquer que le *t* final s'était déjà fortement affaibli. Grâce à cet artifice, plus grande était la ressemblance graphique entre les mots qui au singulier ne pouvaient rimer ensemble. Le genre et le nombre de la terminaison ont donc une incidence sur le système d'équivalences des groupes consonantiques.

L'interprétation qui a été donnée de ces équivalences de consonnes latentes est la suivante :

> Les règles conçues par les théoriciens du début du XVIIe siècle et les usages établis à leur suite peuvent se résumer dans cette formule : seules peuvent rimer les consonnes écrites qui dans la liaison aussi formeraient une rime pure. C'est

[53] Se pose la question de la prononciation des noms propres comme *Mars* ou *Titus* : étaient-ils prononcés sans le *s*, ou n'était-ce qu'une rime pour l'œil ?

donc la valeur phonique qui paraît dans la liaison et non l'orthographe qui importe.

[Elwert, 1965, p. 99].

Ainsi, en situation de liaison, les consonnes *t* et *d*, sont réalisées par [t] ; *c* et *g* par [k] et *s*, *z* et *x* par [z]. On explique ainsi que *Césars* puisse rimer avec *hasards*, mais pas *César* avec *hasard*, que *corps* puisse rimer avec *efforts*...

Les ans, qui m'ont usé l'esprit comme le corps,
Abattraient tous les deux sous les moindres efforts ;

Corneille, *Pulchérie*, vers 575-576.

Le principe d'équivalence des consonnes latentes repose sur une réalisation phonique potentielle, sans doute jamais réalisée, puisque la liaison ne se fait jamais d'un vers sur le vers suivant, si l'on considère que le vers constitue un mot phonologique[54]. L'absence de liaison d'un vers à l'autre rend peu cohérente l'explication, comme le fait remarquer Cornulier[55] [1995, p. 214]. Dans la tradition de Martinon, il considère qu'au XVIIe siècle tous les signes graphiques se prononçaient y compris les consonnes finales (les oppositions voisé/non voisé, m/n étant neutralisées), ce qui rend inutile l'hypothèse de liaison potentielle. On peut aussi considérer comme nous le suggère justement B. Habert, que les équivalences de consonnes latentes peuvent être interprétées comme une mémoire d'un état antérieur de la langue.

Nous verrons plus tard que pour les consonnes d'appui, il n'y a quasiment jamais d'équivalence entre *t* et *d*, entre *c* et *g*... Si les consonnes dites latentes étaient encore prononcées au XVIIe siècle comme les consonnes internes, il serait étonnant qu'un *t* puisse être équivalent à un *d*, un *c* à un *g*... Pour tenir l'hypothèse phonique de la rime, il faut admettre un affaiblissement de ces consonnes finales.

---

[54] Milner [1974, 1982] pose comme principe constituant que le vers peut être considéré comme un mot phonologique et que toutes les règles phonologiques qui gouvernent la question du e muet et de la liaison dans la langue s'appliquent au vers. Ainsi, en langue la liaison se fait toujours à l'intérieur d'un mot phonologique mais jamais d'un mot phonologique à l'autre. En transposant, Milner pose que toutes les liaisons se font à l'intérieur du vers mais jamais d'un vers à l'autre (*cf.* chapitre 0). Le mot phonologique est à la fois une unité syntaxique et un unité d'intonation.

[55] Cette idée ne tient pas, selon lui, pour trois raisons : inexactitude descriptive (Madrid/esprit) ; paradoxe métrique (pas de liaison de vers à vers) et un anachronisme (selon lui, les consonnes finales n'étaient pas encore amuïes, ce qui reste à démontrer).

Nous reprenons donc la proposition d'Elwert, qui a pour défaut de simuler une liaison là où le vers n'en fait jamais, mais qui a une grande efficacité descriptive :

Deux groupes consonantiques en fin de vers sont équivalents si et seulement si, en cas de liaison, la réalisation de la consonne latente est identique.

La décomposition de la rime en VCT nous permet de mieux catégoriser et observer les différents groupes de mots-rimes.

III.3.2. Le rôle de la graphie

Nous connaissons mal la prononciation du français au XVIIe siècle. Il paraît en tout cas évident que l'orthographe n'était pas le simple reflet de la prononciation : pour ne prendre qu'un exemple, il ne reste plus de trace du *s* implosif devant consonne à partir du XIIIe siècle, comme l'a montré Cerquiglini [1995, p. 26-36], alors qu'il se maintient dans l'orthographe. Les variations orthographiques que l'on observe dans le vers (*encor*/*encore*, *hasar*/*hasard*) nous confirment dans l'idée que les consonnes finales étaient pour une large part latentes, ainsi que les *e* muets en fin de mots.

Il s'ensuit que les différences de genre et de nombre ne devaient déjà plus être perceptibles par l'oreille. En tout cas, elles ne le sont plus pour nous. Certes, Bernardy [1986] et Milner et Regnault [1987] recommandent pour distinguer les terminaisons V des terminaisons Ve, d'allonger la voyelle V qui précède (ami/amie)... Mais comment l'oreille pourrait-elle distinguer *connue* de *connues*, ou *absolu* de *absolus*, qui ne peuvent pourtant pas rimer ensemble ? Seule la mise en situation de liaison met en évidence la différence. La rime a une dimension graphique évidente, puisque la plupart des oppositions de terminaisons n'étaient probablement plus et ne sont plus perceptibles par l'oreille.

Les différences en genre et nombre ne sont pas les seules qu'il nous soit aujourd'hui difficile de distinguer. Il en est de même pour toutes les terminaisons masculines avec consonnes latentes. Ainsi, *trésor* ne peut rimer avec *mort* alors que les réalisations phonétiques ont de grandes

chances d'être identiques. Certains phonèmes ne peuvent rimer ensemble s'ils n'ont pas la même graphie (*au* ne rime quasiment jamais avec *eau*). Plus généralement, l'examen des fréquences des couples de mots rimes et plus encore l'examen des éditions originales montre une nette préférence pour la ressemblance graphique de la séquence VCT. Cela devrait être analysé plus systématiquement sur les éditions d'époque.

Il y a, en effet, chez Corneille et Racine, un réel souci de cohérence dans le système de la rime. Ils utilisent ou créent des variantes orthographiques pour maintenir la cohérence de l'ensemble, prouvant encore le rôle de la graphie. Nous l'avons vu précédemment dans les variations graphiques *encor*/*encore*, *die*/*dise*, *hasar*/*hasard*.

Les variantes orthographiques sont mobilisées pour enrichir l'identité graphique. Ainsi, chez Racine, d'après l'édition de Pierre Mélèse, on trouve pour *paraître* et ses dérivés trois orthographes selon le contexte à la rime :

> C'est donc là cette ardeur qu'il me faisoit paraistre ?
> Cet humble Adorateur se déclare mon Maistre :
>
> Racine, *Alexandre le Grand*, vers 681-683, éd. P. Mélèse.

> Avant que sur ses bords l'Inde me vît paroistre,
> L'éclat de sa vertu me l'avoit fait connoistre.
>
> Racine, *Alexandre le Grand*, vers 1013-1014, éd. P. Mélèse.

> Si je leur donne temps, ils pourront comparestre.
> Çà, pour nous élargir, sautons par la fenestre.
>
> Racine, *Les Plaideurs*, vers 63-64, éd. P. Mélèse.

La variation orthographique permet d'augmenter l'identité graphique des deux rimes, même si celle-ci n'est pas indispensable pour assurer la qualité de la rime puisque l'on trouve aussi :

> Non, je ne vous croy point. Mais quoy qu'il en puisse estre,
> Pour jamais à mes yeux gardez-vous de paraistre.
>
> Racine, *Bérénice*, vers 915-916, éd. P. Mélèse.

Bien d'autres exemples de ces variantes orthographique qui augmentent la ressemblance graphique des vers apparaissent chez Racine :

Helas ! s'il estoit vray... Mais non, il a cent fois
Rassuré mon amour contre leurs dures Lois.
Racine, *Bérénice*, vers 641-642, éd. P. Mélèse.

Invisible d'abord elle entendoit ma voix,
Et craignoit du Serrail les rigoureuses loix.
Racine, Bajazet, vers 203-204, éd. P. Mélèse.

A travers ces phénomènes se manifeste dans la pratique de nos deux poètes la recherche d'une esthétique graphique de la rime[56]. L'identité sonore ne suffit pas à faire la rime[57].

III.4. LE ROLE DE LA CONSONNE D'APPUI

Nous abordons dans cette section le rôle du phonème d'appui, ce qui nous conduira à mettre en évidence une hiérarchie des voyelles, le rôle de la consistance graphique de la rime, celui de la qualité syntaxique du morphème final et celui de la fréquence du vocabulaire.

Le vers alexandrin dans le *Roman d'Alexandre* est organisé en laisses, strophes de longueur variable de vers de même mètre. Tous les vers de la laisse ont la même rime. Cette organisation en laisses facilite l'identification de la séquence définissant la rime. Elle est l'intersection commune aux finales des vers de la laisse.

Un parcours du *Roman d'Alexandre* montre que la rime est constituée de la voyelle occupant la dernière position et de son environnement droit, à savoir tout ce qui suit cette voyelle. L'environnement gauche, quand il s'agit d'une consonne, n'intervient jamais dans la rime. La séquence est limitée à VCT. En revanche, la semi-voyelle précédant la voyelle V fait toujours partie de la rime. Le point de départ de la rime est invariant : voyelle ou semi-voyelle et tout ce qui suit. La rime dans les laisses du roman d'Alexandre porte sur VCT ou GVCT, où V est le noyau vocalique de la dernière position métrique et G la semi-voyelle, C le groupement consonantique et T la terminaison.

---

[56] Il faudrait pouvoir remonter aux manuscrits pour vérifier qu'il s'agit bien d'un effort de l'auteur, et pas seulement de l'imprimeur.

[57] Cette manière de jouer avec la graphie pour que la rime pour l'œil soit meilleure a été observée par G. Lote (cité par Aroui [1996]) chez les poètes du XVIe au XVIIIe siècles. Lote note que cette recherche d'une ressemblance graphique n'est cependant pas systématique.

Dans le vers classique, plus précisément dans le vers de Corneille et Racine, la définition du contenu de la rime se complexifie et varie selon la nature de la voyelle qui est au cœur de la dernière position. Ainsi, l'analyse des groupes de mots-rimes chez Corneille et Racine, et plus précisément l'analyse en réseau, nous montrent que ce qui précède la dernière voyelle métrique peut jouer un rôle dans la rime. En effet, tandis qu'en général il n'y a, pour une séquence VCT donnée, qu'un seul groupe de mots-rimes, parfois, il y a quasiment autant de groupes que de consonnes à gauche de la séquence.

Dans quels cas et comment intervient la consonne d'appui ?

Nous avons cherché à tester les quatre hypothèses suivantes :

La première pose qu'il existe une hiérarchie des voyelles, liée à leur intensité dans la réalisation acoustique et à leur épaisseur graphique. En fonction de la position d'une voyelle sur cette échelle, la consonne d'appui est plus ou moins nécessaire. Autrement dit, plus une voyelle a une saillance élevée, moins l'identité de la consonne d'appui est nécessaire pour qu'il y ait rime.

La seconde hypothèse consiste à dire que plus l'environnement droit de la voyelle est maigre (faible nombre de graphèmes), plus il faut un renforcement de l'environnement gauche de la voyelle pour qu'il y ait rime.

La troisième, déjà reconnue dans les traités de métrique, consiste à dire que, lorsque la fin de vers coïncide avec un morphème grammatical, la rime doit s'étendre sur la gauche et inclure la consonne d'appui.

Enfin, la quatrième pose que plus le nombre de mots partageant la même terminaison est élevé dans le corpus considéré, plus l'identité de la consonne d'appui est fréquente. La richesse du vocabulaire pour une terminaison donnée aurait une incidence sur le rôle de la consonne d'appui.

Nous évoquerons enfin le rôle joué par la glissante dans la rime.

Comme le faisait remarquer Dominique Billy,

> La caractérologie des types [de rime] revêt un caractère hétéroclite, des critères graphiques, parfois grammaticaux, etc. intervenant à côté de critères phoniques et statistiques qui prévalent et constituent le complexe central à travers lequel est recherchée cette valeur moyenne (...) où la pauvreté phonétique est rachetée par la rareté et inversement.
>
> [Billy, 1984, p. 75].

Nous cherchons donc à démêler les différentes composantes qui influent sur l'extension de la rime à partir de l'examen empirique d'environ 35 000 couples de mots-rimes.

III.4.1. Hiérarchie des voyelles

Pour les deux premières hypothèses, nous avons limité l'étude à cinq ensembles de groupes de mots-rimes : les voyelles nues (terminaison masculine, singulière), les voyelles nues suivies d'une terminaison féminine singulière (Ve), les voyelles nues suivies d'une terminaison masculine plurielle (Vs), les voyelles nues suivies d'une terminaison féminine, plurielle (Ves) et enfin les voyelles suivies de la consonne *r* (VT où T $\in \{\varnothing$, e, s, es$\}$, VC où C=r). Autrement dit, pour les quatre premiers ensembles, il n'y a pas de groupe consonantique entre la voyelle cœur de rime et la terminaison.

Pour chacun de ces ensembles nous avons regardé quelle était la proportion de couples de mots-rimes dans lesquels la consonne d'appui était différente.

voyelle p. 12	nb de groupes de mots-rimes	nb mots	nb de couples de mots-rimes	nb couples avec consonne d'appui différente	%
é	39	539	1315	1	0
i	15	82	280	1	0
u	4	61	263	7	3
a	14	40	34	2	6
eu[58]	2	11	86	46	53
eau	1	24	86	62	72
oi	1	22	836	815	97
ui	1	6	319	318	100
ai	4	8	6	1	-

Clef de lecture : la rime *u* apparaît dans quatre groupes de mots-rimes qui regroupent 61 mots différents et constituent 263 couples de mots-rimes. Parmi ces couples, la consonne d'appui est différente dans 7 couples (3%).

Tableau 13. Présence de la consonne d'appui pour les séquences V

Dans le vers de Corneille et Racine, deux mots en *é* ne peuvent rimer ensemble que si la consonne d'appui est identique. Et si la voyelle est précédée par plus d'une consonne, il y a une nette préférence pour l'identité de tout le groupe consonantique précédant. On trouve de manière exceptionnelle :

> Et quand par des voleurs il fut assassiné,
> Ce digne favori l'avait accompagné.
>
> Corneille, *Œdipe*, vers 995-996.

> Seigneur, mon sang m'est cher, le vôtre m'est sacré :
> Serai-je sacrilège, ou bien dénaturé ?
>
> Racine, *La Thébaïde*, vers 57-58.

Il n'existe qu'un unique exemple où l'environnement gauche est différent pour une terminaison en *é* :

> Cet heureux temps n'est plus. Tout a changé de face,
> Depuis que sur ces bords les dieux ont envoyé
> La fille de Minos et de Pasiphaé.
>
> Racine, *Phèdre*, vers 35-36.

Il est peu probable qu'au XVIIe siècle un yod [j] s'intercale entre le *a* et le *é* de *Pasiphaé*. Ces vers sont donc exceptionnels par leur traitement de la rime.

[58] L'un des groupes de mots-rimes correspond au rimème *eu*, l'autre à *ieu* : jamais un mot en *eu* ne rime avec un mot en *ieu*.

Il en est quasiment de même pour les rimes en *i* que pour les rimes en *é* : la consonne d'appui est toujours identique sauf dans un cas, situé dans *Les Plaideurs* :

> Ma robe vous fait honte : un fils de juge ! Ah, fi !
> Tu fais le gentilhomme. Hé ! Dandin, mon ami,
>
> Racine, *Les Plaideurs*, vers 89-90.

Le *u* nu est à peine plus vigoureux : sur 263 couples de mots-rimes en *u*, il y a sept cas où la consonne d'appui est différente (*voulu* rime avec *pu*, qui ailleurs rime avec *vu*, qui rime avec *tribu*, *interrompu*, *vaincu*...). La Figure 7 montre bien comment ce groupe de mots-rimes s'organise en sous-groupes (cliques) rarement connectés entre eux : c'est principalement le cas pour les terminaisons en *–du*, *–tu*, et *–lu*. Certains termes jouent le rôle de connecteurs entre ces cliques : *rendu*, *tu*, *pu*, *vu*, *voulu*. Voici quelques vers à titre d'illustration :

> Mais cet enfant fatal, Abner, vous l'avez vu :
> Quel est-il ? de quel sang ? et de quelle tribu ?
>
> Racine, *Athalie*, vers 545-546.

> Qu'est-ce qui s'est passé ? qu'as-tu vu ? que sais-tu ?
> Depuis quel temps, pourquoi, comment t'es-tu rendu ?
>
> Racine, *Mithridate*, vers 481-482.

On notera que dans les deux cas les consonnes d'appui ne diffèrent que par le trait voisé/non voisé. Ce type de rimes sans appui n'apparaît que chez Racine. Corneille utilise même les ressorts de la liaison pour éviter une rime aussi nue :

Permettez...
ANGELIQUE.
Cependant, à qui me laisses-tu ?
Tu frustres donc mes vœux de l'espoir qu'ils ont eu,
Corneille, *La Place royale*, vers 1079-1080.

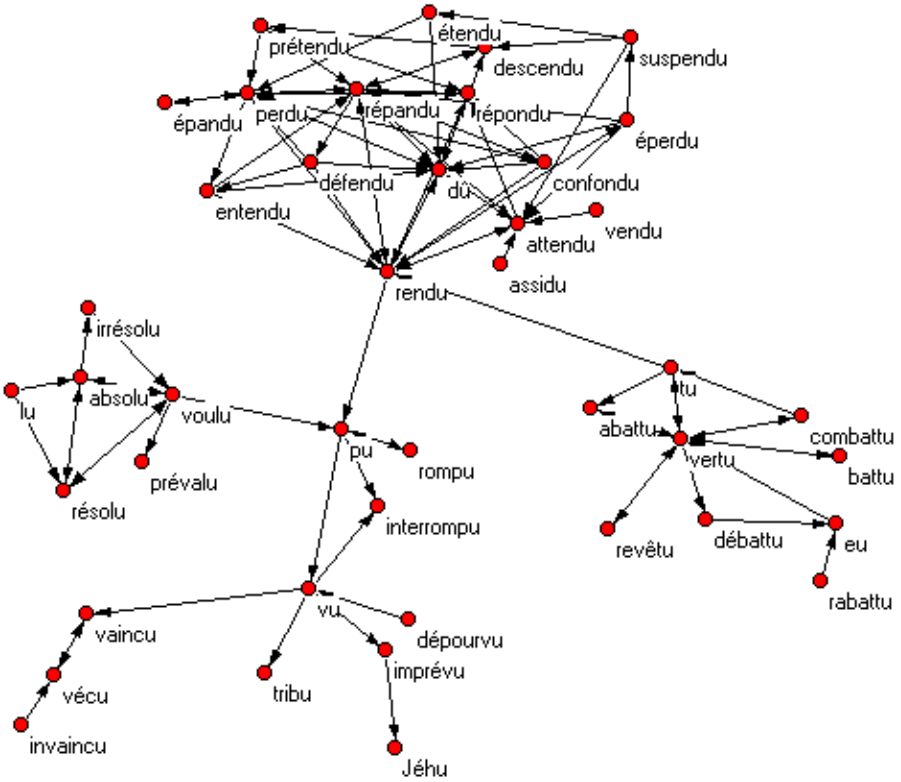


Figure 7. Le réseau des mots-rimes en u

Pour la voyelle *a*, les cas où la consonne d'appui est différente sont également marginaux : les deux exemples proviennent des *Plaideurs*.

CHICANNEAU.
Hé ! Ne te fâche point.
ISABELLE.
Adieu, monsieur.
ACTE_II, SCéNE_IV
L'INTIME.
Or çà,
Verbalisons.
CHICANNEAU.
Monsieur, de grâce, excusez-la :
Racine, *Les Plaideurs*, vers 373-374.

Et voilà comme on fait les bonnes maisons. Va,
Tu ne seras qu'un sot.
LEANDRE.
Vous vous morfondez là,

Racine, *Les Plaideurs*, vers 109-110.

Dès que l'on passe à des voyelles métriques orthographiées par digraphes ou trigraphes, l'identité de la consonne d'appui devient, au contraire, un phénomène minoritaire. L'épaisseur graphique semble donc jouer un rôle. Ainsi, dès que la voyelle métrique comprend plus d'un graphème, on a plus de 70 % de couples de mots-rimes avec des consonnes d'appui différentes, alors que pour les voyelles d'une lettre on a en moins de 14 %.

Pour les groupes de mots-rimes en Ve, le rôle de la consonne d'appui varie en fonction de la voyelle, comme précédemment.

voyelle p. 12	**nb de groupes de mots-rimes**	**nb mots**	**nb de couples de mots-rimes**	**nb couples avec consonne d'appui différente**	**%**
ée	41	423	554	2	0
ie	6	209	940	189	20
ue	1	71	299	140	47
oie	1	16	175	160	91
oue	1	8	14	13	93
aie	2	7	8	7	-

Tableau 14. Présence de la consonne d'appui pour les séquences Ve

Pour les terminaisons en *ée*, l'identité de la consonne d'appui est systématique. Elle ne souffre que deux exceptions, équivalence entre un *p* et un *b*, et entre un *v* et un *f* (*ph*), qui ne diffèrent que par le trait voisé/non voisé :

Elle s'en est, seigneur, mortellement frappée,
Et dans son sang, hélas ! Elle est soudain tombée.

Racine, *La Thébaïde*, vers 1469-1470.

Quand je sus qu'à son lit Monime réservée,
Avait pris, avec toi, le chemin de Nymphée !

Racine, *Mithridate*, vers 59-60.

En revanche pour *i*, le nombre des cas avec consonne d'appui différente est sensiblement plus élevé que pour les terminaisons masculines. Est ainsi creusé l'écart entre les valeurs de *é* et de *i*.

Les deux tableaux suivants portent sur les terminaisons plurielles (masculines et féminines) :

voyelle p. 12	nb de groupes de mots-rimes	nb mots	nb de couples de mots-rimes	nb couples avec consonne d'appui différente	%
ées	26	97	115	0	0
ies	11	43	42	7	17
ues	3	13	13	4	31
aies	1	2	4	0	-

Tableau 15. Présence de la consonne d'appui pour les séquences Ves

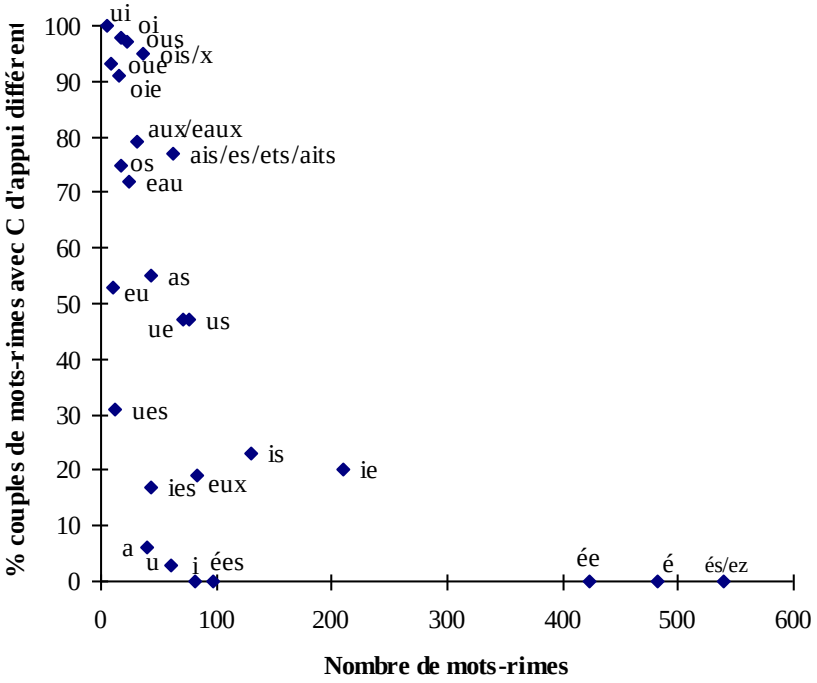
voyelle p. 12	nb de groupes de mots-rimes	nb mots	nb de couples de mots-rimes	nb couples avec consonne d'appui différente	%
és/ez	76	482	557	0	0
eux	1	83	968	185	19
is	9	130	588	137	23
us	4	76	230	109	47
as	1	42	534	295	55
os	1	18	73	55	75
ais/es/ets /aits	3	62	317	244	77
eaux/au x	3	32	68	54	79
ois/oix	1	36	402	381	95
ous/oux	1	18	931	915	98

Tableau 16. Présence de la consonne d'appui pour les séquences Vs

Roubaud [communication personnelle] a recherché dans d'autres textes du XVIIe siècle des cas de rimes non soutenues pour les terminaisons en *-é, -ée, -és, -ées, -er, -a, -i, -u*. Les cas sont rares, à l'exception de La Fontaine, qui n'a visiblement pas un usage « classique » de la rime[59].

[59] *Garni* rime avec *vi* dans *L'Étourdi* de Molière ; *souci* avec *transi* dans la *Diane* de Rotrou ; *Thésée* avec *épée*, *renommée* et *embarrassée* dans *Phèdre et Hippolyte* de Pradon ; *cheminée* avec *enfumée*, *échoué* avec *Cimothoé*, *égarée* avec *Anthée* dans le *Virgile travesti* de Scarron ; *cela* avec *plaira* chez Montfleury, *finira* avec *cela* chez Donneau de Visé et c'est à peu près tout sur une trentaine d'œuvres examinées par Roubaud. En revanche, dans

La richesse lexicale de chaque terminaison (*cf.* Figure 8), c'est-à-dire le nombre de mots différents qui la contiennent, a également une incidence sur le rôle de la consonne d'appui.



Clef de lecture : dans CORRAC, il y a 40 mots-rimes se terminant par *a*. Sur l'ensemble des couples de mot-rimes en *a*, 6 % ont des consonnes d'appui différentes. Pour *é*, aucun couple n'a de consonne d'appui différente.

Figure 8. Rôle de la consonne d'appui en fonction de la richesse lexicale des terminaisons

Les tableaux et la Figure 8 montrent d'importantes variations dans le rôle de la consonne d'appui selon les voyelles. Pour certaines, l'identité de la consonne d'appui est plus fréquente que pour d'autres. Le phonème [e] a un statut particulier, puisque dans tous les cas, l'identité de la consonne d'appui est nécessaire pour la rime : c'est le seul phonème qui ne peut à lui seul former une rime. Il n'est plus assez « fort » pour constituer la rime même dans un environnement à rimes plates, ce qui n'est pas le cas pour les

---

les fables de La Fontaine, les exemples sont plus nombreux : *chargé* rime avec *relevé*, *félicité* avec *pelé*, *passé* avec *né*, *parlé* avec *gré*, *approchez* avec *entendez*, *écouter* avec *manger*, *comptez* avec *propriétés* et *prés* ; *là* avec *monta*, *précipita* et *échappa*, *éclata* avec *gens-là*, *s'écria* avec *déjà*.

autres voyelles. Il apparaît donc une hiérarchie entre les voyelles, dans laquelle le [e] a une très faible consistance, très supérieure cependant au *e* muet qui ne peut même pas avoir le statut de cœur de rime. Ceci constitue une première évolution par rapport au *Roman d'Alexandre* dans lequel le graphème *e* suffisait à lui seul à constituer la rime.

Pour les autres voyelles, la proportion de couples de mots-rimes avec consonne d'appui différente est assez variable et pourrait révéler l'existence d'une hiérarchie des poids des voyelles : plus une voyelle a un poids élevé, plus elle a de chances d'être suffisante pour constituer la rime (toujours avec équivalence de l'environnement droit). Soit f la fonction qui pour une rime donnée indique le taux de couples de mots-rimes avec consonnes d'appui différente. D'après nos tableaux, nous voyons que f(é) < f(i) < f(u) < f(a), que f(ée) < f(ie) < f(ue), que f(és) < f(is) < f(us) < f(os) < f(as), que f(ées) < f(ies) < f(ues)... Le vers classique fait donc apparaître une hiérarchisation entre les voyelles pour la rime.

III.4.2. Principe de consistance graphique minimale

Tout se passe comme si dans la rime un principe de consistance graphique minimale était à l'œuvre. Plus la séquence VCT occupe un espace graphique important, moins le recours à la consonne d'appui est nécessaire. Un jeu subtil d'équilibre entre V, C et T intervient également.

Deux catégories de faits confirment cette hypothèse.

Premièrement, les tableaux précédents et la Figure 8 montrent que plus l'extension graphique de la voyelle métrique (le fait qu'elle soit orthographiée par une, deux ou trois lettres) est grande, plus elle est autonome par rapport à la consonne d'appui. Alors que, quand la voyelle finale est nue, il y a une préférence très majoritaire pour l'identité de la consonne d'appui, l'inverse apparaît pour les voyelles composées de plusieurs graphèmes. Ces dernières occupent un espace suffisant pour constituer la rime à elles seules. Ici encore, la graphie joue son rôle de différenciation entre les voyelles métriques.

Deuxièmement ces tableaux montrent le rôle que joue l'environnement droit de la voyelle. Plus il est riche, moins le recours à la consonne d'appui

est fréquent. Reprenons dans les tableaux précédents les voyelles *é*, *i*, *u*, suivies des terminaisons ∅, *e* et *s* (pour les terminaisons *es*, les effectifs sont trop faibles). Pour *i* et *u*, la proportion de couples sans consonne d'appui identique est plus élevée quand la terminaison est féminine ou masculine plurielle que quand la voyelle est nue (terminaison masculine).

Voyelle p. 12	**%**	**Voyelle p. 12**	**%**	**Voyelle p. 12**	**%**	**Voyelle p. 12**	**%**
é	0	ée	0	és/ez	0	ées	0
i	0	ie	20	is	23	ies	17
u	3	ue	47	us	47	ues	31

Par ailleurs, nous avons vu que chez Corneille, dans les finales en *u*, la consonne d'appui était toujours identique. Pourtant, dès que la voyelle est couverte à droite, même Corneille n'a plus systématiquement recours à la consonne d'appui :

> Et par l'objet présent les sentiments émus
> N'attendent pas toujours des rois qu'on n'a point vus.
>
> Corneille, *Suréna*, vers 399-400.

L'examen d'un nombre un peu plus élevé de groupes de rimes en *i* confirme cette hypothèse : dans le groupe de mots-rimes en *ire* (VCT=*ire*), l'identité de la consonne d'appui est minoritaire (33%), alors qu'elle est systématique pour le groupe *i* (VCT=*i*∅∅), de 80% pour le groupe *ie*...

Séquence VCT	**nb de groupes de mots-rimes**	**nb mots**	**nb de couples de mots-rimes**	**nb couples avec consonne d'appui différente**	**%**
i	15	82	280	1	0
ie	6	209	940	189	20
ies	11	43	42	7	17
is	9	130	588	137	23
ir	6	123	777	29	4
ire	1	48	510	340	67
irs	1	8	102	68	67
ise	1	39	126	95	75

Dans le vers classique, l'identité de la consonne d'appui pour les groupes VCT, où C est représenté par au moins une consonne, semble assez minoritaire.

Les deux constats précédents (la consonne d'appui est d'autant moins nécessaire que la voyelle métrique est transcrite par un nombre élevé de voyelles et que l'environnement droit de la voyelle est important) montrent que les principes d'équivalences qui définissent la rime sont complexes. Il

faut qu'il y ait une séquence graphique minimale, et le rôle de la consonne d'appui est d'enrichir cette séquence quand la voyelle et son environnement droit ne sont pas suffisants.

III.4.3. Les morphèmes grammaticaux

La présence de la consonne d'appui ne dépend pas seulement de la qualité de la voyelle, ni de la consistance de la séquence, elle dépend aussi de la qualité syntaxique de la terminaison. Quand la terminaison correspond à un morphème grammatical, très souvent, la consonne d'appui est nécessaire pour qu'il y ait rime. Ainsi, dans tous les groupes de mots-rimes qui correspondent à des verbes à l'imparfait ou au conditionnel, la consonne d'appui est identique :

> LYSE.
> Ai-je bien pris mon temps ?
> CLITON.
> Le mieux qu'il se pouvait.
> C'est une honnête fille, et Dieu nous la devait :
> Corneille, *La suite du Menteur*, vers 169-170.

> Il m'assurait qu'Othon jamais ne m'obtiendrait,
> Et disait à demi qu'un refus nous perdrait.
> Corneille, *Othon*, vers 563-564.

De même, pour les terminaisons de l'infinitif, les cas où la consonne d'appui est différente sont extrêmement rares (*cf.* Tableau 17). Dans les groupes de mots-rimes en *er* et *ir*, la très grande majorité des mots-rimes correspondent à des verbes à l'infinitif. On trouve quelques cas de noms et adjectifs : *air*, *clair*, *amer*, *mer*, *Alger*, *Jupiter*, *cher*, *cocher*, *rocher*, *acier*, *coursier*, *dernier*.... pour la séquence *er* ; *plaisir*, *soupir*, *vizir*... pour la séquence *ir*.

voyelle p. 12	groupe de mots-rimes (numéro)	nb mots	nb de couples de mots-rimes	nb couples avec consonne d'appui différente	%
er/air	23	605	1948	2 ou 3	0
ir	6	123	777	29	4

Tableau 17. Terminaisons de l'infinitif

Pour les terminaisons en *er*, nous avions à peu près autant de groupes de mots-rimes que de consonnes d'appui attestées. Un seul groupe échappait à la règle, puisqu'il regroupait des mots finissant en *ner*, *der*, *gner*, *ler*. La représentation en réseau du groupe des mots finissants en *ner*, *der*, *ler*,

montre la présence de sous-réseaux différenciés par la consonne d'appui et que les liens entre les sous-réseaux sont peu nombreux (*cf.* Figure 9).

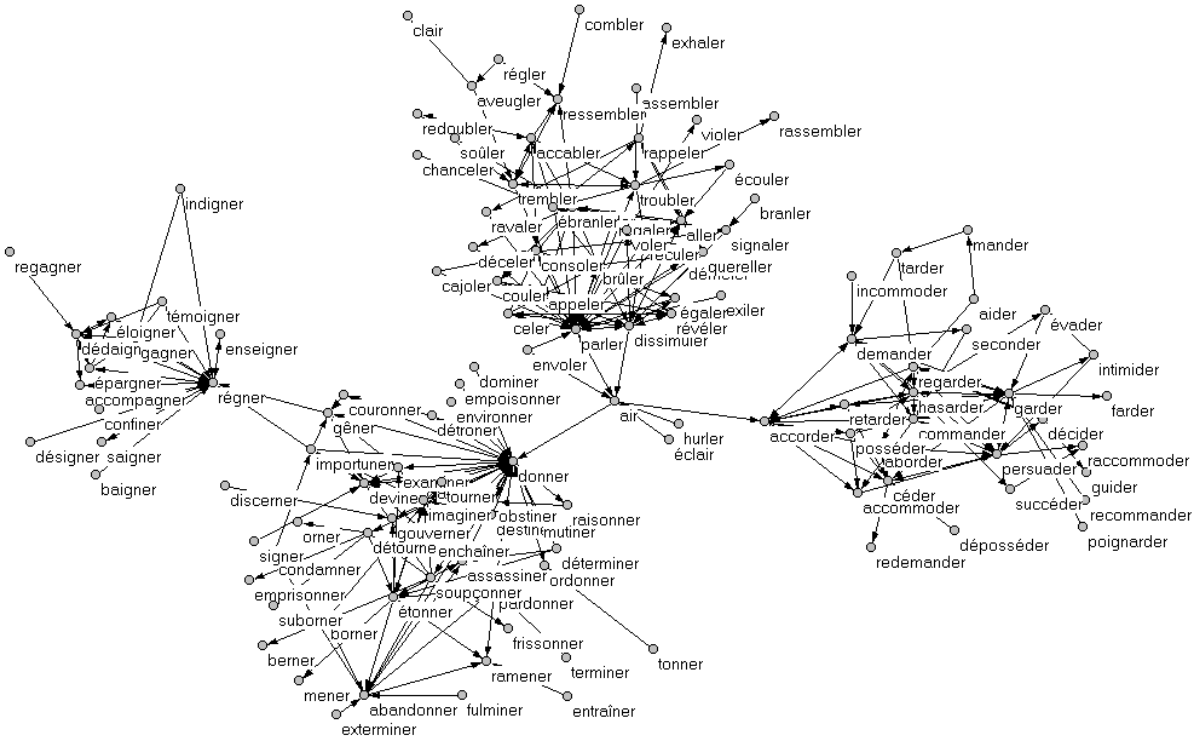


Figure 9. Le réseau des mots-rimes en der-ler-ner

Le regroupement de ces sous-ensembles est principalement dû au mot *air*, qui acquiert par là même un statut très particulier :

> Pour Clarice et Lucrèce, elles en ont quelque air ;
> Philiste avec Alcippe y vient vous accorder ;
>
> Corneille, *La suite du Menteur*, vers 289-290.

> Examiner sa taille, et sa mine, et son air,
> Et voir quel est l'époux que je vous veux donner.
>
> Corneille, *Le Menteur*, vers 391-392.

> La foi que j'en reçus s'en est allée en l'air,
> Et ces traits de sa plume osent encor parler,
>
> Corneille, *Mélite*, vers 913-914.

Les derniers vers montrent que ce mot échappe à notre hypothèse de départ (la rime ne repose que sur le dernier mot). L'article élidé est constitutif de la rime, de même que le *n* de *son* qui retrouve sa valeur consonantique en situation de liaison. En revanche, dans les deux premiers vers, il est clair qu'il n'y a pas de consonne d'appui identique. Il est assez significatif de constater que la liaison entre des groupes indépendants est

faite par l'intermédiaire d'un seul mot qui diffère par la graphie (*air* au lieu de *er*), par la prononciation ([ɛ r] au lieu de [e]) mais aussi par sa catégorie syntaxique (nom *versus* verbe à l'infinitif) de la très grande majorité des mots du groupe. Avec *air*, il n'y a pas de risque de redoublement de l'identité graphico-phonétique par une identité de terminaison grammaticale, ce qui serait le cas entre deux infinitifs en *er*, avec consonne d'appui différente. Nous avons aussi la preuve ici que la relation « rime avec » n'est pas toujours transitive, contrairement à ce que nous avions posé au départ. En effet, si *donner* peut rimer avec *air* et *air* avec *parler*, *donner* ne pourra jamais rimer avec *parler*. Deux terminaisons peuvent rimer ensemble à condition qu'il ne s'agisse pas du même morphème grammatical.

Pour les terminaisons en *ir*, sur 744 couples de mots-rimes, 29 ont des consonnes distinctes. Remarquons que dans les couples du type haïr/obéir, nous avons pris le parti de considérer qu'il y a un phonème consonantique ø, et donc que le phonème d'appui est identique.

Sur ces 29 cas, seuls deux font intervenir deux verbes à l'infinitif (*agir* / *obéir* ; *attendrir* / *rougir*).

> Heureuse si mes pleurs vous peuvent attendrir,
> Une mère à vos pieds peut tomber sans rougir.
>
> Racine, *Iphigénie*, vers 931-932.

Dans les autres cas, l'un des deux membres du groupe est un nom.

L'identité des morphèmes de l'infinitif ne suffit pas pour constituer la rime, car cette identité a un caractère trop fréquent et trop trivial. Surtout, ce type de rime introduit une identité morpho-syntaxique contradictoire avec la rime qui se veut préférentiellement une équivalence graphico-phonétique.

Le dernier exemple concerne les terminaisons en *ent* et *ant* qui correspondent principalement à des adverbes en *ment* et à des participes présents. La constitution des groupes de mots-rimes a également fait apparaître un grand groupe, avec des consonnes d'appui différentes. L'analyse du réseau différencie clairement des sous-réseaux entre lesquels les liens sont rares : sur 658 relations entre des mots terminant en *ent*/*ant*, dans 652 cas, la consonne d'appui est identique. On a six exceptions qui font pratiquement toujours intervenir l'adjectif grand (*grand*/*galant* ; *grand*/*tant* ; *différent*/*vend* ; *également*/*grand* ; *grand*/*vend* ; *grand*/*sanglant*). Voici l'une d'entre elles :

Flavie ; et de deux maux qu'on craint également
Celui qui nous arrive est toujours le plus grand,

Corneille, *Attila*, vers 1127-1128.

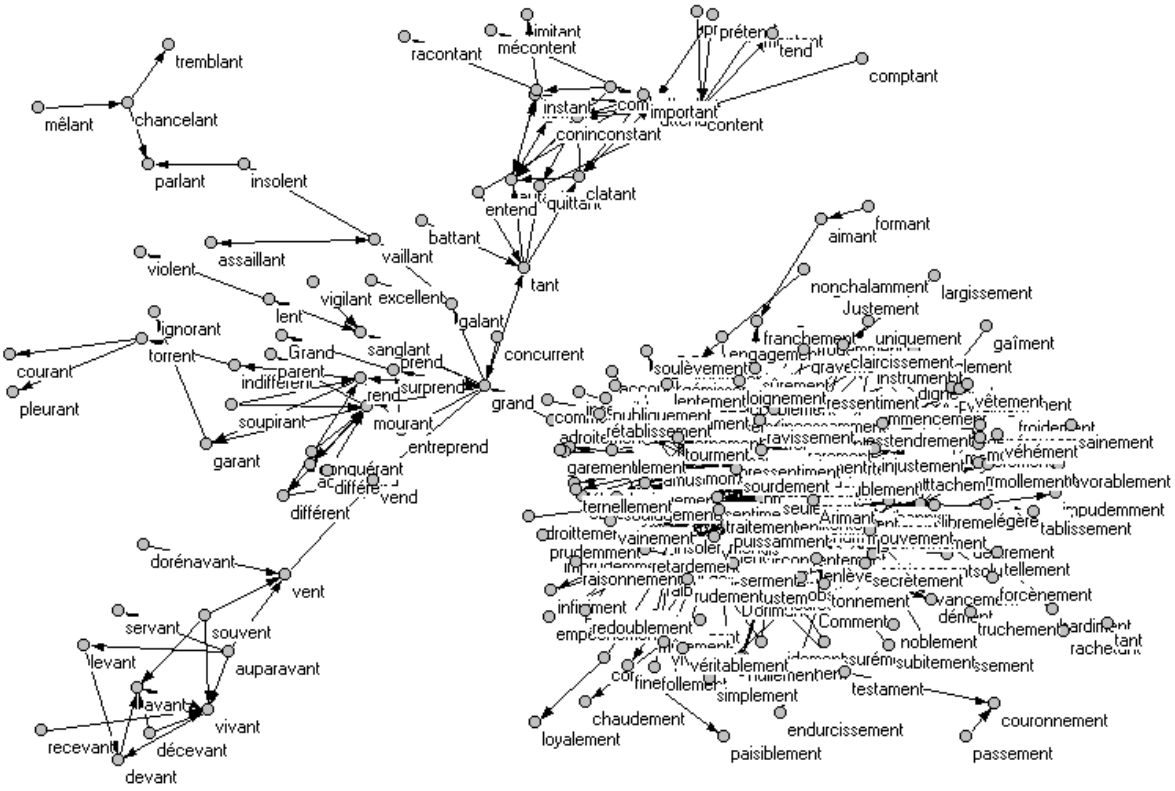


Figure 10. Réseau des rimes en ent-ant

En conclusion, quand les deux mots-rimes s'achèvent par le même morphème grammatical (terminaison de verbe, d'adverbe...) l'identité de la consonne d'appui est quasiment toujours indispensable. Il faudrait étudier les traditions poétiques antérieures pour identifier le moment où ce soutien est devenu indispensable.

Et cela tient au fait que le morphème grammatical est identique, et non pas seulement à la qualité phonographique de la terminaison. En effet, si l'un des deux mots-rimes ne présente pas ce morphème grammatical final, la rime n'a pas besoin de s'étendre à gauche de la voyelle V (air(nom)/accorder ; souvenir/soupir (nom)).

III.4.4. Incidence de la richesse lexicale des rimes

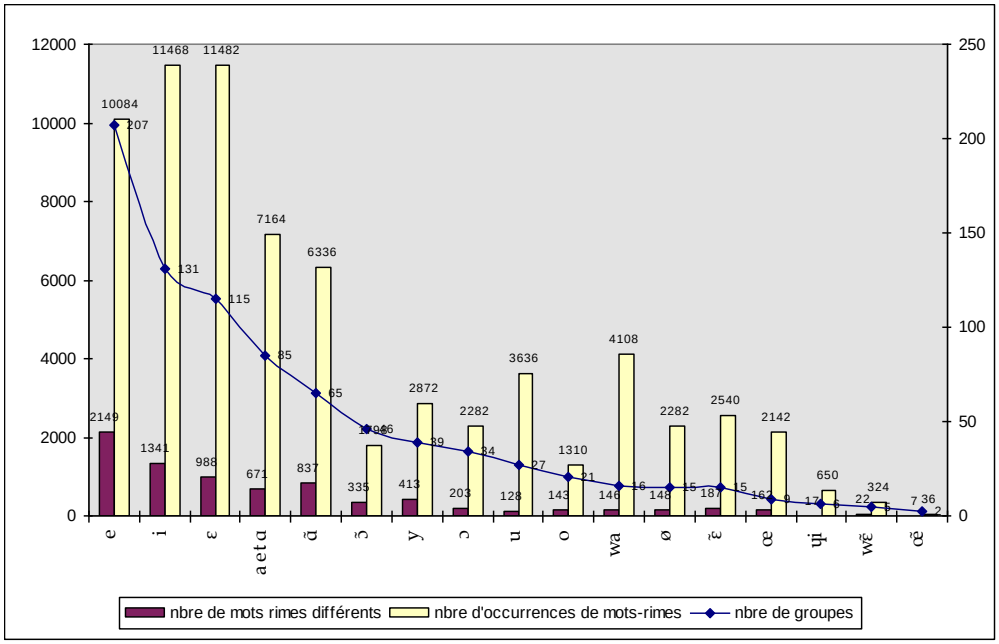
Il nous reste une dernière hypothèse à tester : l'incidence de la richesse lexicale des rimes sur le rôle de la consonne d'appui. La richesse lexicale d'une séquence VCT relative à un ensemble de vers est le nombre de mots-rimes du lexique ayant la terminaison VCT. On cherche à montrer qu'il y a une sorte d'équilibre entre la richesse lexicale d'une séquence VCT et l'extension ou la « richesse » de la rime. Plus une séquence VCT se trouve communément dans le lexique, plus la rime sera « riche » en termes de phonème et/ou graphèmes communs : elle fera le plus souvent intervenir la consonne d'appui.

Pour vérifier cette hypothèse, il aurait fallu pour chaque séquence VCT calculer le nombre de mots-rimes et le pourcentage de couples de mots-rimes avec consonne d'appui distinctes, comme nous l'avons fait pour quelques séquences, et mettre en évidence des corrélations. La difficulté étant de neutraliser les effets liés au type de voyelle, à l'épaisseur graphique de la séquence VCT et à la qualité syntaxique de la terminaison.

Nous avons tenté une approximation de ces calculs. Pour une voyelle donnée, un nombre élevé de groupes de mots-rimes signifie l'existence d'un nombre important de séquences VCT distinctes mais aussi la présence systématique d'une consonne d'appui (il y a alors pour une VCT donnée autant de groupes que de consonnes d'appui avant V). Ainsi, la voyelle /e/, pour laquelle il n'existe que cinq séquences VCT (e - [-], e - [e], e - [es], e - [s], e r [-]), admet 207 groupes de mots-rimes différents, parce que la consonne d'appui intervient quasiment toujours dans la définition de chaque groupe de mots-rimes. Inversement, la voyelle /ɔ/ qui admet 38 séquences VCT distinctes (- ɔ - [-], - ɔ - [s], - ɔ b [e], - ɔ d [e], - ɔ d/t [-], - ɔ f r [e], - ɔ k [e], - ɔ k [es], - ɔ l [e]...) admet 54 groupes de mots-rimes distincts, et aucun groupe ne fait intervenir la consonne d'appui[60]. La Figure 11 donne

[60] Le nombre de groupes de mots-rimes est plus élevé que le nombre de séquences VCT, parce que certains groupes, qui ne partagent pas le même vocabulaire, n'ont pas été rapprochés. Il en est ainsi des deux groupes suivant caractérisés par le même rimème :

pour chaque voyelle phonétique, la richesse lexicale (nombre de mots-rimes ayant V comme noyau vocalique), le nombre d'occurrences de mots-rimes et le nombre de groupes de mots-rimes.



Clef de lecture : Il y a 2149 mots-rimes différents avec /e/ comme noyau vocalique, qui représentent 10084 occurrences. 207 groupes de mots-rimes différents ont été constitués.

Figure 11. Groupes de mots-rimes pour chaque voyelle

Cette figure montre qu'il y a plutôt moins de groupes de mots-rimes pour les voyelles phonétiques qui s'écrivent avec des digraphes ou trigraphes (partie droite de la courbe), cela nous renvoie à la question du rôle de la consistance graphique.

Groupe de mots-rimes	Rimème	Nb mots V	Nb occ N	Mot (fréquence à la rime sur CORRAC)
92	- ɔ - [-]	24	172	Chicanneau(2) anneau(4) bandeau(2) bateau(1) beau(25) berceau(7) bourreau(14) cerveau(3) chapeau(2) château(2) couteau(6) drapeau(1) eau(7) fardeau(6) flambeau(11) marteau(1) monceau(1) nouveau(27) panneau(2) peau(2) rideau(2) tableau(1) tombeau(40) vaisseau(3)
469	- ɔ - [-]	2	2	ciseau(1) fuseau(1)

Mais cette courbe montre surtout une corrélation très forte (r^2=0,9854) entre la richesse lexicale et le nombre de groupes de mots-rimes (*cf.* Figure 12).

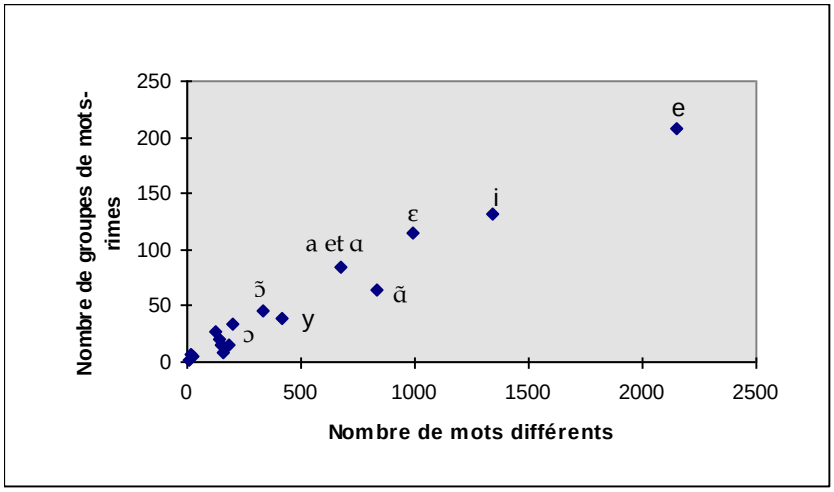


Figure 12. Corrélation entre nombre de groupes et richesse lexicale

Ces calculs sont grossiers mais témoignent de l'incidence de la richesse lexicale sur les groupes de mots-rimes, dont le nombre dépend en partie du rôle de la consonne d'appui. Une évaluation plus précise de notre hypothèse reste à faire mais nous avons quand même le sentiment que plus une séquence VCT est rare dans le corpus, plus il est difficile de trouver des couples de mots-rimes avec consonne d'appui identique. Ainsi, pour le rimème *or* (- ɔ r [-]), composé de 12 mots, la consonne d'appui n'est identique que dans un couple (*Castor-Nestor*).

Groupe de mots-rimes	Rimème	Nb mots V	Nb occ N	Mot (fréquence à la rime sur CORRAC)
159	- ɔ r [-]	8	42	Alidor(3) Hector(10) Nicanor(1) Rosidor(2) encor(18) mogor(1) or(2) trésor(5)
508	- ɔ r [-]	2	2	Castor(1) Nestor(1)
792	- ɔ r [-]	2	2	Encor(1) butor(1)

Clef de lecture : le premier chiffre correspond au numéro du groupe. Il est suivi par la transcription graphico-phonétique de la terminaison commune (rimème), puis par le nombre de mots dans le groupe (V), et par le nombre d'occurrences (N). Viennent ensuite tous les mots-rimes du groupe ainsi que leur fréquence.

Le degré de saillance des traits de la rime semble lié à la fréquence. Deux vers qui finissent en *or* se ressemblent plus que deux vers qui finissent en *ir*, parce que les mots rimes en *or* sont très rares dans le corpus, ainsi que

les occurrences des couples de mots-rimes (12 mots en *or* pour 23 couples, 123 en *ir* pour 777 couples). Il aurait fallu, comme nous l'a suggéré Benoît Habert, pouvoir évaluer la richesse lexicale des rimèmes non pas seulement en considérant les mots-rimes du corpus, mais en mobilisant tout le vocabulaire du corpus, voire en s'appuyant sur d'autres corpus ou sur des lexiques de l'époque. Cela serait un moyen d'évaluer la richesse lexicale des rimèmes en langue.

Pour l'écriture, quand le stock de mots avec une terminaison donnée est petit, il est plus difficile de trouver deux mots qui riment richement (un phonème au moins en plus de ceux nécessaires à la rime suffisante). Pour la lecture-audition, la reconnaissance de l'identité à la rime de deux mots aux terminaisons rares sera plus aisée que celle de mots aux finales très fréquentes. Il y aura de plus dans le deuxième cas un risque de confusion dans la reconnaissance des fins de vers puisque la probabilité de rimes internes sera plus grande.

III.4.5. Les glissantes dans la rime : diérèse et hiatus

Nous avons jusqu'à présent porté notre attention sur le rôle de la consonne d'appui. Quelle est la portée de la rime quand la voyelle cœur de la séquence VCT est précédée par une glissante ; phonème qui peut se réaliser soit comme une voyelle haute, dorénavant notée H, soit comme une semi-voyelle, notée G[61] ? La glissante est-elle systématiquement intégrée à la séquence VCT ou uniquement dans les cas de synérèse ?

Jusqu'à présent, même quand la consonne d'appui intervient, la rime classique semble trouver sa matière dans la dernière position métrique mais pas dans les précédentes. Cela semble montrer que les positions structurent non seulement le mètre mais aussi la rime.

Le rôle de la glissante dans la rime met à jour des cas de discordance entre positions métriques et rimes : ces cas nous renvoient à la question de la diérèse. Voici quelques exemples tirés de Corneille et Racine :

[61] Nous reprenons les notations H et G proposées par Klein [1991].

Mais si bien qu'à présent voulant vous mari/er
Vous ne trouveriez pas la fille d'un huissier

Corneille, *La suite du Menteur*, vers 309-310.

Et moi je ne veux plus, tant tu m'es odi/eux,
Partager avec toi la lumière des cieux.

Racine, *La Thébaïde*, vers 1177-1178.

Malgré tout son orgueil, ce monarque si fier
À son trône, à son lit daigna l'associ/er,

Racine, *Bajazet*, vers 467-468.

Dans ces jours solennels à l'orgueil dédi/és
Seule et dans le secret je le foule à mes pieds

Racine, *Esther*, vers 279-280.

.

Dans chacun de ces exemples ce qui constitue la rime correspond tantôt à deux positions métriques tantôt à une seule[62], (*ma-ri-er* vs *huis-sier*, *odi-eux* vs *cieux*, *fier* vs *associ-er*, *dédi-és* vs *pieds*...).

Ces exemples sont rares, particulièrement chez Corneille, mais ils méritent quelque attention. En effet, ce phénomène est déjà présent dans le *Roman d'Alexandre* où, dans les laisses rimées, alternent des cas de diérèse et de synérèse sur le mot-rime. Ainsi :

Car de lui commença li doners et li biens.
Il conquist les hermins, Persans et Hurï/ens

Le Roman d'Alexandre, vers 197-198.

Biens correspond à une syllabe, *Hu-rï-ens* a trois. La lecture de l'ensemble de la laisse montre que le *i* entre dans la définition de la rime, à moins que ce ne soit un yod qui s'introduit même dans les cas de diérèse.

On pourrait être tenté de considérer que la rime n'inclut pas la semi-voyelle. Or d'une manière quasiment systématique, la voyelle haute qui précède la dernière voyelle métrique est un élément constitutif de la rime. Par exemple, il n'existe aucun cas où un mot en *ier* rime avec un mot en *er* sans *i*, idem pour les terminaisons en *iez*, *ier*, *ié*.... Pour les terminaisons en *ons*, il n'y a qu'un unique cas de rime entre un mot en *ions* et en *ons*.

[62] *Cf.* les tableaux sur la diérèse dans le chapitre IV.1.

Tiens, voilà ton paîment.
L'INTIME.
Un soufflet ! Ecrivons :
Lequel Hiérome, après plusieurs rébellions,

Racine, *Les Plaideurs*, vers 416-417.

Nous n'avons trouvé que quelques rares cas de rimes en *on* et *ion* :

1	maison-provision	(R, Les Plaideurs, I, IV)
1	protection-nom	(C, Sertorius, II, I)
1	union-nom	(C, Sertorius, III, I)
1	on-exécution	(R, Les Plaideurs, I, VII)
1	donc-création	(R, Les Plaideurs, III, III)

Je prétends un peu plus, mais dans cette union
De votre lieutenant m'envieriez-vous le nom ?

Corneille, *Sertorius*, vers 952-953.

Avocat, Ah ! Passons au déluge. Avant donc
La naissance du monde et sa création

Racine, *Les Plaideurs*, vers 800-801.

Et je gagne ma cause. A cela que fait-on
Mon chicanneur s'oppose à l'exécution

Racine, *Les Plaideurs*, vers 213-214.

On notera que ces rimes qui violent la norme statistique sont pratiquement toujours issues des *Plaideurs*, la rime *donc* / *création* étant particulièrement non conforme. Tout se passe comme si Racine cherchait à marquer le plus possible l'écart entre le vers de tragédie et celui de comédie, encore un alexandrin, mais malmené à force d'être découpé et pauvrement rimé, et concordant moins que jamais avec la structure syntaxique.

De même, les rimes en *eu* sont distinctes des rimes en *ieu* et dans l'ensemble des rimes en *eux*, il n'y a que trois cas (sur près de 1 000) du type ieux/eux et 5 de type ueux/eux :

1	eux-glorieux	(C, *Œdipe*, V, V)
1	yeux-veux	(C, *Héraclius*, V, III)
1	yeux-vœux	(C, *Psyché*, V, III)
1	deux-incestueux	(R, *Britannicus*, IV, II)
1	incestueux-deux	(R, *La Thébaïde*, IV, I)
1	incestueux- vœux	(C, *Héraclius*, IV, III)
1	vertueux-eux	(C, *Horace*, V, III)
1	vœux-respectueux	(C, *Pulchérie*, III, III)

L'éclat de ces vertus que je ne tiens pas d'eux
Reçoit pour récompense un trépas glorieux.

Corneille, *Œdipe*, vers 1839-1840.

Soldats, sans plus tarder, qu'on l'immole à ses yeux ;
Et sois après sa mort mon fils, si tu le veux.

Corneille, *Héraclius*, vers 1693-1694.

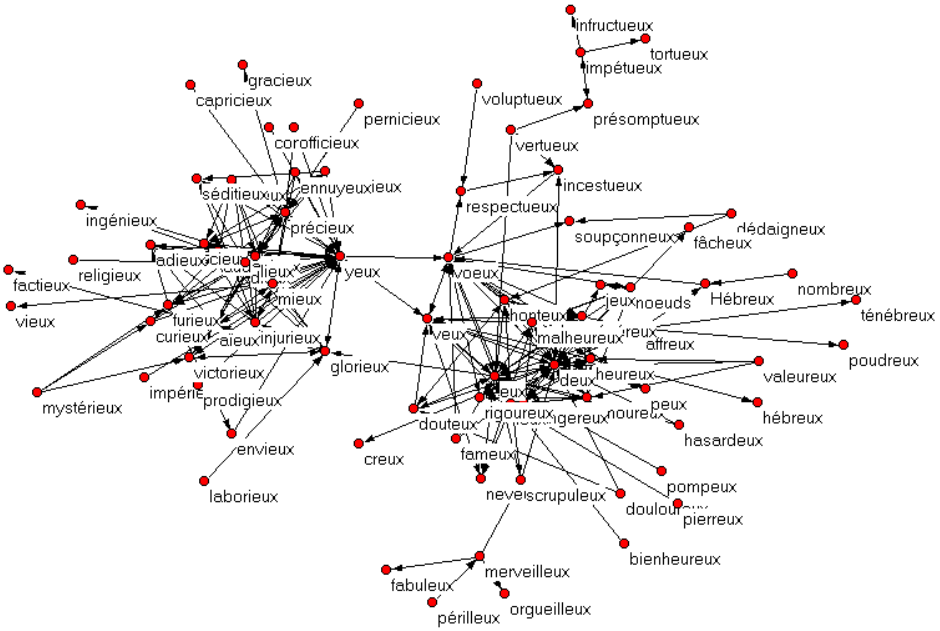


Figure 13. Groupe de mots-rimes en eux/ieux

Alors que la consonne d'appui est tantôt nécessaire tantôt non, la semi-voyelle intervient d'une manière beaucoup plus systématique dans la définition de la rime. Donc il nous faut considérer que la rime ne se limite pas à la séquence VCT, où V est le noyau de la dernière position, mais qu'elle fait intervenir la glissante (G) qui la précède. La rime porte sur GVCT. Or, dans les cas de diérèse, cela signifie que la rime s'appuie sur les deux dernières positions métriques, et non pas seulement sur la dernière.

Dans la transcription phonétique que nous avons adoptée dans le métromètre, les représentations que nous avons choisies pour différencier diérèse et synérèse semblent maximiser l'écart entre les deux. Dans des situations de diérèse, la représentation est de type (...H)(V...), tandis qu'en cas de synérèse, on a (...G V...), où H est la voyelle haute et G la semi-

voyelle. Ainsi *marier* est représenté par (m a)**(r i)(e)** tandis que *huissier* est représenté par (ɥ i)(**s j e**) ; dans cette transcription phonétique, ces deux mots n'ont comme partie commune que [e]. Or, il est tout à fait envisageable d'introduire une transcription de la diérèse avec présence de la semi-voyelle ainsi *marier* serait transcrit (m a)(**r i)(j e**). Ce choix de transcription permettrait d'enrichir d'une manière satisfaisante la définition de la rime. Ainsi le yod fait partie de la rime, qui ne s'appuie alors que sur la dernière syllabe. On maintient par le choix de transcription adopté une certaine cohérence entre la structure métrique et celle de la rime.

Le choix de cette transcription permettrait de plus de résoudre une des incohérences apparentes du vers dans le traitement du hiatus. Le hiatus externe était interdit dans le vers à partir de l'époque classique (« Tu es » n'est pas admis), tandis que le hiatus interne, celui que crée systématiquement la diérèse était tout à fait licite (« tuer » est admis). D'un côté, les frontières entre les mots sont atténuées par l'interdiction du hiatus externe, de l'autre côté, les diérèses multiplient les heurts vocaliques. Ceci présente quelque incohérence. En considérant que dans les cas de diérèse, il y a à la fois maintien de la voyelle haute et semi-vocalisation, on rend plus fluide la diérèse en introduisant entre les deux voyelles une glissante ([j],[w] ou [µ]). Au lieu de considérer que diérèse et synérèse s'opposent par l'absence ou la réalisation de la semi-vocalisation, elles se distinguent par le maintien ou non de la voyelle haute, la semi-vocalisation étant systématique.

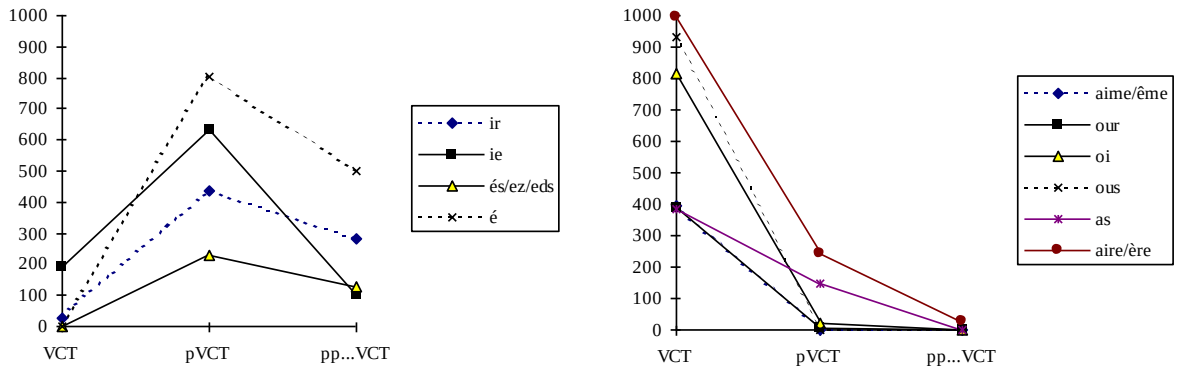
Le choix d'une représentation phonétique qui en cas de diérèse maintient la voyelle haute et la semi-voyelle permet au niveau phonétique d'enrichir la définition de la rime, tout en maintenant la règle selon laquelle la rime ne s'étend pas à gauche au-delà de la dernière syllabe.

Mais si, dans la transcription phonétique, cette modification permet de lever l'apparente incohérence entre la structure métrique et celle de la rime, qui s'appuie tantôt sur une ou sur deux syllabes, au niveau graphique cette incohérence est maintenue. Ainsi, se maintient une forme de discordance entre le découpage syllabique et la rime.

III.5. DEFINITION CONTEXTUELLE DE LA RIME PAUVRE, SUFFISANTE ET RICHE

Nous avons vu que l'extension de la rime variait en fonction de la terminaison. Comment peut-on alors qualifier la richesse de la rime, si celle-ci dépend de la séquence VCT ?

Prenons à titre indicatif quelques séquences VCT. Pour chacune des séquences, nous avons classé les couples selon que l'équivalence portait sur la séquence VCT, sur VCT plus un phonème avant et sur VCT plus au moins deux phonèmes avant. Les effectifs de couples de mots-rimes se répartissent comme le montre la Figure 14.



Clef de lecture : Sur l'ensemble des couples de mots-rimes finissant en *ir* (VCT=*ir*) dans CORRAC, il y a 29 couples pour lesquels la rime se limite à *ir*, 435 couples où la rime inclut un phonème d'appui à gauche de VCT (pVCT) et 280 couples où la rime porte au moins sur deux phonèmes supplémentaires (pp...VCT).

Figure 14. Richesse de la rime

Sur ces quelques exemples, apparaissent deux types de courbes de répartition différenciés : d'un côté, il y a des groupes de mots-rimes pour lesquels la présence du phonème d'appui est majoritaire, de l'autre des groupes pour lesquels sa présence est rare. Cela semble lié au nombre de mots de chaque VCT. Les premiers sont des groupes de mots-rimes de plus de 100 mots, les seconds ont moins de 50 mots (à l'exception des terminaisons en *aire*/*ère* dont le vocabulaire a plus de 100 mots et dont le profil se rapproche pourtant des séquences pauvres en vocabulaire). Selon la

fréquence de la terminaison dans le corpus, le seuil qui définit la rime suffisante varie.

Il ne peut donc y avoir de définition absolue de la qualité de la rime, elle se définit en référence à la langue pour un type de terminaison donné. Nous pouvons dire, parce que nous avons étudié les groupes de mots-rimes, que *désir* et *soupir* constituent une rime pauvre, *mourir* et *périr* une rime suffisante et *secourir* et *mourir* une rime riche, parce que la situation dominante (largement majoritaire) pour les rimes en *ir* est l'identité de la consonne d'appui. En revanche, *jour* et *amour* constituent une rime suffisante, et *séjour* et ***jour*** une rime riche (*cf.* Figure 14). La définition d'une nomenclature des rimes est donc éminemment complexe puisqu'elle fait intervenir de très nombreux paramètres : le type de voyelle, l'épaisseur graphique de la rime, la qualité syntaxique de la terminaison et la fréquence du rimème. Il reste à examiner pour chaque séquence de rime (ou rimème), jusqu'où s'étend la rime dans la majorité des cas pour définir dans le contexte ce qu'est la rime suffisante.

Maintenant que les différents critères ont été identifiés, on peut être tenté de proposer un modèle qui permet de déterminer si deux vers riment ensemble et quelle est la richesse de la rime. Les paramètres de ce modèle devront être définis empiriquement[63]. On pourrait aussi, c'est une approche

---

[63] A titre indicatif, nous proposons une fonction linéaire à seuils. Bien d'autres types de modèles pourraient être proposés et testés. Pour améliorer le modèle, il faudrait procéder à des tests économétriques pour définir les paramètres et tester diverses formes de modèles. Ce travail empirique, intéressant, dépasse le cadre de cette thèse. Soit la finale de vers VCT, P est le phonème d'appui, P-1 le phonème précédant le phonème d'appui....

$$F(a,b) = d(V_a,V_b)d(C_a,C_b)d(T_a,T_b) \times [1 + \alpha d(P_a,P_b) + \beta d((P-1)_a,(P-1)_b) + \gamma d((P-2)_a,(P-2)_b) + \delta l(a,b) - \varepsilon g(VCT_a,VCT_b)]$$

où d(i,j) = 1 si i et j sont équivalents ; = 0 si i et j ne sont pas équivalents.

g(i,j) = 1 si i et j sont deux morphèmes grammaticaux identiques ; =0 si i et j ne sont pas deux morphèmes grammaticaux identiques ;

l(i,j) = longueur de la séquence de lettres communes entre i et j[63] ;

α, β, γ, δ, ε sont des paramètres à estimer.

F(a,b) = 0 : il n'y a pas rime

On teste dans un premier temps l'équivalence de VCT, ensuite on examine l'équivalence des phonèmes antérieurs à VCT : plus le nombre de phonèmes équivalents augmente, plus la valeur de la fonction croît. Puis la dimension graphique est prise en compte, à travers la longueur de la séquence de lettres communes. Enfin, la nature syntaxique de la terminaison est examinée : s'il s'agit de deux morphèmes grammaticaux identiques, la valeur de F(a,b) décroît.

alternative, représenter la terminaison de chaque mot-rime par un vecteur de traits phonétiques et graphiques[64] et mesurer la distance entre les mots-rimes de chaque couple, en fonction du nombre de traits communs. Cela permettrait d'évaluer pour chaque rimème la distance moyenne qui pourrait définir l'extension de la rime suffisante.

Alors seulement, on pourrait entreprendre des études diachroniques. À partir de quand les voyelles ont-elles commencé à avoir un statut différent pour la rime ? Quand est-ce que le /e/ est devenu trop faible pour constituer une rime à lui seul ?… En étendant l'investigation aux œuvres de Hugo, Baudelaire, Rimbaud et Mallarmé (disponibles sur support électronique), on pourrait également valider ou invalider l'hypothèse souvent avancée sur l'évolution de la rime au XIXe siècle. Les métriciens considèrent généralement que la rime s'est enrichie au XIXe siècle, comme si elle cherchait à compenser l'affaiblissement du système métrique interne et en particulier la discordance croissante entre mètre et syntaxe.

Cornulier [1985] a fait une tentative dans ce sens en comparant quelques vers du XIXe siècle à *Phèdre*. Il montre que la progression de la

---

Il faudrait ensuite définir pour chaque VCT les seuils (rime pauvre, suffisante, riche) en fonction de la voyelle cœur de rime et du nombre de mots ayant une terminaison.

Prenons quelques exemples de terminaisons en *ir* et *our*, en fixant $\alpha = 1$, $\beta = 1$, $\gamma = 1$, $\delta = 0{,}1$, $\varepsilon=0{,}8$.

F(*attendrir*, *rougir*) = 1*1*1[1+0+0+0+0,2-0,8] = 0,4 (*ir* terminaison d'infinitif)

F(*désir*, *soupir*) = 1*1*1[1+0+0+0+0,2-0] = 1,2

F(*mourir*, *périr*) = 1*1*1[1+1+0+0+0,3-0,8] = 1,5 (*ir* terminaison d'infinitif)

F(*secourir*, *mourir*) = 1*1*1[1+1+1+0+0,5-0,8] = 2,7

F(*avenir*, *tenir*) = 1*1*1[1+1+1+0+0,4-0] = 3,4

F(*jour*, *amour*) = 1*1*1[1+0+0+0+0,3-0] = 1,3

F(*séjour*, *jour*) = 1*1*1[1+1+0+0+0,4-0] = 2,4

Pour VCT = ir : si F(a, b) = 0 a ne rime pas avec b ; si 0 < F(a, b)<= 1, la rime a, b est pauvre ; si 1<F(a, b)<2 la rime est suffisante et si F(a, b) > = 2 elle est riche.

Pour VCT = *our :* si 0 < F(a, b) < 2, la rime est suffisante, si F (a, b) >= 2, elle est riche.

La définition des seuils pourrait être modélisée en tenant compte de la nature de la voyelle, de la richesse lexicale de la terminaison et du nombre de lettres de la séquence VCT.

Ce modèle n'est présenté qu'à titre d'orientation : la forme linéaire a été ici adoptée car elle est plus facile à manier. En l'état, notre travail permet seulement de pointer les paramètres qui constituent la structure feuilletée de la rime, et qui définissent sa qualité.

[64] Nous devons cette suggestion à Benoît Habert.

rime riche n'est pas très marquée, bien que le *sentiment* d'enrichissement de la rime soit fort. Il propose une hypothèse interprétative :

> Il se pourrait donc que l'idée traditionnellement admise suivant laquelle il y a chez les poètes romantiques un développement considérable de la rime riche repose sur un faible développement de la rime riche, et un développement considérable de l'*effet* de rime riche.
>
> [Cornulier, 1985, p. 120].

Pour confirmer cette hypothèse, il faudrait comparer des corpus de taille plus grande avec des critères plus fins que ceux qui ont été utilisés (Cornulier a conscience des limites de ses décomptes : pour lui une rime est riche si l'identité porte sur l'avant-dernière voyelle métrique, ce qui est une définition très élargie de la rime riche, qui ne prend pas en compte les variations selon le contexte vocalique[65]). Le RIMARIUM devrait faciliter ce genre d'investigation.

III.6. CONCLUSION

Ce premier examen de la rime chez Corneille et Racine démontre que celle-ci est à la fois graphique et phonétique : elle est bien faite pour un œil-oreille.

L'exploration des groupes de mots pouvant rimer ensemble a permis d'identifier les équivalences graphiques et phoniques en fin de vers, et d'abstraire pour chaque groupe un *rimème*, c'est-à-dire une séquence graphico-phonique commune à tous les mots-rimes du groupe.

Les éléments *nécessaires, mais pas toujours suffisants*, pour que deux vers riment sont *l'équivalence graphico-phonétique* des deux séquences VCT, c'est-à-dire des dernières voyelles métriques, V, des consonnes ou groupes consonantiques, C, qui les suivent et des terminaisons, T, (féminines/masculines ; singulières/plurielles). Le cœur de la rime est la dernière voyelle métrique ; C et T peuvent être réduits au vide. La notion d'équivalence graphico-phonétique est plus adéquate que celle d'identité pour définir la rime, puisqu'il y a des équivalences de graphèmes et de phonèmes.

---

[65] Une rime peut être dite riche si elle inclut la consonne d'appui pour certaines voyelles, mais pas pour d'autres.

L'équivalence VCT n'est pas suffisante dans deux cas :

- quand les deux séquences VCT correspondent au même *morphème grammatical* (*-ait, -er, -ant...*), surtout quand il est fréquent. Cet aspect est souligné dans les traités et manuels de métrique ;
- quand la dernière voyelle métrique est un [e] suivi uniquement d'une terminaison dans l'ensemble {Ø, e, s, es}[66]. Cette quasi-règle ne figure pas dans les traités sur le vers.

Dans ces deux cas, pour qu'il y ait rime, dans le vers classique, il faut obligatoirement que le phonème d'appui (consonne ou glissante) soit identique (*envo**ler*** / *par**ler*** ; *chance**lant*** / *par**lant*** ; *irri**té*** / *autori**té** ; sacri**fié*** / *all**ié*** ; *desti**nées*** / *an**nées***...). La rime serait trop banale s'il n'y avait pas l'identité du phonème d'appui, les terminaisons grammaticales et les mots en *é, és, ée, ées* étant très fréquents.

Nous avons pour la première fois montré que l'extension de la rime variait selon la séquence finale (VCT) et essayé de fournir quelques éléments d'explication, ou du moins de description, de ces variations. En général, sauf dans les deux cas vus précédemment, l'équivalence de la dernière voyelle métrique et de son environnement droit (VCT) peut être suffisante pour qu'il y ait rime. Mais elle tend à prendre appui sur la gauche quand l'épaisseur (graphique, phonique et/ou syntaxique) de la séquence VCT est insuffisante. Nous avons pu voir que, en fonction de chaque type de VCT, la probabilité que la rime soit enrichie par l'identité d'au moins un phonème à gauche de VCT varie. Les facteurs qui semblent avoir une incidence sur le recours au phonème d'appui sont :

- la nature de la voyelle de la dernière position métrique. Nous pouvons ainsi construire une échelle des voyelles : plus le poids de la voyelle est faible, plus le recours à la consonne d'appui est nécessaire.
- l'espace graphique occupé par VCT : plus le nombre de lettres de la séquence VCT est grand, moins le recours au phonème d'appui est nécessaire ;

[66] Nous avons trouvé sur notre corpus trois exceptions sur plus de 2 500 couples de rimes en é : *envoyé/Pasiphaé* ; *frappée*/*tombée*, réservée/*Nymphée* et Jacques Roubaud en a trouvé six sur une trentaine de textes examinés (sans compter celles que l'on trouve chez La Fontaine qui sont beaucoup plus nombreuses).

- le nombre de mots dans le lexique ayant une terminaison VCT donnée : plus une séquence VCT est rare dans le lexique, moins le phonème d'appui est nécessaire ;
- la nature du phonème précédant la dernière voyelle métrique. En effet, la probabilité que ce phonème intervienne dans la rime est très grande (proche de 1) quand le phonème est une glissante.

Tous ces facteurs vont dans le même sens. Plus la finale VCT est remarquable, saillante, moins la rime a besoin de s'étendre sur la gauche. Intuitivement, on dirait qu'il faut que la ressemblance entre les deux fins de vers soit « suffisante » pour qu'il y ait rime. Tout se passe comme s'il fallait atteindre une masse critique, un cumul de traits similaires, pour qu'il puisse y avoir rime. Or le même ne se définit que par contraste avec le différent : quelques traits *rares* en commun sont suffisants ; si les traits en commun sont banals, il en faut une plus grande quantité.

IV. DIERESE, E "MUET", LIAISON : TROIS LIEUX DE VARIATION PHONOLOGIQUE[67]

Nous avons vu que le métromètre, l'outil d'analyse du vers, devait répondre à deux objectifs : l'identification des positions métriques et la description de ces positions. Nous nous intéressons dans ce chapitre au premier objectif. Celui-ci consiste à reconnaître les noyaux vocaliques qui constituent le cœur de chaque position métrique et, en second lieu, à délimiter les frontières des positions métriques, c'est-à-dire à rattacher au noyau vocalique les consonnes qui lui correspondent. Ces opérations qui pourraient être simples ne le sont pas. En effet :

- les groupements vocaliques (suites de deux ou trois voyelles) correspondent selon les cas à une ou deux positions métriques ;
- le *e* muet est un segment instable qui correspond ou non à une position métrique ;
- les consonnes de liaison (ou latentes) sont instables. Leur promotion peut modifier le dénombrement et modifier le contour des syllabes.

La question de la syllabe

L'identification systématique des positions métriques nous renvoie à la question de la syllabe. Les travaux générativistes lui accordent une faible place. Chomsky et Halle dans *Sound Pattern of English* [1968, 1973] considèrent que le concept de syllabe n'est pas nécessaire en phonologie.

Il existe des travaux importants, en particulier en psycholinguistique, qui lui accordent au contraire une place importante. Les travaux

[67] Ce chapitre reprend et approfondit des parties de l'article sur le métromètre [Beaudouin et Yvon, 1994 et 1996].

d'Emmanuel Dupoux et Jacques Mehler sur la segmentation de la parole montrent que, dans le flot acoustique, la reconnaissance des syllabes est plus rapide que celle des phonèmes. Leurs expériences consistent à demander à des locuteurs de réagir dès qu'ils reconnaissent une syllabe ou un phonème. La syllabe, en français du moins, semble jouer « un rôle primordial dans la représentation mentale de l'onde parlée » [Dupoux et Mehler, 1992].

Cette notion de syllabe est utile dans le cadre de l'étude du vers français puisque celui-ci est défini par un nombre de syllabes. Nous reprenons le modèle de la syllabe proposé par Kaye et Lowenstamm [1984, p. 124] :

> La structure syllabique est analysée [par Kaye et Lowenstamm] en deux constituants : l'attaque et la rime[68], laquelle s'analyse à son tour comme un constituant obligatoire, le noyau, et un composant facultatif, la coda, qui comme son nom l'indique, est située après le noyau.
>
> [Encrevé, 1988, p. 138].

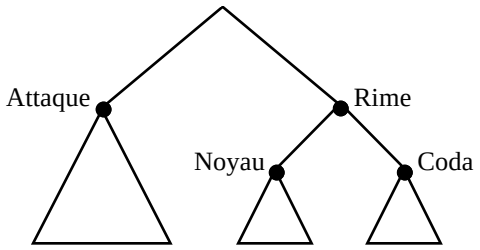


Figure 15. Le modèle de la syllabe d'après Kaye et Lowenstamm [1984]

Voilà par exemple comment se décompose en constituants syllabiques le mot *école*. Certains constituants sont vides, comme par exemple l'attaque de la première syllabe de *école*.

[68] Tel est le terme utilisé par Kaye et Lowenstamm, un peu malheureux dans notre contexte, car comme nous le verrons, il arrive souvent que l'attaque fasse partie de la rime métrique.

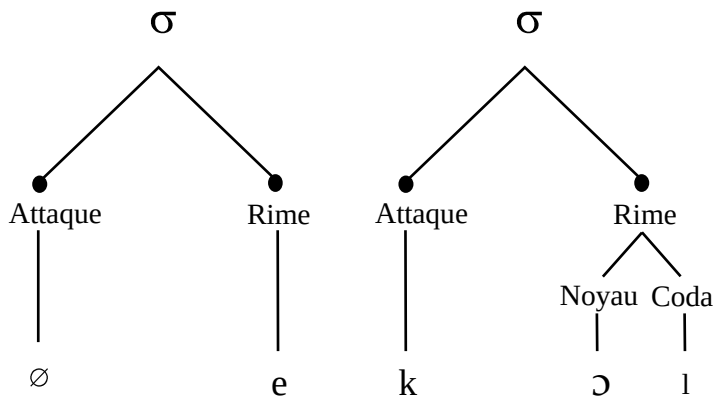


***Figure 16. Constituants syllabiques de* école**

Voilà, semble-t-il, pour ce qui est des certitudes. La composition précise de ces attaques, noyaux et coda est beaucoup plus controversée, suivant le degré de réalité accordé à cette unité de découpage, et suivant les niveaux d'analyse et de représentation (syllabe lexicale, syllabe phonémique, syllabe phonétique…).

Le noyau vocalique est la seule composante indispensable de la syllabe, attaque et coda sont optionnelles. D'ailleurs, pour certaines des questions qui nous occupent, un modèle complexe de la syllabe n'est pas nécessaire, puisque pour le dénombrement, et donc la reconnaissance d'un mètre, il suffit d'identifier les noyaux vocaliques. Dans le système de décompte proposé par Milner [1974], et que nous reprenons en partie, « l'élément phonologique crucial est la voyelle » [p. 6]. Ainsi, dans le vers suivant, pour l'identification et le décompte des positions métriques, la reconnaissance des noyaux vocaliques est suffisante. Les parenthèses délimitent les positions métriques.

Modérez ces transports, voici venir l'infante.

(m o̲)(d e̲)(r e̲)(s e̲)(t r ɑ̲̃ s)(p ɔ̲ r)(v w a̲)(s i̲)(v ə)(n i̲ r)(l ɛ̲̃)(f ɑ̲̃ t)

Corneille, *Le Cid*, vers 1142.

Le rattachement des consonnes au noyau vocalique

Laissons provisoirement de côté la question de l'identification des noyaux vocaliques.

Les voyelles étant toutes identifiées et correspondant au centre des syllabes, la difficulté suivante réside dans le rattachement des consonnes. Voici les conventions de découpage syllabique que nous avons adoptées[69].

A l'intérieur d'un mot, le découpage respecte au mieux le schéma suivant : une consonne simple entre voyelles est toujours rattachée à la voyelle qui la suit. Une consonne à l'initiale d'un mot est également rattachée à la voyelle qui la suit. Les groupes de deux consonnes sont toujours scindés (*funes/te)*, sauf pour les groupes formés d'une obstruante et d'une liquide[70] (*bras*). Ce découpage se fait au nom du *Maximum Onset Principle*, qui tend à maximiser l'attaque de la syllabe à condition que celle-ci soit acceptable en début de syllabe [Durand, 1990, p. 198-199].

Et son bras valeureux n'est funeste qu'à moi,

(e)(s ɔ̃)(b r ɑ)(v a)(l œ)(r ø)(n ɛ)(f y)(n ɛ s)(t ə)(k a)(m w a)

Corneille, *Le Cid*, vers 1152.

Les groupes de trois consonnes ou plus sont rares, et sont (dans notre corpus de vers) toujours scindés après la première consonne (*cf. mar/bre*).

Et mouillant de ses pleurs le marbre de ses pieds,

(e)(m u)(j ɑ̃)(d ə)(s e)(p l œ r)(l ə)(m a r)(b r ə)(d ə)(s e)(p j e)

Racine, *Britannicus*, vers 1729.

Enfin, les consonnes finales (non latentes) sont toujours syllabées avec la voyelle précédente. Ce dernier point est sans aucun doute le plus sujet à caution. Il est admis qu'en langue, au sein d'un groupe syntaxique, des phénomènes d'enchaînement puissent avoir lieu, entraînant le rattachement de consonne(s) finale(s) à la voyelle initiale du mot suivant. L'enchaînement et le non enchaînement de ces consonnes fixes sont très variables mais, selon Pierre Encrevé, « ces phénomènes phonétiques n'ont jamais été constitués (...) en enjeu social » [1988, p. 29], contrairement à la liaison. Nous avons pris l'option de ne pas enchaîner les consonnes finales à la voyelle du mot suivant. Ainsi dans le vers suivant, le *r* final de *devoir* est maintenu dans sa syllabe d'origine au lieu d'être rattaché à la suivante.

---

[69] Nous avons repris les règles de découpage conventionnelles, sans essayer d'en construire la théorie.

[70] Pour être précis, il s'agit des groupes {pbtdkgfv} + r, et des groupes {pbkgfv} + l.

Cependant mon devoir est toujours le plus fort,

(s ə)(p ɑ̃)(d ɑ̃)(m õ)(d ə)(v w a_r)(ɛ)(t u)(ɜ u r)(l ə)(p l y)(f ɔ r)

au lieu de (vw a)(r ɛ)

Corneille, *Le Cid*, vers 1167.

La correspondance entre fin de syllabe et fin de mot (lexical) est donc systématique dans notre système de transcription, sauf en un cas précis : la liaison, que nous examinerons en détail dans la section IV.3 de ce chapitre.

Les noyaux vocaliques

Il existe en français trois lieux de variation phonologique qui ont une incidence directe sur l'identification des composantes syllabiques : le *e* muet, la liaison et la diérèse. En langue, les réalisations du *e* muet, de la liaison et de la diérèse (*versus* synérèse) varient selon de multiples facteurs, tels que l'origine du locuteur, le contexte de l'élocution, l'état émotionnel, la vitesse de locution, le niveau de langue...

Un *e* muet, s'il est prononcé, devient le noyau vocalique d'une syllabe ; s'il n'est pas prononcé il disparaît dans la représentation phonétique.

Une femme mourante :
(y)(n ə)(f a)(m ə)(m u)(r ɑ̃ t) ou (y n)(f a m)(m u)(r ɑ̃ t)

De même, une consonne de liaison suivie par un mot à initiale vocalique sera soit éliminée dans la représentation phonétique, si la liaison n'est pas réalisée, soit rattachée à la première syllabe du mot suivant (enchaînement), soit maintenue attachée à la dernière syllabe de son mot d'origine (liaison sans enchaînement).

Vous semblez interdite :
(vu)(s ɑ̃)(b l e)(ɛ̃)(t ɛ r)(d i t),
ou (vu)(s ɑ̃)(b l e)(z ɛ̃)(t ɛ r)(d i t)
ou (vu)(s ɑ̃)(b l e z)(ɛ̃)(t ɛ r)(d i t)

Enfin, selon que, dans un groupement vocalique commençant par une voyelle haute (*i*, *ou*, *u*), on a ou non semi-vocalisation de la voyelle haute (synérèse/diérèse), une ou deux syllabes sont délimitées.

Lyon : (li)(ɔ̃) ou (l j ɔ̃) ;
remercier : (r ə)(m ɛ r)(s i)(e) ou (r ə)(m ɛ r)(s j e)

Dans une chanson contemporaine, un même suffixe peut constituer une ou deux syllabes :

Notre père qu'êtes audacieux

(n ɔ)(t r ə)(p ɛ r)(k ɛ t)(z o)(d a)(s j ø)

Seriez-vous insidi-eux

(s ə) (r j e)(v u)(z ɛ̃)(s i)(d i)(ø)

Alain Bashung, « Etrange été ».

Ainsi, les variations du nombre et du contour des syllabes phonétiques résultent de ces trois lieux de variation phonologique.

En langue, le *e* muet, la diérèse et la liaison sont des lieux de variations dont les règles sont incertaines et difficiles à formaliser, comme en témoigne les travaux qui leur sont consacrés (par exemple Dell [1984] sur le *e* muet, Encrevé [1988] sur la liaison et le *e* muet et Klein [1991] sur la diérèse / synérèse). On pourrait naïvement croire que ces variations sont subtilement utilisées par le poète comme des facteurs d'élasticité du vers[71]. Il n'en est rien ; si les variations existent dans le vers, elles sont régies par des règles contextuelles strictes, heureusement plus simples à formaliser que pour la prose.

Voyons ces variations. Certains *e* sont maintenus, d'autres non. Dans le vers suivant le *e* final de *Œnone* est maintenu tandis que celui de *couronne* est élidé.

Œnone, fais briller la couronne à ses yeux,

(e)(n ɔ)(n $ə_f$[72])(f ɛ)(b r i)(j e)(l a)(k u)(r ɔ n)(a)(s e)(z_l[73] y ø)

Racine, *Phèdre*, vers 800.

Des groupements vocaliques comptent tantôt pour une position métrique tantôt pour deux. Ainsi, dans les vers suivants, la séquence *ien* correspond tantôt à une seule position métrique (*siens*), tantôt à deux (*li/ens*) :

[71] Certains poètes amateurs le croient volontiers, et construisent l'identité métrique en jouant librement sur les variations. Dans ce cas la réalisation orale devient indispensable puisque la métrique est trop personnelle.

[72] $ə_f$: désigne un *e* muet faible.

[73] Un petit *l* en indice indique qu'il s'agit d'une consonne de liaison.

Souffrez que de vos cœurs rapprochant les liens,

(s u)(f r e)(k ə)(d ə)(v o)(k œ r)(r a)(p r ɔ)(ʃ ɑ̃)(l e)**(l i)(ɛ̃)**

Je me cache à vos yeux, et me dérobe aux siens.

(ʒ ə)(m ə)(k a ʃ)(a)(v o)(zɺ j ø zɺ)(e)(m ə)(d e)(r ɔ b)(o)**(s j ɛ̃)**

Racine, *Britannicus*, vers 1073-1074.

Même la liaison peut jouer un rôle métrique, comme dans le vers suivant :

Du choix d'un successeur Athènes incertaine,

(d y)(ʃ w a)(d œ̃)(s y k)(s e)(s œ r)(a)(t ɛ)**(n ə zɺ)(ɛ̃)**(s ɛ r)(t ɛ n)

Racine, *Phèdre*, vers 485.

En effet, si on ne fait pas la liaison, il manque une position métrique :

Athènes incertaine => a t ɛ **n ə z # ɛ̃** s ɛ rt ɛ n => a t ɛ **n ɛ̃** s ɛ r t ɛ n

=> (a)(t ɛ)**(n ɛ̃)**(s ɛ r)(t ɛ n)

Pour le dénombrement correct du nombre de positions des alexandrins, la connaissance et l'application de ces règles est indispensable. Dans cette partie, nous allons examiner en détail les règles de la diérèse, du *e* muet et de la liaison et la manière dont elles ont été intégrées dans le métromètre. Il s'agit bien de règles de dénombrement que l'interprétation peut évidemment contourner.

IV.1. LA DIERESE

Les principes qui définissent la position métrique sont directement inspirés de ceux qui définissent la syllabe de langue. Pour traiter la question de la diérèse, qui concerne les noyaux vocaliques, il nous suffit de prendre le modèle le plus simple possible de la syllabe. Celui proposé par Roubaud [1986] est largement suffisant : la syllabe de langue contient un noyau vocalique V entouré de non-voyelles[74], N : NVN (N peut être réalisé par une ou plusieurs consonnes, par une consonne et une semi-voyelle ou par aucun élément).

Considérons la correspondance entre les voyelles graphiques (à l'exception du *e* muet) et le V du modèle, autrement dit les règles de transformation du modèle abstrait en modèle concret :

[74] Ici, la composition exacte de l'attaque et de la coda ne sont pas indispensables.

1. Une voyelle seule entourée de consonnes est en relation bijective avec le noyau vocalique V du modèle (c'est le cas idéal : dans *visibilité*, chaque V est réalisé par une voyelle).

Quand on a affaire à des groupements de voyelles, divers cas se présentent :

2. Le groupement correspond à un seul noyau vocalique V. C'est le cas pour *au*, *eau*, *eu*, *ou*..., pour des groupements vocaliques commençant par *u* après *g* ou *q* (*qui*, *guère*...) et pour des groupements contenant un *e* « inerte », comme dans *chant**ai**e**nt* ou *remerci**e**ment*[75]. Ceci est également le cas pour les séquences *oi* et *oin* qui correspondent au niveau phonétique à des diphtongues de forte cohésion qui ne donnent jamais lieu à une diérèse dans le vers, ni dans la prose sauf à des fins stylistiques marquées [*cf.* Klein, 1991].

3. Le groupement correspond à plusieurs noyaux vocaliques V. C'est le cas pour *cré-ation*, *ré-el*, *ha-ï*... En général, l'une des voyelles porte un accent graphique.

4. Le groupement commence par une voyelle *i*, *u* ou *ou* : *ia*, *io*, *ue*, *ui*, *oua*.... Selon le contexte, l'origine géographique du locuteur, la situation d'énonciation, le niveau de langue, cette voyelle initiale correspondra phonétiquement à une voyelle haute ou à une semi-voyelle. On aura soit une réalisation des deux voyelles (diérèse : maintien de la voyelle haute), soit d'une seule (synérèse : semi-vocalisation de la glissante *i*, *u*, *ou*). Ainsi, le mot *hier* aura pour transcription phonétique (i)(ε r) en cas de diérèse et (j ε r) en cas de synérèse. En cas de synérèse, la dernière voyelle devient le noyau V tandis que la semi-voyelle (ou semi-consonne) est assimilée au groupement N. Dès l'époque classique, il y a eu une tendance importante à la semi-vocalisation de la voyelle haute. Mais aujourd'hui encore, on trouve des locuteurs qui prononcent les deux voyelles dans *Ly-on*, *vi-olence*... Marc Klein [1991] part justement de l'hypothèse que dans les « parlers parisiens » le traitement des glissantes (voyelles hautes/semi-voyelles) est soumis à d'importantes variations. Dans le vers, la situation est également complexe, quoique

[75] Ces *e* ne sont devenus inertes qu'à partir du XVIe siècle. Au Moyen-Age, ces *e* internes comptaient.

plus systématique[76] et l'alternance entre diérèse et synérèse est également constatée.

C'est le quatrième point qui nous importe ici parce que l'alternance diérèse / synérèse a une incidence sur l'identification des noyaux vocaliques et sur le décompte. **Faire la diérèse consiste à dissocier dans la prononciation deux voyelles consécutives dont la première est une glissante** (voyelle haute *i*, *u* ou *ou*).

L'alternance diérèse/synérèse mérite un traitement spécifique pour plusieurs raisons : parce qu'elle fait intervenir des niveaux complexes de langue (lexique et syntaxe) et parce qu'elle a une incidence importante sur le dénombrement du vers.

Les cas de diérèse et de synérèse ne se définissent pas d'une manière simple. On aurait pu imaginer que la diérèse se fasse systématiquement sur des groupements particuliers, mais ce n'est pas le cas :

Et que j'entende dire aux peuples ind**ie**ns
(e)(k ə)(j ã)(t ã)(d əʃ)(d i r)(o)(p ø)(p l ə ʒ)(ɛ̃)(d **i)(ɛ̃)**

Racine, *Alexandre le Grand*, vers 11.

Hé b**ie**n, je l'avouerai, que ma juste colère
(e)(b **j ɛ̃**)(j ə)(l a)(v u)(r ɛ)(k ə)(m a)(j y s)(t ə)(k ɔ)(l ɛ r)

Racine, *Alexandre le Grand*, vers 233.

On aurait pu au moins espérer que pour une même forme graphique, on ait le même traitement. Ce n'est pas non plus le cas : les homographes[77] *fier* adjectif et *fi-er* verbe n'ont pas le même comportement dans le vers :

[76] Nous verrons qu'en dehors de rares exceptions, les auteurs adoptent les mêmes règles de décompte.

[77] Ils ne sont certes pas homophones, mais en partant de la forme écrite, seule une analyse syntaxique permet de différencier les deux mots.

Qui dès qu'à ses regards elle ose se f**ie**r,
(k i)(d ɛ)(k a)(s e)(r əf)(g a r zl)(ɛ l)(o)(z ə)(s ə)**(f i)(e)**

Racine, *Britannicus*, vers 421.

Des sentiments d'un cœur si f**ie**r, si dédaigneux,
(d e)(s ɑ̃)(t i)(m ɑ̃)(d œ̃)(k œ r)(s i) (**f j ɛ r**)(s i)(d e)(d ɛ)(ɲ ø)

Racine, *Phèdre*, vers 68.

Seule une bonne désambiguïsation morpho-syntaxique des homographes hétérophones permettra de distinguer ces deux cas.

Dans *Dire le vers*, Milner et Regnault [1987] n'évoquent à aucun moment le rôle de la diérèse dans la diction du vers. D'ailleurs, les mises en scène que nous avons vues montrent qu'il s'agit d'un phénomène soumis à une forte variation : tandis que certains comédiens distinguent nettement chacune des voyelles du groupement vocalique en cas de diérèse, d'autres les traitent comme une seule unité vocalique avec un léger allongement. Ainsi dans la mise en scène d'*Horace* par Marion Bierry en 1998 au Théâtre de l'Œuvre, chaque cas de diérèse donnait lieu à une distinction appuyée des deux voyelles, tandis que dans celle de Jean-Pierre Miquel en 1974, à la Comédie-Française, les diérèses n'étaient marquées que par un léger allongement du groupement vocalique. Mais il ne s'agit là que de choix de diction.

Est-elle un phénomène suffisamment important pour que l'on s'y attarde ? A-t-elle une forte incidence sur le décompte ? Nous allons voir que oui.

Sur l'ensemble du corpus de Corneille et Racine, mais aussi des différentes œuvres de Frantext qui ont complété ce corpus, nous avons repéré tous les mots comprenant des groupements susceptibles de porter la diérèse (groupements vocaliques commençant par *i*, *u* et *ou*). Nous désignons ces mots comme étant à « diérèse potentielle ». Ensuite, nous avons identifié parmi les mots à diérèse potentielle ceux qui portaient effectivement la diérèse (« diérèse attestée »).

	Global	Diérèse potentielle	Diérèse attestée
V	32724	10,8 %	5,6 %
N	1062777	4,8 %	1 %
N/V	32,5	14,3	5,8
V1/V	41,6 %		50,4 %

V : nombre de vocables (étendue du vocabulaire)[78] ;
N : nombre d'occurrences ou de mots (étendue du texte) ;
V 1/V : pourcentage d'hapax (mots de fréquence 1).

Clef de lecture : 10,8 % du vocabulaire comportent un groupement vocalique commençant par une voyelle haute (groupement vocalique susceptible de porter une diérèse). 5,6 % du vocabulaire (soit un peu plus de la moitié) porte effectivement une diérèse. Ce tableau ne correspond qu'aux groupements vocaliques commençant par *i*, *u* ou *ou* (y compris les cas où intervient la règle des trois consonnes). Ont été exclus les groupements avec *e* muet (du type *vie*, *vue*, *remercieront*, *partaient*...), les groupements *eui*, *uei*, *oui* devant */j/* (*deuil*, *accueil*, *bouilloire*, etc.) et les groupements vocaliques uV après *q* ou *g*.

Figure 17. Part du vocabulaire affecté par la diérèse

Premier constat : la diérèse ne concerne qu'une faible partie du vocabulaire (5,6 %) et une plus faible partie encore des occurrences (1 %). Pourtant, le phénomène de la diérèse n'est pas négligeable pour le décompte. En moyenne, un alexandrin est composé de huit mots (*cf.* chapitre VIII). En faisant la diérèse systématiquement sur tous les mots à diérèse potentielle, un vers sur trois en moyenne sera mal compté. En ignorant le phénomène, en ne faisant que des synérèses, un vers sur douze sera mal dénombré. Le moindre mal est d'ignorer tous les cas de diérèse mais cela conduit à un taux d'erreur non négligeable (8 %). D'où la nécessité d'étudier en détail les cas de diérèse.

Second constat : les mots portant une diérèse sont des mots rares. Ils ont une fréquence moyenne très faible par rapport à l'ensemble du vocabulaire (5,8 contre 32,5) et le taux d'hapax (mots n'apparaissant qu'une seule fois) est sensiblement plus élevé (50,4 % contre 41,6 %). On peut penser que ces mots, parce qu'ils sont plus rares, ont mieux su conserver leur forme complexe. Ceci semble confirmer l'hypothèse sur la tendance à la synérèse que nous verrons ultérieurement : plus un mot est fréquemment employé, plus la tendance à la synérèse est forte.

[78] Pour qu'il n'y ait pas de confusion avec le signe V, désignant une voyelle, ni avec N, désignant les non-voyelles, nous avons utilisé une police spécifique pour le V (*V*)et le N (N) de Muller [1977] qui correspondent au nombre de mots différents et au nombre d'occurrences dans un corpus.

A titre de remarque, il est assez étonnant de constater que la diérèse crée un hiatus interne - la rencontre de deux voyelles -, alors que le hiatus n'apparaît pas dans le vers classique quand les deux voyelles appartiennent à deux mots différents (hiatus externe). C'est Malherbe qui a imposé cette nouvelle contrainte au vers, formalisée dans *L'Art Poétique* de Boileau :

> Gardez qu'une voyelle à courir trop hâtée
> Ne soit d'une voyelle en son chemin heurtée
>
> Boileau, *L'Art Poétique*, Chant Premier.

Pourtant Malherbe lui-même a commis dans sa jeunesse des hiatus externes, qu'il se gardera bien de rééditer...

> Je demeure en danger que l'âme qui est née
> Malherbe, *Poésies*, « Les larmes de Sainct-Pierre, Imitées du Tansille. Au Roy », vers 250.

Ce hiatus externe est évité dans le vers classique et sera rejeté par son auteur lui-même alors que le hiatus interne (rencontre de deux voyelles à l'intérieur d'un mot)

> Avant qu'il me cherchât, un orgueil inqui-et
>
> Racine, *Alexandre le Grand*, vers 239.

ne pose aucune difficulté.

L'absence de hiatus externe dans le vers classique français n'est pas simple à interpréter. En tout cas, elle ne peut être mise sur le compte de la recherche d'une euphonie, contredite par les hiatus internes. On notera que le vocabulaire de la métrique distingue bien les phénomènes intérieurs aux mots des phénomènes de frontière : on parle de diérèse / synérèse en position intérieure, de hiatus / synalèphe en position de frontière de mot. Et c'est bien là ce qui différencie les deux phénomènes : l'absence ou la présence d'une frontière de mot.

IV.1.1. L'interprétation classique de la diérèse

Pourquoi dans un groupement vocalique commençant par une glissante (*i, ou, u*), celle-ci devient-elle alternativement une voyelle haute ou une semi-voyelle ? Pourquoi tous les groupements vocaliques n'ont-ils pas le même comportement ? Quelles sont les origines de ces variations ?

Les métriciens définissent usuellement la diérèse et la synérèse par rapport à la langue, ou du moins par rapport à leur propre représentation de la langue. Ainsi :

> Certaines syllabes uniques en prose sont dédoublées dans l'élocution versifiée (...). C'est ce qu'on nomme la DIERESE. A l'inverse deux syllabes en prose peuvent être contractées en une seule dans l'élocution versifiée (...). C'est ce qu'on nomme la SYNERESE.
>
> [Bernardy, 1988, p. 109-110].

Le présupposé des métriciens est qu'en langue la synérèse est devenue quasi-systématique, en tout cas qu'il s'agit d'un phénomène non soumis à variation ; ce que démentent les observations sur la prononciation actuelle. Ce n'est donc pas le privilège du vers.

De nombreux métriciens considèrent que le phénomène de la diérèse est intrinsèquement lié à l'évolution de la langue et que la perspective historique permet de comprendre une grande partie des règles de la diérèse. Nous présentons ci-dessous l'interprétation la plus couramment admise sur les règles de la diérèse.

A l'origine, le décompte des groupements vocaliques se fondait sur des données étymologiques. Les règles ont l'apparence de la simplicité, comme le montre Grammont [*1908*, 1965, p. 16] :

1. « Les deux voyelles françaises constituent une diphtongue qui représente une seule voyelle latine ou une voyelle et une consonne ; elles ne font qu'une syllabe. » Par exemple, *bien*, dérivé de *bene*, compte pour une seule syllabe.
2. « Les deux voyelles françaises représentent deux voyelles latines, qui étaient séparées en latin par une consonne (...) ou y étaient déjà contiguës. » Dans l'une et l'autre condition, elles constituent deux syllabes. C'est le cas pour *li-er*, qui vient de *ligare*, ou pour *di-adème*, dérivé de *diadema*.
3. « Des deux voyelles françaises, l'une est la finale d'un mot simple et l'autre l'initiale d'un suffixe qui s'est adjoint au radical pour en tirer un dérivé en français même. » Ici encore, les deux voyelles appartiennent à deux syllabes différentes. *Vertu-eux* est ainsi construit sur *virtus, -tutis*

La représentation que les métriciens se font de l'évolution de l'alternance diérèse / synérèse est la suivante. La diction poétique, à l'origine fondée sur une prononciation conforme à l'étymologie, était pratiquée dans la langue courante. Mais l'écart se serait peu à peu creusé entre la langue courante et la langue poétique. Ainsi, dans la prononciation ordinaire, les voyelles /i/, /ç/ et /u/ (voyelles hautes) suivies de voyelles ont tendance à se transformer en semi-consonnes, tandis que la langue poétique reste davantage fidèle à la tradition. La perception et la reconnaissance de cet écart ont conduit les métriciens de l'âge classique à énoncer et imposer des règles, puisque l'usage courant ne pouvait servir de référence. Les métriciens auraient en quelque sorte figé le vers par des règles réactionnaires.

Cependant, même dans le vers, le respect des règles étymologiques est peu à peu altéré par les phénomènes liés à l'évolution de la langue. L'évolution linguistique a certes affecté le vers mais d'une manière plus lente et moins systématique. On observe une certaine tendance à la synérèse et l'introduction d'une règle qui force à semi-vocaliser la glissante *i* après une séquence obstruante-liquide (règle des trois consonnes).

La tendance à la synérèse fait que certains termes, qui devraient être sujets à la diérèse pour des raisons étymologiques, ne le sont pas dans le vers : c'est le cas de *dieux, diable, oui* (les exemples sont finalement assez rares) et de certaines formes conjuguées. Mazaleyrat [1974, p. 46-53] propose deux éléments d'explication de la tendance à la synérèse qui entraîne un non respect de la règle étymologique :

1. L'*influence de l'usage* : la fréquence de l'emploi entraîne un infléchissement des règles. Ainsi, la synérèse s'est imposée pour les termes très employés comme *oui*, *diable*... alors que pour les termes plus savants, la diérèse a été mieux respectée. L'évolution serait comparable à une usure due à l'usage réel (fréquence d'emploi) plutôt qu'à un changement historique : c'est bien d'*usage* qu'il s'agit. Malgré tout, le nombre d'exceptions à la règle étymologique est assez limité.
2. Les entraînements de l'analogie : les groupes phonétiquement proches ont tendance a être comptés de la même manière, même s'ils sont étymologiquement distincts. Et, là encore, la tendance à la synérèse

prévaut sur la diérèse. Ainsi, toutes les formes verbales finissant en *-ions* ont été assez tôt comptées comme une seule position (sauf celles qui correspondent à des verbes en *-ier* : *justifi-ons*). Pourtant, l'imparfait et le conditionnel, dérivés de *-ebamus*, auraient dû donner lieu à deux positions. C'est par analogie avec les subjonctifs dérivés de *-eamus*, qui dès qu'ils sont apparus en français ont été comptés en synérèse, que la diérèse a disparu pour l'imparfait et le conditionnel.

La règle dite « des trois consonnes » proscrit la succession de trois consonnes et force donc à vocaliser le *i* s'il est précédé de deux consonnes, la dernière étant une liquide (*l* ou *r*). Dans la phonologie actuelle, il s'agit de la règle *OLG (interdiction d'une succession Obstruante-Liquide-Glissante) décrite par Klein [1991] qui force à vocaliser la glissante. Par exemple, *plier* se prononce /pli e/ et non /plje/. Ce phénomène de langue, apparu entre le XVIe et le XVIIe siècles, s'étend au vers. Par exemple, *grief* n'a aucune raison étymologique de porter une diérèse, puisque la diphtongue représente une seule voyelle latine *a* (*gravis)*.

> Non qu'il ne me soit grief que la tombe possede
>
> Malherbe, *Poésies*, « XC, Consolation à Monsieur du Perier, gentil-homme d'Aix en Provence, sur la mort de sa fille ».

Pourtant après Malherbe, il est toujours décomposé en deux positions dans le vers.

> Gri-efs et faits nouveaux, baux et procès-verbaux.
>
> Racine, *Les Plaideurs*, vers 225.

Il semble que Corneille soit à l'origine de cette évolution :

> Notre poésie a cette obligation, avec plusieurs autres, à M. Corneille, qui, dans sa tragi-comédie du *Cid*, a osé le premier faire *meurtrier* de trois syllabes. Je suis un des premiers qui ait imité en cela M. Corneille, ayant remarqué que les dames et les cavaliers s'arrêtaient comme à un mauvais pas, à ces mots de *meurtrier*, *sanglier*, *bouclier*, *peuplier*, et qu'ils avaient peine à les prononcer.
>
> Ménage, *Observations sur les poésies de Malherbe*, cité par Bernardy [1988, p. 110].

Cette règle a introduit à l'âge classique de nouveaux cas de diérèse. Elle pousse à reconnaître deux positions là où peu de temps auparavant on n'en comptait qu'une. Ainsi compte-t-on :

də /vri /ɔ̃	au lieu de	də /vrjɔ̃
sɑ̃ /gli/e	au lieu de	sɑ̃ /glje.

Alors que le respect de l'étymologie creuse l'écart entre vers et prose, la tendance à la synérèse, comme la règle « des trois consonnes », ont tendance à rapprocher la diction du vers de celle de la prose.

En résumé, on considère que le décompte des groupements de voyelles à l'époque classique résulte de l'interaction de forces contradictoires : le respect de l'étymologie, la tendance à la synérèse et la règle des trois consonnes, qui introduit de nouveaux cas de diérèse.

Cette interprétation historique est séduisante, mais nous n'avons pas les moyens de la vérifier. Y a-t-il réellement eu un affaiblissement de la diérèse ? Cela mériterait un examen approfondi. Une lecture superficielle du *Roman d'Alexandre*, datant du XIIe siècle, semble indiquer qu'il y a à peu près le même nombre de diérèses à cette époque que chez Corneille et Racine ; mais ce ne sont pas les mêmes : certaines ont disparu, tandis que de nouvelles ont été introduites.

Il faudrait aussi examiner le rôle qu'a pu jouer l'évolution de l'orthographe. L'orthographe du français a été considérablement enrichie du Moyen-Age à la Renaissance [Cerquiglini, 1995], et cet enrichissement s'est en particulier traduit par l'introduction de nouvelles voyelles hautes. Ainsi *sanglier* s'écrit au Moyen-Age *sangler*. Or l'introduction dans l'orthographe de voyelles hautes multiplie les possibilités de diérèse. C'est parce que *sangler* devient *sanglier*, qu'ultérieurement on pourra y dénombrer trois syllabes. L'évolution de l'orthographe n'a-t-elle pas eu une incidence inverse à celle de l'usage, en introduisant des nouveaux groupements vocaliques susceptibles de donner lieu à des phénomènes de diérèse ?

La question du tréma mérite aussi quelque attention. Dans les textes du Moyen-Age mais aussi dans la poésie espagnole jusqu'au XXe siècle, et sans doute ailleurs, le tréma semble marquer la diérèse dans le vers comme un élément différenciateur par rapport à la prose.

> Car l'air estut müet, le firmament croissir
>
> Le Roman d'Alexandre, vers 24, p. 70.

> Lento lo embiste, y con süave estilo
>
> Luis de Góngora, *Soledades*, « Primera soledad », vers 40.

Comme nous n'avons pas les moyens d'étudier systématiquement l'évolution du couple diérèse/synérèse en nous appuyant sur des données étymologiques, nous nous contenterons de dresser un portrait synchronique

de la diérèse à l'époque classique telle qu'elle est traitée par Corneille et Racine. Nous verrons comment des textes supplémentaires enrichissent ce portrait. Ce portrait sera constitué d'un ensemble de *règles* faisant intervenir des données lexicales et morphosyntaxiques, et de *listes d'exceptions*.

IV.1.2. Examen empirique de la diérèse chez Corneille et Racine

Ferdinand de Gramont, dans *Les Vers français et leur prosodie* [1876], a constitué un tableau des regroupements de voyelles et de la manière dont ils sont comptés. Dans la lignée de ce travail, nous avons procédé à un examen systématique et exhaustif des suites de voyelles contiguës (commençant par *i*, *u*, ou *ou*) dans les œuvres théâtrales complètes de Corneille et Racine (corpus CORRAC). Nous avons utilisé une des fonctions du logiciel Hyperbase conçu par Etienne Brunet [1989, 1993], le « concordancier », qui restitue le texte à droite et à gauche de la chaîne de caractères recherchée. Les vers sont présentés triés en fonction de l'environnement droit ou gauche de la séquence recherchée, comme le montre la Figure 18. Malheureusement, quand la séquence est au début ou à la fin du vers, le vers est parfois incomplet, ce qui peut être préjudiciable pour l'évaluation du décompte. Si tout le corpus était constitué exclusivement d'alexandrins, il suffirait de l'hémistiche pour évaluer la présence ou non de la diérèse. Mais, comme quasiment tous les mètres sont attestés, le vers complet est nécessaire pour la reconnaissance préalable du mètre, puis pour l'évaluation de la diérèse. Dans les cas d'incertitude, les liens hypertextuels nous ont permis d'aller directement consulter le texte original.

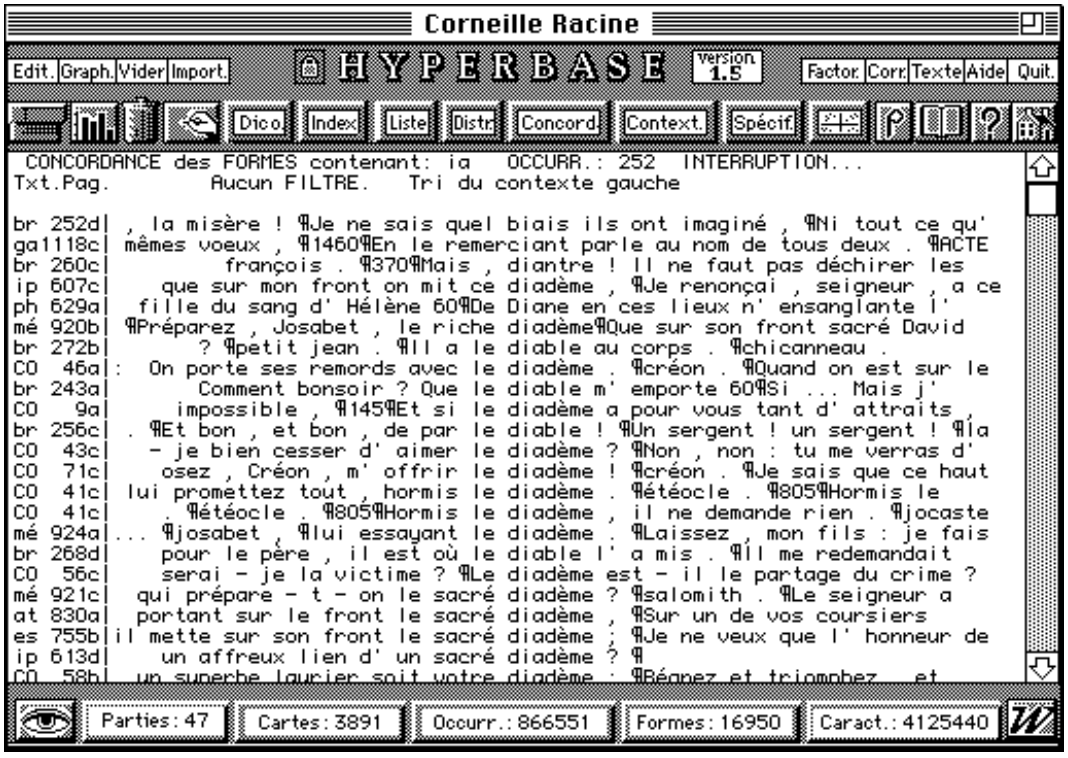


Figure 18. Recherche de vers contenant le groupement « ia » (extrait)

Dans la majorité des cas, un hémistiche suffit pour trancher la question de la diérèse, s'il ne contient qu'un mot à diérèse potentielle, c'est-à-dire une seule séquence glissante-voyelle. La séquence concernée sera comptée comme une ou deux positions de manière à en trouver six. En revanche, s'il existe plusieurs syllabes candidates à la diérèse ou qu'il y a des mots avec un *h* initial dont la nature (aspiré ou non) n'est pas connue, il faudra avoir recours à d'autres vers comportant le même terme pour résoudre l'équation. D'où l'intérêt que les vers soient triés.

Prenons par exemple le vers suivant :

> De Diane en ces lieux n'ensanglante l'autel.
>
> Racine, *Iphigénie*, vers 60.

L'hésitation qu'il peut y avoir pour faire porter la diérèse sur *Diane* ou *lieux* est levée par l'existence d'autres vers qui contiennent l'un ou l'autre terme :

> Jusque sur le bûcher Di-ane est descendue,
>
> Racine, *Iphigénie*, vers 1786.

Cet examen nous a amené à constituer un ensemble de règles et des listes d'exceptions. Le système de description synchronique de la diérèse/synérèse est ainsi constitué de règles qui intègrent des données

phonétiques, lexicales et syntaxiques, et d'exceptions à chacune de ces règles. Il correspond aux règles de décompte des groupements de voyelles dans les œuvres de Corneille et Racine, ce qui ne constitue finalement qu'un faible échantillon de vers, au regard de l'immense production alexandrine. Il y a néanmoins de fortes probabilités pour que de nombreux écrivains de cette époque, voire d'époques ultérieures, aient suivi les mêmes règles. Nous verrons que l'introduction de nouveaux textes a permis d'introduire de nouvelles règles, d'allonger les listes d'exceptions, mais n'a pas invalidé les règles déjà trouvées, ce qui est un résultat assez satisfaisant.

IV.1.3. Tableaux récapitulatifs des règles de la diérèse

Les tableaux suivants présentent le traitement dans les vers de Corneille et Racine de groupements vocaliques commençant par les voyelles hautes *i*, *u* et *ou*. Les lettres *i*, *u* et *ou* ont respectivement les réalisations phonétiques /j/, /ɥ/ ou /w/ s'il n'y a pas de diérèse, et /i/, /ç/ ou /u/ s'il y en a une. On a rajouté en italique d'autres cas trouvés dans les œuvres qui constituent le complément de notre corpus : Hugo, Baudelaire, Mallarmé, Rimbaud... (*cf.* Annexe VI).

Les exemples de diérèse qui relèvent de la règle des trois consonnes présentée précédemment n'ont pas été traités dans le tableau.

	Diérèse (deux voyelles métriques)		Synérèse (une voyelle métrique)		Mots variables
Gp voy.	**Catégorie grammaticale**	**Exemples**	**Catégorie grammaticale**	**Exemples**	
ia (iâ)	imparfait et passé simple des verbes en *-ier*[79]	étudi-ai, envi-ait...			
	noms propres	Eli-acim, *Vi-a, Panagi-a, Di-az, Geni-alis...*		*Mathias (Hugo), Brescia (Hugo)*	
	noms, adjectifs	li-aison, bi-ais[80], insati-able, di-adème, mari-age, invari-able... *li-ard (Hugo), milli-ard (Rimbaud)* *di-aloguer, osti-aire (Hugo), immédi-atement (Hugo), opini-âtrement (Baudelaire),*		diable, viande, diantre *endiablé, irrémédiable*[81] *(Baudelaire),, diacre (Hugo), piastre (Hugo), jusquiame (Hugo)*[82]	*miasme*
	verbes rire et sourire	*souri-ant, souri-ait, ri-ant*			
Cya ∅ïa		*hy-acinthe, Asty-age, Cy-admis, Marsy-as, dry-ade, ï-ambe*			
iec		*Babie-ça*			
ied		*Ovi-edo*		pied, *Winkelried*	
			verbes en *-seoir*	sied, messied, assied	
ief				*relief, fief*	grief
iel		*essenti-el, torrenti-el, pestilenti-el, Dani-el*		ciel, fiel, Penafiel, *Spielberg, immatériel*	
iem		*pati-emment*			
ien			formes conjuguées des verbes tenir et venir et leurs dérivés	tient, parvient, viens, tienne...	
			pronoms et adjectifs possessifs	mien, tien, sien, mienne, tienne...	
			bien, rien et dérivés	bienheureux, bien, rien...	

[79] Exemples de verbes en *-ier* : *remercier, étudier, fier, sacrifier, lier, expier, marier, envier.*

[80] Pour F. de Gramont, *biais* n'a qu'une voyelle métrique. On trouve pourtant chez Racine (unique occurrence) : « Je ne sais quel bi-ais ils ont imaginé" (*Les Plaideurs*)

[81] Dans un unique exemple, *irrémédiable* ne porte pas de diérèse. Ceci repose sur une fausse étymologie, le terme n'est pas construit sur *diable*, mais sur *remédier*.

[82] Autres exceptions repérées par F. de Gramont : *fiacre, galimatias.*

	Diérèse (deux voyelles métriques)		Synérèse (une voyelle métrique)		Mots variables
Gp voy.	**Catégorie grammaticale**	**Exemples**	**Catégorie grammaticale**	**Exemples**	
	noms, adjectifs en *-ien*	impati-ence, impati-ent, audi-ence, li-en, chirurgi-en, *bohémi-en, tyroli-en, magici-en,...* patrici-enne, *musici-enne*		entretien, chien, Praticien, chrétien[83], chrétienne... *vaurien (Baudelaire)*	ancien, gardien
ien, yen	noms propres	Malli-en, Ori-ent, Argi-en, *Phrygi-en, Apuli-ens...* *Liby-ens, Fabi-ens...*		Amiens, *Enguien, Damiens (Hugo), Etienne, Vienne, Sienne*	
ier, yer	verbes en *-ier*	envi-er, mari-er, fi-er...	noms, adverbes, adjectifs	fier, fierté, cellier, *vierge* *premier, dernier...* *pierre*	
			noms propres	*Janvier, Montpellier, Montdidier, Ogier, Pierre, Renyer*	hier lyerre
ies		hardi-esse[84], *li-esse*		*Fiesque (Hugo)*	sieste
iet		inqui-et	toujours	serviette, assiette, *émiette*[85]...	
ieu	adjectifs en *-ieux*(se) ; noms qui se finissent par un x au singulier ; adverbes construits sur les adjectifs en *-ieux*	audaci-eux, religi-eux, *radi-euse, les ambiti-eux...* compendi-eusement (Racine), *silenci-eusement (Baudelaire, Rimbaud), victori-eusement et délici-eusement (Mallarmé)*	noms	vieux (seul adjectif : féminin irrégulier) lieu(x), dieu(x), *épieu(x)*, adieu, parbieu, cieux, mieux... Dieu, Mathieu	
	adjectifs en *-ieur*	extéri-eur, *intéri-eur, antéri-eur, mendi-eur*		monsieur, sieur, plusieurs	
	adjectifs construits sur verbes *-rire*	*ri-eur, ri-euse*			
iez	verbes en *-ier*	remerci-ez, fi-ez...	tous les verbes à la deuxième personne du pluriel, sauf verbes en *-ier* et *rire*	pourriez, prétendiez...	
	verbe *rire*	ri-ez			
ié	verbes en *-ier*	envi-é, li-é	noms finissant en *-tié*	amitié, moitié, pitié	

[83] Elwert [1961] rappelle que la synérèse s'est imposée au XVIe siècle pour *chrétien*, au XVIIe pour *ancien*.
[84] F. de Gramont a aussi trouvé *li-esse, acqui-esce.*
[85] Cependant Mazaleyrat compte *mi-ette*.

	Diérèse (deux voyelles métriques)		Synérèse (une voyelle métrique)		Mots variables
Gp voy.	**Catégorie grammaticale**	**Exemples**	**Catégorie grammaticale**	**Exemples**	
		pi-été, alli-é, ali-éner		assiégeant, piédestal, fiévreux, piéton, tiédeur (Hugo), maniéré (Baudelaire)	
iè		inqui-ète, *espi-ègle(Hugo*)	noms	piège, siècle, carrière, première, dernière, tiède, genièvre, don Diègue, *Molière*, *siège*...	
	verbes en -*ier*	li-èrent,			
io		*pati-o, scenari-o, (Hugo)*			
iob		*Ni-obé*			
ioc		*médi-ocre, di-océsain (Hugo)*		*pioche (Hugo)*	
iod		*péri-ode*			
iog		*héli-ogabales (Hugo)*			
iol		*vi-ol, vi-oler, invi-olable, vi-olence, vi-olenter, vi-olete(Malherbe), vi-olet (Boileau), vi-olons (instrument, Hugo), vi-ole (instrument, Mallarmé), bari-oler (Hugo, Mallarmé), éti-oler (Mallarmé), glori-ole (Mallarmé)*			*viol, invi-olé fiole, fi-ole*
iom		*idi-ome (Hugo)*			
ion	1ère personne pluriel, verbes en -*ier*	sacrifi-ons	imparfait 1ère personne du pluriel sauf verbes en -ier	passions, prenions	devisi-ons
	noms propres	Anti-ochus, I-onie, Hermi-one, Ili-on..., *Ethi-opie, Liri-ope, Héli-opolis, Ari-oste, Mégari-os, Ri-os (Rimbaud), Ori-on, Alli-oth, Ni-obé(Hugo), Bali-ol*		*Biot (Hugo)*	
	noms communs en -*ion*	passi-on, divisi-on, li-on, milli-on... *mari-onnette, visi-onnaire, révoluti-onnaire (Hugo)...*		*scorpions (Baudelaire)*	
iop		di-optrique (Rimbaud)			
ios		curi-osité, *di-oscure (Hugo), lysi-os (Hugo), pluvi-ose (Hugo)*			
iot		*lori-ot, chari-ot, idi-ot (Hugo), héli-otropes (Rimbaud), Alli-oth*			
iou		chi-ourmes (Hugo)		Cadédiou (Corneille), *pioupiou (Rimbaud*)	

	Diérèse (deux voyelles métriques)		Synérèse (une voyelle métrique)		Mots variables
Gp voy.	**Catégorie grammaticale**	**Exemples**	**Catégorie grammaticale**	**Exemples**	
Cyo ∅yo	Noms propres et noms communs	Y-olande (Racine) *Ly-on, alcy-on, my-osotis, my-ope*		*yoles (Rimbaud)*	
iu	noms propres	Mari-us, Héracli-us		*Fiume (Hugo)*	
iu	noms communs	*opi-um pandémoni-um, médi-um reli-ure*			
ua	verbes en *-uer*[86] participe présent, imparfait, passé simple	attribu-ait, vou-ai, remu-ant, tu-a, salu-ait...		*puant (Rimbaud)*	
	noms, adjectifs, verbes	*nu-age, nu-ance, su-ave, sanctu-aire...* *immu-able,* *persu-ader, dissu-ader*			jaguar
	noms propres	*Bellu-a, Ju-ana (Rimbaud)*		*Juan, Guadalquivir, Aguas, Aranjuez, Huesca*	
ue	verbes en *-uer* infinitif, 2ème personne du pluriel	continu-er, diminu-ez, perpétu-er...			
		mutu-el, ponctu-ellement, spritu-el Terru-el, tru-elle influ-ence mu-et, mu-ette su-eur, lu-eur, respectu-eux, tumultu-eux		écuelle (Hugo)	duel
ué	verbes en *-uer* participe passé (verbe ou adjectif)	diminu-é, tu-é			
		nu-ées, hu-ées, prostitu-ée, pu-éril			
	noms propres	Assu-érus			
ui, uy		*bru-it (verbe), bru-ire (Hugo)*	verbes en *-uire*	conduire, détruire, instruire,	bruire
		ru-ine, ru-inera *bru-ine, casu-iste, intu-iti-on, flu-ide, pu-îné, dru-ide, tru-ie (Hugo)*		fuite, aujourd'hui, ennui, huit, juif, nuit, puis, puisque, bruit, fruit	suicide
			pouvoir, être, suivre	puis, puissance, suis, suivre	
		Lu-is		*Ruy, Tuy*	

[86] Exemples de verbes en *-uer* : *diminuer, exténuer, tuer, dénuer, contribuer...*

	Diérèse (deux voyelles métriques)		Synérèse (une voyelle métrique)		Mots variables
Gp voy.	**Catégorie grammaticale**	**Exemples**	**Catégorie grammaticale**	**Exemples**	
uo	verbes en *-uer* (1ère personne du pluriel)	salu-ons, continu-ons			
		impétu-osité, *monstru-osité, du-o, septu-or*			
oua	verbes en *-ouer*[87]	désavou-ant, avou-ait...			
	dérivés de verbes en *-ouer*	lou-ange, lou-able			
		douai-re, cou-ard, rou-age, tatou-age, Chou-an (Hugo) échou-age (Rimbaud)		*bivouac, Souabe (Hugo), fouaillez (Hugo)couacs (Rimbaud), gouailleurs (Mallarmé)*	
oue	verbes en -*oue*r	secou-er	Verbe fouetter	fouette	
		jou-et		fouet, girouette	
oué	verbes en *-ouer* : participe passé	jou-é, secou-é, nou-ée			
	noms propres	Mérou-é-é (Mérové dans d'autres éditions)			
oui	verbes *-ouir*[88] et *-ouïr*	ou-is, *évanou-ir, épanou-issait*			
		ou-ïe, lou-is, Lou-is		oui, *fouine, Baudouin (Hugo)*	

[87] Exemples de verbes en -ouer : *secouer, nouer, jouer, louer, avouer, désavouer*...

[88] Exemples de verbes en -ouir : *jouir, éblouir, s'évanouir...*

Ce tableau donne les règles de la diérèse en fonction des groupements vocaliques. Certaines règles peuvent être énoncées de manière plus générale :

- Règle des trois consonnes : il y a toujours diérèse dans les contextes obstruante-liquide-*i*-voyelle : *voudri-ez*, *pri-er*...
- Règle des infinitifs : on fait toujours la diérèse pour les verbes ayant un infinitif en -*ier*, -*uer* et -*ouer* et pour les verbes construits sur *rire*.
- Pour tous les autres verbes, on ne fait pas la diérèse sur les terminaisons de conjugaisons en -*ions* ou -*iez*.

Marc Klein [1993, p. 106, 107] considère à propos du français parlé d'aujourd'hui que la tendance à la synérèse varie selon la glissante :

> Toutes choses par ailleurs égales, la synérèse est plus probable avec la glissante la moins sonore /i/ qu'avec /u/, et moins probable avec la plus sonore /y/.
>
> [Klein, 1993, p. 106].

Parmi les facteurs qui peuvent avoir une incidence sur l'alternance diérèse/synérèse, il identifie :

> - le nombre de syllabes lexicales : la disyllabicité renforce le hiatus (diérèse), la polysyllabicité renforce la synérèse ;
> - la cohésion morphologique : le rendement de la synérèse est fort à l'intramorphémique HV[89], à la joncture de suffixes et de certains préfixes H+V, faible à la joncture d'autres préfixes H=V, nul à la joncture de mot H#V ;
> - la variation dialectale : l'échelle /I/</U/</Y/ est beaucoup mieux marquée, donc plus observable, dans les parlers septentrionaux (notamment en région parisienne) que méridionaux.
>
> [Klein, 1993, p. 106].

Nous avons cherché à voir s'il existait effectivement une hiérarchie des voyelles hautes au regard de la synérèse. Nous n'avons pas pu le faire « toutes choses égales par ailleurs », c'est-à-dire en contrôlant les facteurs externes, mais nous présentons quand même les différences de traitement de la diérèse selon la voyelle haute. Nous avons extrait de l'ensemble de nos corpus les mots contenant une voyelle haute suivie d'une voyelle et examiné leur comportement au regard de la diérèse / synérèse pour chacune des voyelles hautes. Que ce soit

[89] H : glissante ; V : voyelle.

en nombre de mots différents ou en nombre d'occurrences, la voyelle haute I est la plus fréquemment attestée dans les groupements vocaliques. Dans 74 % des mots et 64 % des occurrences contenant une séquence glissante-voyelle, cette glissante est /I/.

Vocables	**I (74 %)**	**U (8 %)**	**Y (18 %)**	**Total**
taux de synérèse (GV)	49 %	10 %	62 %	48 %
taux de diérèse (HV)	51 %	90 %	38 %	52 %

Clef de lecture : Dans 49 % des mots, la séquence IV se réalise en GV (semi-vocalisation de la voyelle haute, synérèse), dans 51 % des mots elle se réalise en HV (diérèse).

Occurrences	**I (64 %)**	**U (3 %)**	**Y (33 %)**	**Total**
taux de synérèse (GV)	77 %	32 %	87 %	78 %
taux de diérèse (HV)	23 %	68 %	13 %	22 %

Clef de lecture : Dans 77 % des occurrences, la séquence IV se réalise en GV (semi-vocalisation de la voyelle haute, synérèse), dans 23 % des mots elle se réalise en HV (diérèse).

Figure 19. Tendance à la synérèse selon les glissantes

Au vu de ces simples tableaux, il ne semble pas que l'échelle de sonorité /I/< /U/ < /Y/ soit valable pour le vers classique : la synérèse est moins probable dans les séquences UV que dans les séquences IV, et moins probable dans IV que dans YV. On a plutôt une hiérarchie du type /Y/</I/</U/. Il faudrait pour mener un examen plus fin, tenir compte de la longueur des mots et de la position de la glissante par rapport aux frontières de morphèmes.

IV.1.4. Formalisation

Il apparaît clairement, au vu des tableaux récapitulatifs, que dans un but d'automatisation il est imprudent de s'appuyer uniquement sur les groupes de voyelles pour déterminer le nombre de positions métriques. On ne peut non plus s'appuyer sur l'unité lexicale : de nombreux homonymes se distinguent justement par leur décompte dans le vers. Seule une analyse morpho-syntaxique permet de lever l'ambiguïté. Ainsi, les verbes en *-ions* (première personne du pluriel) ne comportent-ils pas de diérèse, à l'inverse des noms finissant en -

ions. L'analyse syntaxique permettra de distinguer *passi-ons* (substantif) de *passions* (première personne du pluriel), et *fier* (adjectif) de *fi-er* (verbe).

Pour parvenir à une bonne modélisation des règles de la diérèse, il faut disposer à la fois d'un vaste lexique, où les homonymes sont distingués, et d'un analyseur morpho-syntaxique de qualité.

La prise en compte informatique la plus naturelle de la diérèse se fait au niveau lexical en ajoutant le trait « diérèse » ou « non-diérèse » aux formes du lexique. Ce marquage est produit soit à l'aide de règles ad hoc au cours du traitement, soit directement dans le lexique (pour les exceptions aux règles).

Un exemple de règle : tous les verbes en *-ier* (infinitif, formes conjuguées, participe présent, participe passé et adjectifs dérivés) se voient affectés du trait "diérèse". Le verbe *privilégier* et ses dérivés constituent les seules exceptions à cette règle : on lui ajoutera donc le trait "non-diérèse" dans le lexique. Autre règle : tous les noms propres ont le trait "diérèse". Les rares exceptions (*Amiens*, *Penafiel*...) seront là encore prises en compte dans le lexique.

La différence entre les dérivations phonologiques appliquées au vers et celles appliquées en général, est sous cet aspect de la diérèse facile à opérationnaliser : il suffit d'ajouter un méta-contexte à l'application de la règle de semi-vocalisation des voyelles hautes (/i/ /ç/ /u/), qui spécifie l'inhibition de la règle pour toute voyelle haute d'un mot portant le trait « diérèse[90] ».

IV.2. LA QUESTION DU *E* MUET

Si la lettre *e* est la plus fréquente en français, le phonème que l'on désigne usuellement par l'expression *e* « muet » est en même temps le plus instable. Sa réalisation varie selon sa position dans l'énoncé (ou plus précisément dans le mot phonologique) et selon le contexte de l'énonciation (origine géographique

[90] Cette solution a le mérite d'être simple, mais le travers d'être probablement incorrecte. Il existe sans doute des mots pour lesquels les différentes voyelles hautes n'ont pas le même statut vis-à-vis de la diérèse. Comment compter les voyelles de *intuition*, ou de *expiation* ? Nous n'avons néanmoins pas rencontré de tels problèmes chez Corneille et Racine. Ces mots apparaissent pour la première fois chez Hugo et dans les deux cas, les deux voyelles hautes du mot ont le même comportement.

et sociale du locuteur, situation pragmatique de l'énonciation, caractéristiques des allocutaires...). Selon les cas, le *e* muet se réalise en /ə/, /ø/, /œ/ ou ne se réalise pas phoniquement. D'autres termes sont utilisés pour le désigner : *e caduc*, *e instable*, *e à éclipses*, *schwa*... Nous conserverons l'expression de *e* muet pour désigner l'ensemble de ces *e* graphiques qui, selon les cas, se réalisent ou ne se réalisent pas. Le terme est très mal approprié, puisque dans bien des cas, ces *e muets* ne sont pas muets. On aurait pu tout simplement employer le terme *e*. Ainsi, pour nous, la catégorie de *e* muet englobe les *e* des monosyllabes (*j**e**, t**e**, qu**e**...)* qui sont graphiquement élidés devant voyelle, les *e* de fins de mots (*diadèm**e**, vi**e**...*) mais aussi les *e* à l'intérieur des mots (*r**e**ssentiment*).

Dans le vers comme dans la langue, le *e* muet est instable. Mais le vers, parce qu'il est nombré, facilite la mise en règle des variations. Le fait que le *e* intervienne ou non dans le décompte métrique permet alors d'établir les règles de sa prononciation. Nous trouvons ici un intérêt de l'étude métrique pour la linguistique générale : l'existence de la contrainte de dénombrement va nous permettre de trancher certains cas, ce que le linguiste, confronté à des réalisations phonétiques variables, fait plus difficilement. Nous avons trouvé cet avantage méthodologique de la contrainte métrique dans l'étude des équivalences phonétiques (*cf.* chapitre III sur la rime).

Dans le vers, il existe pour le *e* muet à un niveau abstrait deux états : soit il est compté métriquement (il occupe alors le cœur d'une position métrique), soit il ne l'est pas. Cette variation est soumise à des règles précises.

Lorsque l'on passe de la structure abstraite du mètre à sa structure réalisée, du modèle du vers à la diction, cette dichotomie de base peut être nuancée en introduisant, comme le proposaient Milner et Regnault [1987], des niveaux de prononciation intermédiaires ; il n'en reste pas moins que le décompte est premier pour l'analyse du *e* muet.

Au-delà du décompte, le *e* muet joue un rôle crucial dans le vers. Tout d'abord, il marque le genre des rimes (rime dite « féminine »). Ensuite, la manière dont il occupe ou non les positions autour de la césure permet de qualifier la physionomie globale du vers et de distinguer les vers selon les types de césure admis (césure classique, épique, lyrique ou italienne).

Après avoir exposé la théorie de Milner et Regnault, telle qu'elle figure dans *Dire le vers* [1987], nous montrerons comment nous l'avons adaptée dans le cadre du métromètre.

IV.2.1. Le décompte du *e* muet dans le vers

Dès 1974, Milner propose, dans « Réflexions sur le fonctionnement du vers français », un modèle de fonctionnement du *e* muet et de la liaison dans le vers (ce texte a été repris et remanié dans *Ordres et raisons de langue* [Milner, 1982]). Ce modèle est repris par Milner et Regnault dans *Dire le vers* [1987, p. 31-45], avec une terminologie affinée. Ces auteurs proposent un cadre d'analyse à deux étages très séduisant. Ils partent de l'hypothèse que vers et langue fonctionnent selon les mêmes principes (phonologiques). Si la langue et le vers diffèrent quant au maintien du *e* muet, c'est que les unités sur lesquelles les règles s'appliquent sont différentes. Dans la langue, l'unité en question est le mot phonologique. Mais qu'est-ce qu'un mot phonologique ?

> Le mot phonologique en français est constitué par un groupe syntaxique majeur (Groupe nominal, Groupe verbal, Groupe adjectif, Groupe prépositionnel).
>
> [Milner et Regnault, 1987, p. 21].

Les auteurs concluent :

> En général est mot phonologique ce qui constitue *à la fois* une unité de souffle, une unité grammaticale et une unité de signification.
>
> [Milner et Regnault, 1987, p. 29].

Cette notion de mot phonologique correspond à l'unité classiquement retenue pour les questions d'accentuation, c'est le groupe de souffle ou ce que Dell désigne par « tronçon » [1984, p. 68]. Comme le rappelle Vaissière [1991, p. 108], le découpage en groupes est loin d'être évident. Dans la pratique, l'identification des mots phonologiques de langue sera le principal point d'achoppement du métromètre.

Pour ce qui est du traitement du *e* muet (et de la liaison, nous le verrons) **dans le vers**, Milner et Regnault posent le « Principe constituant du vers » : le vers classique doit être considéré comme un unique mot phonologique. Les règles qui s'appliquent en langue aux mots phonologiques s'appliqueront au vers entier.

Arrêtons-nous un instant sur le « Principe constituant du vers ». Est-ce le vers tout entier qui doit être considéré comme un seul mot phonologique ? Revenant sur la théorie de Milner, Lusson et Roubaud [1974] considèrent que chaque hémistiche doit aussi être traité comme un mot phonologique.

> Nous dirons que l'alexandrin est un *segment métrique* (MV) à 12 positions, concaténation de deux segments métriques à 6 positions chacun.
>
> [Lusson et Roubaud, 1974, p. 43].

Dans le vers classique, les deux options peuvent être indifféremment retenues pour le traitement du *e* muet. C'est justement la particularité du vers classique, comme l'a montré Roubaud [1977], d'être à la fois un et deux segments métriques. Par son traitement du *e* muet, le vers admet simultanément les deux interprétations : un ou deux mots phonologiques. En revanche, si nous avions à analyser le rôle du *e* muet dans l'alexandrin épique, il faudrait absolument considérer qu'il est constitué de deux mots phonologiques soumis aux mêmes règles.

Dans le cadre d'une mécanisation des règles, il était plus facile de choisir la première option (vers = mot phonologique). La seconde (hémistiche = mot phonologique) implique de repérer préalablement la frontière entre les deux hémistiches, ce qui est loin d'être simple.

Dorénavant, bien que l'essentiel de la théorie figure dans l'article de Milner de 1974, nous reprendrons les formulations de Milner dans *Ordres et raisons de langue* [1982] et de Milner et Regnault dans *Dire le vers* [1987].

Milner pose comme base de fonctionnement général du *e* muet en langue les règles suivantes[91] :

---

[91] Il est clair que la question du *e* muet, probablement une des questions les plus complexes et controversées de la phonologie du Français, ne saurait se résumer à ces trois principes simples. Pour le vers néanmoins, ils constituent un bon cadre de départ pour la formalisation qu'on va développer. On se reportera à Dell [1985] pour une discussion plus complète et à Encrevé [1988, p. 212-232].

(1) *Chute du e final* :
e tombe à la fin du mot phonologique, quelque soit la nature de l'élément suivant.
(2) *Elision :*
e muet tombe à l'intérieur du mot phonologique s'il est suivi d'une voyelle.
(3) *Loi des trois consonnes* (optionnelle) :
e tombe à l'intérieur du mot phonologique s'il est précédé et suivi de consonnes et si sa chute ne produit pas une suite de trois consonnes ou plus.

[Milner, 1982, p. 287-288].

Seules les lois (1) et (2) sont obligatoires et elles seules s'appliquent au vers comme à la langue. Par application du « Principe constituant », les deux règles suivantes permettent alors de rendre compte du traitement du *e* muet dans le vers :

(a) à l'intérieur du vers, aucun *e* muet ne tombe, sauf par élision[92] devant une voyelle ou un *h* non aspiré ;

(b) à la fin du vers, le *e* muet tombe obligatoirement[93].

Ainsi, d'après ces deux règles (a et b), seuls deux *e* muets comptent dans le vers suivant qui en comporte cinq :

Refuse cette joie, et s'en trouve accablée

Corneille, *Le Cid*, vers 54.

Le premier et le second comptent tous les deux puisqu'ils sont suivis d'une consonne (règle (a)), le troisième et le quatrième ne comptent pas (règle (a)), enfin le dernier est élidé puisqu'il est en fin de vers (règle (b)).

---

[92] Contrairement aux apparences, l'application de cette règle ne va pas de soi, puisqu'il faut être capable de déterminer quel est le phonème qui suit le *e* muet, et donc, entre autres, de connaître les lois de la liaison. Nous reviendrons sur ce point.

[93] C'est par commodité que nous avons pris cette option, parce que le *e* muet en fin de vers n'intervient pas dans la reconnaissance du mètre alexandrin. Mais ce *e* muet joue un rôle pour la rime, pour l'alternance en genre et pour la structure des strophes. Nous avons d'ailleurs dû le réintégrer pour l'analyse de la rime, et nous tenterons dans une prochaine version du métromètre de garder la trace de ces *e* muets élidés. D'ailleurs jusqu'au XVIe siècle, les alexandrins « féminins » étaient considérés comme des vers de 13 syllabes et les « masculins » de 12 (*cf.* les traités du XVe cités par Roubaud [1977, p. 37-38]), ce qui laisse supposer que les *e* muets finaux n'étaient donc pas considérés comme élidés.

Pour être complet, il faut ajouter un ensemble de règles sur le *e* muet après voyelle.

- A l'intérieur du mot (*cf. défiera, tueras*…), il est clair que le *e* muet après voyelle n'a quasiment jamais de rôle métrique :

  Je tuerais ma maîtresse avec mon ennemi.

  Corneille, *L'Illusion comique*, vers 344.

- De même les *e* de la désinence de la troisième personne du pluriel à l'imparfait et au conditionnel (*aient*), ainsi que le *e* de *soient* ne constituent jamais de position métrique.

  Que ce soient vérités, que ce soient impostures,

  Corneille, *Mélite*, vers 976.

  Ah ! que tant de malheurs m'auraient favorisé,

  Corneille, *Clitandre*, vers 651.

Chez Corneille comme chez Racine ce type de terminaison n'apparaît quasiment jamais à la rime. Graphiquement, cette terminaison devrait constituer une rime féminine. Mais puisque dans le vers le *e* muet après voyelle ne compte pas, on en conclut qu'il est vraiment mort ; la terminaison devrait plutôt constituer une rime masculine. Il y a une contradiction évidente entre la graphie et le décompte qui serait insoutenable pour l'alternance en genre des rimes.

Ces deux catégories de *e* muets, qui ne jouent aucun rôle métrique, méritent bien le nom de *e* mort qu'avait introduit Roubaud dans le cadre du programme Dynastie [1986, p. 70].

Sur l'ensemble du corpus, on trouve cependant deux cas, que l'on doit à Corneille, de *e* muets après voyelle qui occupent une position métrique, qui ne sont donc pas morts.

Comme toutes les deux jou-ent leurs personnages !

Corneille, *La suite du Menteur*, vers 1014.

Les sœurs cri-ent miracle, et chacune ravie

Corneille, *Médée*, vers 73.

Dans les vers précédents, *jou-ent* et *cri-ent* constituent deux syllabes métriques.

Ce sont les uniques exemples de séquences V*ent* à l'intérieur du vers, alors que celles-ci sont fréquentes en fin de vers, et ont valeur de rimes féminines. Dans ces terminaisons V*ent*, et il en va de même pour les terminaisons en V*es*, le *e*

n’est pas mort, on l’évite à l’intérieur du vers car sa réalisation métrique est en contradiction avec la prononciation. On trouve même un exemple de terminaison en *oient* à la rime à valeur de rime féminine (les rimes des vers antérieurs *ennemi*/*ami* et des vers postérieurs *chasser*/*embrasser* sont masculines) :

> Ne t’étonne donc plus si je veux qu’ils se voient :
> Je veux qu'en se voyant leurs fureurs se déploient,
>
> Racine, *La Thébaïde*, vers 887-888.

Les *e* muets de ces terminaisons ne sont pas morts.

- Quant aux autres *e* muets après voyelle, ils n’apparaissent dans le vers classique que s’ils sont suivis d’une autre voyelle ou d’un *h* non aspiré, autrement dit, uniquement s’ils peuvent être « fusionnés » avec la voyelle suivante.

> Et vous avez montré, par une heureuse audace,
> (e)(v u)(zɩ a)(v e)(m ɔ̃)(t r e)(p a r)(y n)(œ)(r ø z)(ɔ)(d a s)
>
> Racine, *Andromaque*, vers 149.

A la suite des recommandations de Malherbe, aucun *e* muet après voyelle ne doit apparaître devant une attaque non nulle. Ainsi, le vers suivant (erreur d’édition) est impossible chez Racine :

> *À venir prodiguer sa vie sur un théâtre => sa voix
>
> Racine, *Britannicus*, vers 1475.

Pour Mazaleyrat [1974, p. 64-65], cette mesure permet de lever la contradiction entre le traitement du *e* muet après voyelle à l’intérieur du mot (*tuerais*), qui est systématiquement amuï dans le vers, comme il l’est dans la langue depuis le moyen français, et celui du *e* muet après voyelle en fin de mot qui tantôt constitue le cœur d’une position métrique tantôt non.

Ces variations témoignent de l’importance de cette difficulté. En effet, si devant une attaque non nulle le *e* après voyelle est maintenu, la conformité avec la métrique classique est respectée (maintien des *e* muets devant consonne), mais le décalage est grand avec la prononciation. Ainsi dans l’exemple précédent, on aurait comme positions métriques : (v i) (ə)(s y r), avec un *e* muet assez artificiel dans la prononciation qui en plus crée une situation de hiatus.

Inversement, s'il est élidé, on se rapproche de la prononciation mais les règles métriques ne sont plus respectées. Cornulier [1995, p. 226] reprend cette même idée en écrivant : « Dans les suites de type "VeC", il faut choisir entre une irrégularité métrique (par négligence de FG1[94] ou par excès d'élision métrique) et une irrégularité (ou particularité) grammaticale de nature toute différente. » Autrement dit, soit le *e* muet est élidé et les règles de l'élision métrique ne sont plus respectées, soit il est maintenu, mais son maintien entre en contradiction avec la prononciation. Malherbe résout abruptement cette contradiction en interdisant tous les *e* muets après voyelle devant une attaque non nulle.

Une question est de repérer les *e* muets qui comptent métriquement, une autre est de les dire. Le propos de Milner et Regnault ne se limite pas au repérage des *e* muets comptés, qui peut être réglé par les règles vues précédemment, mais vise la diction des vers. Ils proposent dans la suite de leur analyse de distinguer cinq[95] niveaux hiérarchisés de diction du *e* muet, qui permettent de prendre en considération les contradictions qui peuvent éclater entre les structures syntaxiques et métriques, et de différencier par leur « force » ces différents *e* muets.

Sans entrer dans les détails de leur analyse, disons simplement que, par exemple, les *e* muets qui se maintiennent dans le vers comme dans la langue doivent se réaliser plus fortement que les *e* muets qui peuvent s'élider en langue, et, a fortiori, que ceux qui s'élident systématiquement. Notre but n'est pas de proposer une représentation de la diction, mais il nous a semblé cependant pertinent que le métromètre distingue deux types[96] de *e* muets qui

---

[94] FG1 (Fiction graphique 1) : « Convention 1 de l'« *e* » optionnel (FG1) : toute graphie d'*e* optionnel est censée correspondre à la voyelle correspondante (élidable seulement par Élision métrique) » [Cornulier, 1995, p. 217].

[95] En fait six, mais un de ces niveaux correspond à la chute pure et simple. Nous le passerons donc sous silence.

[96] Ceci pose un problème de transcription, puisqu'il n'existe qu'un symbole dans l'alphabet phonétique international pour représenter le *e* muet : /ə/. Nous avons donc introduit un caractère supplémentaire, noté /$ə_f$/, qui représente la version "affaiblie" du précédent.

correspondent à deux situations phonologiques distinctes et qui ont des rôles métriques différenciés : les *e* muets forts et les *e* muets faibles.

IV.2.2. Formalisation

Concrètement, dans le métromètre, trois règles phonologiques gouvernent la prononciation du *e* muet, qui est toujours présent dans la structure profonde. Les deux premières concernent son élision au contact de voyelles (aussi bien avant qu'après), et sont identiques à celles qui figurent dans la liste des dérivations phonologiques « standard » du français. Encore une fois, le métromètre ne prétend pas rendre compte d'une norme de prononciation des *e* muets mais simplement distinguer les *e* muets qui ont un rôle métrique de ceux qui n'en n'ont pas, autrement dit de ceux qui ne comptent pas. Le terme d'élision, aujourd'hui largement utilisé pour les *e* qui ne comptent pas, ne convient peut-être pas, s'il reste effectivement une trace de la présence de ce *e* muet. Peut-être faudrait-il lui préférer le terme de synalèphe (fusion des deux voyelles, du *e* muet avec la voyelle qui le suit) qui était utilisé dans les traités du XVIe siècle (*cf.* le traité de Fabri cité par [Roubaud, 1977, p. 62]). Voici nos trois règles :

(1') *Le e muet s'élide devant une voyelle.*

(2') *Le e muet s'élide après une voyelle.*

Cette seconde règle, qui permet d'éliminer aisément les cas de *e* mort, est valable pour l'alexandrin de Corneille et Racine, qui respecte la norme la plus stricte, celle qui interdit des séquences Ve#C. Elle échoue uniquement dans le décompte de *crient* et *jouent*, puisqu'elle conduit à une élision du *e* alors que le vers le maintient (*cri-ent*, *jou-ent*). Dans une prochaine version, nous envisagerons d'affiner cette seconde règle.

La troisième règle traite des *e* muets en fin de mot (lexical). Elle a pour but[97] d'élider le dernier *e* muet de la ligne analysée[98], ainsi que « d'affaiblir »

[97] Et c'est bien le seul. En particulier, nous ne suivons pas Milner et Regnault dans leur proposition de faire sonner les rimes féminines qui se terminent par une voyelle haute (/i/, /u/, et /ou/) en transformant le *e* muet final en la semi-voyelle correspondante (*une vie* lu / ç n ə / v i j / à la rime). Pourquoi en effet appliquer un traitement spécifique à ces trois voyelles ?

tous les autres *e* muets terminaux qui tomberaient en langue. C'est cette règle qui permet de distinguer les *e* muets faibles des *e* muets forts.

(3') *Sauf cas particulier (monosyllabe, "règle des trois consonnes"), les e muets s'affaiblissent en fin de mot. Le dernier disparaît complètement.*

Dès que nous avons commencé à travailler sur la rime, l'élimination du *e* muet en fin de vers nous a paru être un mauvais choix. En effet, elle gommait la distinction entre rime féminine et masculine essentielle à la rime. Nous avons dû réintroduire cette distinction. Il nous faudra garder ultérieurement une trace de ces *e* muets élidés.

Autrement dit, sont considérés comme *e muets forts* :

- les *e* muets de monosyllabes non élidés graphiquement ;
- les *e* muets internes qui, s'ils étaient élidés, conduiraient à une succession de trois consonnes.

Sont considérés comme *e muets faibles :*

- les *e* muets en fin de mots polysyllabiques ;
- les *e* muets à l'intérieur d'un mot (sauf s'ils sont forts).

Donnons quelques exemples de traitement qui montrent nos convergences et divergences avec Milner et Regnault. Soit le vers analysé en détail p. 45.:

Dans une prochaine version, il nous faudra garder la trace du *e* muet surtout pour l'étude de la rime.

[98] Rappelons que l'entrée courante n'est pas toujours un vers, mais peut également être un fragment de vers, ce qui conduit à un certain nombre d'erreurs, en appliquant cette règle, quand la réplique qui suit le fragment initial commence par une consonne. C'est une des raisons qui nous a conduit à reconstituer les vers.

Où suis-je ? Qu'ai-je fait ? Que dois-je faire encore ?

Racine, *Andromaque*, vers 1393.

Les deux premiers *e* muets ainsi que le quatrième sont faibles (élidés en langue). En effet, dans les cas d'inversion, le verbe et l'enclitique forment une seule unité polysyllabique. Le troisième est fort (maintenu en langue), et les deux derniers sont élidés (règles (1') et (3')), ce qui conduit à la transcription suivante :

(u)(s ɥ i)(ʒ $ə_f$)(k ɛ)(ʒ $ə_f$)(f ɛ)(k ə)(d w a)(ʒ $ə_f$)(f ɛ r)(ɑ̃)(k ɔ r)

Notre système n'établit donc pas de différence entre les deux premiers *e* muets, alors que Milner et Regnault jugent le premier plus faible que le second. Pour le reste, nos analyses concordent.

IV.3. LA LIAISON : INCIDENCE SUR LE DECOUPAGE SYLLABIQUE

Fallait-il que, dans la représentation phonétique des vers produite par le métromètre, les phénomènes de liaison soient représentés ?

Nous avons introduit la liaison pour diverses raisons : les unes tiennent à la langue, les autres au vers, et la dernière relève de la pragmatique et tient aux mises en scènes contemporaines.

Commençons par les raisons de langue : la liaison est obligatoire en français dans un très grand nombre de cas. Ne pas indiquer les consonnes de liaison dans la représentation donnerait une vision incomplète de la structure phonétique du vers. Par ailleurs, nous avons choisi de représenter les positions ou syllabes métriques sous leur forme phonétique. Or la délimitation des frontières varie selon qu'il y a ou non liaison et selon son mode de réalisation (avec ou sans enchaînement).

Viennent ensuite deux raisons qui tiennent à la nature du vers. Premièrement, le décompte correct du vers implique un traitement de la liaison. En effet, quand le premier mot finit par un *e* muet suivi d'une consonne de liaison, c'est la promotion de la consonne de liaison qui permet le maintien du *e* muet. Dans le vers suivant de *Phèdre*,

Me nourrissant de fiel, de larmes abreuvée,

Racine, *Phèdre*, vers 1245.

si la liaison n'est pas faite, le *e* muet de *larmes* tend à s'élider ou à être absorbé (synalèphe) par le *a* de *abreuvée* : il manque alors au vers une position métrique.

(m ə)(n u)(r i)(s ɑ̃)(d ə)(f j ɛ l)(d ə)(l a r)(m ə z[99])(a)(b r œ)(v e) avec liaison

(m ə)(n u)(r i)(s ɑ̃)(d ə)(f j ɛ l)(d ə)(l a r m)(a)(b r œ)(v e) sans liaison

(11 positions métriques)

Deuxièmement, une des règles respectée dans le vers classique est l'absence de hiatus externe, c'est-à-dire de la rencontre de deux voyelles, l'une étant à la fin d'un mot, l'autre au début du suivant. Ne pas représenter la liaison, c'est multiplier les cas de hiatus externe dans la transcription.

Si mon cœur est *es*clave et tous ses vœux captifs,

(s i)(m ɔ̃)(k œ r)(ɛ)(t l ɛ s)(k l a v)(e)(t u)(s e)(v ø)(k a p)(t i f) avec liaison

(s i)(m ɔ̃)(k œ r)(ɛ)(ɛ s)(k l a v)(e)(t u)(s e)(v ø)(k a p)(t i f) sans liaison => hiatus : ɛ ɛ

Corneille, *Suréna*, vers 12.

Nous accordons enfin une place particulière à la liaison parce que notre corpus est constitué d'œuvres inlassablement jouées au théâtre, et qu'au fil des représentations de Corneille et Racine que nous avons suivies depuis cinq ans, il est clair que la place aujourd'hui accordée par metteurs en scène et comédiens à la liaison est très importante (*Dire le vers* n'est peut-être pas étranger à cette situation). Le taux de liaison y est particulièrement, voire anormalement, élevé par rapport à la langue et même par rapport à des mises en scène antérieures. Nous avons ainsi entendu en 1998, dans *Horace* mis en scène par Marion Bierry :

[99] Nous reviendrons ultérieurement sur la notation des consonnes de liaison et sur leur place dans la syllabe.

Tigres, zallez combattre, et nous, zallons mourir.

Corneille, *Horace*, vers 694.

Le tableau suivant donne à titre indicatif les consonnes latentes et leur réalisation en situation de liaison ainsi que quelques exemples de réalisation de ces liaisons observés dans le cadre d'une mise en scène d'*Horace* par M. Bierry. Les exemples auraient pu être tirés de bien d'autres mises en scène contemporaines, puisque la plupart portent une attention toute particulière à la bonne réalisation des liaisons. Mais cette mise en scène représentait la position la plus extrême puisque toutes les liaisons possibles étaient faites de manière très appuyée.

Gra-phie	Réalisation phonétique	Exemples « attestés » dans *Horace*, mise en scène de Marion Bierry, 1998	
n	n	Ne le considérez qu'en objet de leur haine, (vers 1524)	
r	r	Je crains pour l'une et l'autre en ce dernier effort, (vers 89)	
c, g	k	Beaucoup par un long(k) âge ont appris comme vous (vers 1461)	
s, x, z	z	Tigres, allez combattre, et nous, allons mourir. (vers 694) J'irai rendre pour vous grâces aux immortels. (vers 346)	Tigres, allez... sans liaison dans la mise en scène de J. -P. Miquel
t, d	t	Quand il ne punit point ce qu'il n'approuve pas. (vers 1434) Il n'y faut plus penser : en l'état où je suis, (vers 569)	

Tableau 18. Les consonnes de liaison : réalisation, exemples

IV.3.1. La liaison dans le vers

Pour l'essentiel, nous suivrons ici à nouveau Milner et Regnault [1987, p. 49-68], qui proposent un mode d'explication du fonctionnement de la liaison dans le vers cohérent avec les hypothèses avancées pour analyser le comportement du *e* muet. On part de l'hypothèse que *e* muet et consonne de liaison sont présents dans la représentation phonologique[100].

[100] D'après Milner [1982, p. 288], l'approche qui consiste à considérer que les consonnes latentes, tout comme les *e* muets, font partie de la forme primitive et qu'elles subissent une

Milner [1982, p. 288] pose deux règles pour les consonnes finales en langue :

> Chute de consonne finale :
> *C tombe à la fin du mot phonologique, quelque soit la nature de l'élément suivant*.
> Troncation :
> *C tombe à l'intérieur du mot phonologique devant une consonne*

Les règles pour le *e* et pour les consonnes de liaison en langue ont la même forme en substituant *e* en C, et V en C pour le contexte. Les règles pour le *e* comme pour la liaison s'appuient sur la notion de mot phonologique.

Autrement dit, toutes les liaisons sont faites à l'intérieur du mot phonologique (les consonnes de liaison sont promues ou prononcées) et

> Il n'y a pas de liaison d'un mot phonologique au mot phonologique suivant.
> [id. p. 50].

Puis, pour passer de la langue au vers, ils appliquent le « Principe constituant » (chaque vers fonctionne comme un mot phonologique), ce qui permet une formulation élégante du fonctionnement de la liaison dans le vers :

> - en droit, toutes les liaisons se font à l'intérieur du vers ;
> - en droit, aucune liaison ne se fait d'un vers à l'autre.
> [id. , p. 52].

En théorie, cette règle, couplée avec celle qui concerne les *e* muets, suffit à fournir un cadre d'analyse unifié qui permet de prendre en compte simultanément ces deux phénomènes, puisque ce sont les consonnes de liaison qui, une fois promues, vont imposer le maintien du *e* muet.

Néanmoins, et pour nuancer les effets de ces règles brutales qui entraînent quelques liaisons particulièrement disgracieuses et surprenantes[101], ils en viennent à distinguer deux types de diction de la consonne latente :

- la *diction directe* qui consiste à syllaber la consonne de liaison avec la syllabe suivante ;

---

opération d'effacement dans certains contextes est due à Bloomfield, puis à Schane [1968] et à Milner.

[101] Liaisons entre mots séparés par une ponctuation forte, entre autres anomalies.

- la *diction indirecte* qui consiste à maintenir dans la syllabe d'origine la consonne de liaison, et qui correspond à une réalisation discrète ou allégée.

Dans le cas de la liaison directe, la consonne latente constitue l'attaque de la syllabe initiale du mot lié (liaison enchaînée), alors que pour la liaison indirecte, elle reste attachée à sa syllabe d'origine (liaison sans enchaînement).

La liaison est directe chaque fois qu'elle est également réalisée en langue.

Dans le vers suivant de *Britannicus*, le segment *vous allumez un feu* forme un seul mot phonologique de langue (qui coïncide avec le premier hémistiche), la liaison entre *vous* et *allumez* et entre *allumez* et *un* sera donc directe.

Vou//z allume//z un feu qui ne pourra s'éteindre.

Racine, *Britannicus*, vers 1351.

La liaison est indirecte quand il n'y a pas de liaison en langue ou quand la liaison coïncide avec une frontière métrique, la césure par exemple[102]. Dans le vers suivant, Milner et Regnault recommandent une liaison indirecte parce que *ses feux* constitue, en langue, un mot phonologique à part entière :

Pour la veuve d'Hector, ses feu^{z} // ont éclaté

Racine, *Andromaque*, vers 108.

Prenons un exemple de mot phonologique en langue qui traverse l'hémistiche. En langue *venger un père* constitue un seul mot phonologique, mais, dans le vers suivant, *venger* coïncide avec la fin du premier hémistiche : la liaison indirecte marque ici le respect de la frontière métrique.

Qu'ils semblent tous venge^{r} // un père, comme vous.

Corneille, *Cinna*, vers 152.

Comme en général, dans le vers classique, le mot phonologique ne traverse que rarement l'hémistiche, **une liaison indirecte indique en général la présence d'une frontière de mot phonologique de langue.**

Cette règle de liaison systématique à l'intérieur du vers souffre de plus de quelques exceptions : pas de liaison quand il y a collision de consonnes

[102] Cette difficulté serait d'emblée résolue si nous avions considéré l'hémistiche et non le vers comme mot phonologique.

identiques (*qu'aussitôt il arrête*), interjections, vocatifs... On ajoutera aussi comme exception la coordination *et* dont le *t* n'est jamais promu. Cela a une conséquence directe dans le vers : on ne trouve jamais de séquence *et*#V, à une seule exception près :

> Je suai sang et eau, pour voir si du Japon
>
> Racine, *Les Plaideurs*, vers 689.

On note ici qu'il s'agit d'un syntagme figé, ce qui justifie peut-être le hiatus, qui ressemble alors plutôt à une diérèse... On voit combien les frontières de mot sont un sujet difficile.

Pour résumer, à l'intérieur du vers, toutes les consonnes latentes doivent être promues et représentées dans la structure phonétique (ne serait-ce que pour mettre en évidence l'absence de hiatus), l'opposition liaison directe/indirecte permettant de différencier les liaisons fortes et obligatoires de celles qui le sont moins.

Avant d'aller plus loin nous voudrions revenir sur le choix du couple liaison directe/indirecte.

IV.3.2. Liaison avec / sans enchaînement ou forte /faible ?

Si la prise en compte des remarques de Milner et Regnault rend immédiat le repérage des consonnes latentes qui vont se maintenir, la seconde partie de leur programme (évaluer leur force) est bien plus difficile à réaliser, puisqu'il faut, en toute rigueur, pouvoir distinguer les liaisons qui se font en langue des autres.

Pour ces auteurs, toutes les liaisons se font à l'intérieur d'un mot phonologique, et jamais d'un mot phonologique à l'autre. Or la réalité est sensiblement plus complexe. La liaison en français a été l'objet de nombreuses études (voir par exemple Encrevé [1988]), qui toutes s'accordent sur le fait que ce phénomène est hautement variable (suivant les locuteurs, les niveaux de langue, les conditions d'élocution).

Reprenons ce qu'en dit Encrevé [1988]. Les phénomènes de liaison peuvent se classer en trois catégories : les liaisons invariables produites systématiquement par tous les locuteurs ; les liaisons variables produites par certains locuteurs et/ou dans certains contextes de discours ; les liaisons erratiques [Encrevé, 1988, p. 47]. L'analyse d'un ensemble de discours politiques enregistrés lui a entre autre permis de montrer que :

> - dans les contextes de liaison où la consonne de liaison est réalisée invariablement (liaison obligatoire), cette consonne est invariablement enchaînée à la voyelle du mot suivant ; (...)
> - dans les contextes de liaison où la consonne de liaison n'est réalisée que variablement (liaison facultative), elle n'est que variablement enchaînée (...).
>
> [Encrevé, 1988, p. 75].

Autrement dit, dans un contexte de liaison (c'est-à-dire devant voyelle), une consonne latente peut être : réalisée et enchaînée ; réalisée et non enchaînée ; non réalisée.

Complétant son corpus par des discours politiques antérieurs aux années 60, Encrevé montre l'évolution de l'usage. Alors que l'enchaînement était systématique en langue soutenue, il est devenu au fil du temps un élément de variabilité, puisque le taux de liaisons non enchaînées tend à croître.

Milner et Regnault font le choix de simplifier la question de la liaison *dans la langue*, en considérant qu'en droit toutes les liaisons se font à l'intérieur du mot phonologique ; ils négligent pour la langue la variabilité de certaines liaisons et le mode de réalisation variable de ces liaisons (avec ou sans enchaînement).

La complexité du traitement de la liaison est réintroduite dans le cadre du vers, lorsqu'ils distinguent liaison directe et indirecte, qui résultent de la confrontation entre la structure syntaxique de langue et la structure métrique.

	Encrevé Réalisation en langue	Milner et Regnault Réalisation en langue	Milner et Regnault Réalisation dans le vers
liaison invariable	liaison enchaînée	liaison directe	liaison directe
liaison variable	liaison enchaînée liaison sans enchaînement pas de liaison	liaison directe	liaison directe
liaison erratique	pas de liaison	pas de liaison	pas de liaison
pas de liaison		pas de liaison	liaison indirecte

Tableau 19. Le traitement de la liaison

Le Tableau 19 appelle quelques remarques : Milner et Regnault utilisent les notions de liaison indirecte et directe pour différencier les cas où il y a discordance, respectivement concordance, entre structure métrique et syntaxique. La description phonétique que Milner et Regnault donnent des liaisons directes/indirectes correspond exactement à ce que Encrevé désigne sous les termes de liaisons avec/sans enchaînement.

Or, ce phénomène de liaison sans enchaînement semble être apparu assez récemment dans le discours politique. C'est d'ailleurs un mode de réalisation de la liaison proscrit par les traités normatifs sur la liaison. Ce type de réalisation de la liaison sans enchaînement ne semble pas encore avoir été adopté par les comédiens. Dans les représentations auxquelles nous avons assisté et dans les enregistrements écoutés, les cas de liaison sans enchaînement sont rares voire inexistants. La rareté des liaisons indirectes rend le couple liaison directe / indirecte non pertinent pour distinguer phonétiquement, dans toutes les liaisons du vers, celles qui sont obligatoires en langue (liaison directe), de celles qui ne sont pas admises en langue (liaison indirecte). L'intensité de la liaison serait un élément de distinction plus pertinent. D'ailleurs, pour Milner et Regnault, la liaison indirecte correspond à une diction allégée et permet de distinguer liaison faible et liaison forte. C'est la distinction de deux niveaux d'intensité dans la liaison que nous retenons à travers cette distinction.

Dans le métromètre, le codage liaison directe/indirecte, même s'il ne correspond en rien à une réalité de la diction, nous est donc fort utile. Il permet de différencier les deux niveaux d'intensité (fort / faible). Il peut aussi être interprété en termes de liaison invariable / variable. Quand la consonne de liaison est maintenue dans sa syllabe d'origine, cela signifie simplement qu'il s'agit d'une liaison variable ; quand au contraire elle est enchaînée, il s'agit

d’une liaison obligatoire. On notera que, dans nos transcriptions, les consonnes de liaison ont un code spécifique, pour les distinguer des consonnes toujours prononcées. Par exemple, le symbole « t_l » désigne la réalisation de la consonne latente /t/, qui vient se placer soit en coda de sa syllabe d’origine, soit en attaque de la suivante.

IV.3.3. Formalisation

Une des difficultés que pose le système de Milner et Regnault est l’identification des mots phonologiques de langue. Ils considèrent que la « détermination des mots phonologiques est simple et facile : en effet, l’analyse syntaxique, qui détermine tout ici, ne réclame rien de plus qu’un peu d’attention » [p. 28], mais celle-ci n’est pas simple à automatiser. Pour affiner le traitement de la liaison, dans la mesure où nous cherchons, non pas à construire une transcription de ce qui devrait être dit, mais une transcription abstraite qui permette de fournir une base pour effectuer des dénombrements, nous nous sommes fixés une référence, à savoir les règles de liaison édictées par Fouché [1959], qui, si elles correspondent souvent à des prononciations « vieillies », ont le mérite d'être très complètes.

Concrètement, à la fin de chaque mot, la possibilité d’une liaison en langue est évaluée compte tenu de la structure syntaxique du segment analysé ; on en déduit une valeur pour le séparateur de mots correspondant. Ainsi, l'espace qui suit *petits* dans *petits enfants* est affecté du trait « liaison autorisée », contrairement à celui qui suit *enfants* dans *des enfants aimables*. Les signes de ponctuation sont bien sûr marqués « liaison interdite ».

Au cours de la dérivation phonologique, la suppression des consonnes latentes est réalisée par des règles inspirées de la règle de troncation des obstruantes finales (la *French Truncation Rule* de Schane [1968], qui supprime les consonnes latentes dans les cas de non-liaison). Cette règle fonctionne essentiellement comme suit pour les obstruantes :

(a) *suivie d’une consonne (ou d'un* h *aspiré), la consonne latente est toujours supprimée ;*

(b) *suivie d'un séparateur « liaison autorisée » et d'une voyelle, la consonne se maintient et est resyllabée avec la voyelle (liaison enchaînée ou directe) ;*

(c) *suivie d'un séparateur « liaison interdite » et d'une voyelle, la consonne se maintient, mais sans resyllabation (liaison sans enchaînement ou indirecte).*

Voici un exemple :

Et mes coursiers oisifs ont oublié ma voix,
(e)(m *e*)(k u r)(s j e z_l)(w a)(z i f z_l)(ɔ̃)(t_l u)(b l i)(*e*)(m a)(v w a)

Racine, *Phèdre*, vers 552.

.

Ce vers comporte trois liaisons. Les deux premières (entre *coursiers* et *oisifs* et entre *oisifs* et *ont*) sont des liaisons indirectes : la consonne latente reste attachée à la voyelle précédente. Entre *ont* et *oublié*, la liaison est directe : la consonne latente est syllabée avec la voyelle suivante et correspond à un cas de liaison obligatoire en langue.

Des règles similaires gouvernent la réalisation des autres consonnes de liaison. Pour /n/, les cas (a) et (c) entraînent de surcroît la nasalisation de la voyelle qui précède, nasalisation inhibée dans le cas (b). Pour /r/, le cas (c) cause l'ouverture du /e/ qui précède en /ɛ/.

Des crimes de Néron approuver les horreurs,
(d *e*)(k r i)(m $ə_f$)(d ə)(n *e*)(r ɔ̃ n_l)(a)(p r u)(v *e*)(l *e*)(z_l o)(r œ r)

Racine, *Bérénice*, vers 353.

Quels charmes ont pour vous des yeux infortunés
(k ɛ l)(ʃ a r)(m ə z_l)(ɔ̃)(p u r)(v u)(d e)(z_l j ɔ z_l)(ɛ̃)(f ɔ r)(t y)(n e)

Racine, *Andromaque*, vers 303.

.

Comme notre traitement ne fait, à aucun moment, l'hypothèse que l'entrée analysée est un vers, nous avons néanmoins dû sacrifier une des règles proposées par Milner et Regnault. Ainsi, quand un mot phonologique de langue traverse l'hémistiche, alors qu'ils recommandent une diction indirecte (puisque les structures métrique et syntaxique ne coïncident pas), nous avons conservé la liaison directe (car nous ne savons pas *a priori* où se trouve l'hémistiche, ni

même s'il en existe un). Nous n'avons pas non plus pris en compte les exceptions relevées par ces auteurs : la liaison ne doit pas être faite avec certaines interjections comme *ah*, *hélas* et elle doit être très affaiblie quand elle entraîne une collision de consonnes. Par exemple, dans l'hémistiche : *Lever les yeux au ciel* (*Andromaque)*, la seconde liaison devra être faite de manière très atténuée.

Il est clair (et cela transparaît nettement dans nos résultats), que la liaison est le point le plus fragile de notre système. De nombreuses exceptions ne sont pas encore prises en compte dans ce système, qui permettraient d'affiner marginalement nos transcriptions. La liaison reste de toute façon un problème difficile, que les outils de traitement automatique du langage gèrent encore de façon très imparfaite.

Voici un exemple d'erreur, lié à une mauvaise analyse de la métaposition (*cf.* chapitre VI.2) :

> Lorsqu'il vient à la reine expliquer son amour
> (l ɔ r)(s k i l)(v j ɛ̃)(t͡l a)(l a)(r ɛ n)(ɛ k)(s p l i)(k *e*)(s ɔ̃)(n͡l a)(m u r)
>
> Racine, *Bérénice*, vers 6.

L'analyse syntaxique conduit à une mauvaise évaluation du degré de liaison entre *vient* et *à la reine*. En effet, *à la reine* dépend syntaxiquement de *expliquer* (expliquer X à Y).

Nous avons donc fait le choix de « promouvoir » toutes les consonnes latentes à l'intérieur du vers, pour éviter des cas d'élision de *e* muets et de hiatus, suivant la règle proposée par Milner et Regnault. Et pour affiner, nous avons codé différemment les *liaisons obligatoires* (celles qui le sont en langue) et les *liaisons facultatives* (non obligatoires, voire interdites en langue) : les premières (liaisons obligatoires) en enchaînant la consonne de liaison à la voyelle suivante, les secondes (liaisons facultatives) en maintenant la consonne de liaison dans sa syllabe d'origine. Le non enchaînement d'une consonne de liaison marque la présence de la frontière d'un mot phonologique de langue, ce qui pourrait jouer le rôle d'un marquage rythmique.

IV.4. SYNTHESE ET CONCLUSION

Récapitulons les phénomènes phonétiques propres au vers classique :

A l'intérieur des mots, les groupements vocaliques GV, c'est-à-dire, glissante (*i, ou* ou *u*) suivie d'une voyelle, correspondent tantôt à une position métrique (synérèse), tantôt à deux (diérèse) selon des règles précises, qui tiennent à l'origine et à la nature morphologique des mots. En moyenne, dans un mot sur deux contenant ce type de groupement, celui-ci constitue deux voyelles métriques. Mais en terme *d'occurrences*, la tendance à la synérèse l'emporte nettement sur la diérèse : seuls 22 % des occurrences de mots avec ce type de groupement vocalique correspondent à des cas de diérèse (rencontre de deux voyelles métriques). Tout se passe comme si la fréquence d'usage tendait à éliminer les cas de diérèse.

Tous les *e* muets (présents graphiquement) à l'intérieur du vers occupent une position métrique sauf :

- s'ils précèdent un mot à initiale vocalique (Ce#V => CV) ;
- s'ils sont en fin de vers ;
- s'ils constituent un *e* mort dans des séquences du type *aient*.

On peut parler d'élision métrique du *e* muet, ou plus justement de synalèphe : fusion des deux voyelles en une seule. Par cette dénomination, on rapproche le fonctionnement du français, de l'italien et de l'espagnol. En effet, dans le vers italien ou espagnol, l'absence d'attaque entraîne systématiquement la fusion de la voyelle nue avec la voyelle post-tonique précédente en une seule position métrique. Le voyelles qui suivent la voyelle accentuée dans le vers italien ou espagnol subissent les mêmes lois d'élision que le *e* muet dans le vers français : synalèphe devant voyelle, décompte devant consonne ou voyelle accentuée.

Tengo una parte aquí de tus cabellos
(t e n)(g u)(n a)(p a r)(t a)(k i)(d e)(t u s)(c a)(b e)(ll o s)
Je garde une partie de tes cheveux

Garcilaso de la Vega, « Égloga primera », trad. P. Bonhomme in [*Anthologie bilingue de la poésie espagnole*, p. 266].

Cet hendécasyllabe comporte deux synalèphes : entre *tengo* (accentué sur la première syllabe) et *una* et entre *parte* (accentué sur la première) et *aquí.* La synalèphe en italien et en espagnol et l'élision du *e* muet en français relèvent de la même logique : éviter les hiatus en cas d'absence d'attaque consonantique dans une syllabe, surtout quand les voyelles ne sont pas toniques (c'est toujours le cas des *e* muets, et très largement le cas des voyelles finales en espagnol et italien : la majorité des mots étant des « *palabras llanas* », des mots « plats », accentués sur l'avant-dernière syllabe).

Le traitement de la liaison permet de promouvoir des consonnes latentes (Cl) et par conséquent de modifier le contour des syllabes. Elle permet surtout d'éviter les cas de hiatus externe (CVCl#V devient CVCV au lieu de CV#V, si la liaison n'était pas faite).

Milner [1974 et 1982] a bien montré le parallélisme qui pouvait être fait entre le traitement du *e* muet et celui de la liaison. Dans les deux cas, c'est le vers, considéré comme un seul mot phonologique, qui constitue l'unité d'analyse.

Nous n'avons pas insisté sur la question du *h* initial qui peut soit jouer le rôle d'attaque syllabique s'il est aspiré, soit correspondre à une attaque nulle. Il va sans dire que pour le décompte il est indispensable de connaître sa nature et cela ne peut être fait qu'à travers des listes. Autrement dit, un *e* muet devant un *h* non aspiré s'élide (premier vers ci-dessous), alors qu'il occupe une position métrique devant un *h* aspiré (second vers).

Aujourd'hui qu'il peut tout, que votre hymen s'avance,
(o)(ʒ u r)(d ɥ i)(k i l)(p ø)(t u)(k ə)(v ɔ t r)(i)(m ɛ n)(s a)(v ɑ̃ s)

Racine, *Bérénice*, vers 383.

Attacha pour jamais une haine puissante ;
(a)(t a)(ʃ a)(p u r)(ʒ a)(m ɛ)(zɭ y)(n ə)(ɛ)(n əʃ)(p ɥ i)(s ɑ̃ t)

Racine, *Bérénice*, vers 383.

Les questions de la diérèse, du *e* muet et de la liaison viennent d'être traitées les unes après les autres comme si chacune était indépendante des autres. Ces phénomènes résultent tous trois de tentatives pour résoudre les difficultés posées par l'absence de consonne initiale, d'attaque, dans la structure syllabique. Comme le rappelle Klein [1991, p. 13], « les phénomènes de liaison, élision et synérèse ne sont que des effets différents d'une même structure d'attaque nulle ».

Allons plus loin et proposons une conjecture : la logique des différentes règles phonologiques dans le vers pourrait se résumer à un effort pour produire la concaténation de douze segments de type CV. Autrement dit, comme le sait tout écolier, le modèle de l'alexandrin serait ta ta ta ta ta ta ta ta ta ta ta ta.

Sans aller jusqu'à considérer que la suite CV soit un modèle universel, celle-ci est la forme de syllabe la plus fréquente en français. Cette structure syllabique représente 55,6 % des occurrences dans le corpus de Wioland [1985, p. 260] comprenant 87 000 syllabes. Pour renforcer cette alternance CVCV, un certain nombre de phénomènes apparaissent dans la langue : élision (du *e* des pronoms et articles ; des *e* muets finaux), liaison, enchaînement ; variantes pour les déterminants possessifs (*mon amie*) ; variantes pour certains démonstratifs (*cet inconnu*, *bel été*...)... La langue cherche à éviter le hiatus (séquences syllabiques de type VV), surtout à l'intérieur des mots phonologiques et met en place un certain nombre de mécanismes pour lisser la séquence des syllabes à l'intérieur du mot phonologique.

Comme dans la langue, nous observons les mêmes phénomènes dans le vers mais réglés de manière plus rigoureuse, surtout dans le vers classique : tendance à la synérèse à l'intérieur du mot (les cas de synérèse sont trois fois plus nombreux que les cas de diérèse), élision systématique du *e* muet devant voyelle, promotion systématique, devant voyelle, des consonnes latentes dans le vers.

De plus, cet effort vers une forme abstraite la plus régulière possible, que l'on observe dans la langue comme dans le vers, s'est vue renforcé à l'époque classique par un ensemble de règles posant des interdits. Le vers classique en respecte deux :

- l'interdiction du hiatus externe (V#V) ;
- l'interdiction des séquences Ve#C et VeCL#C.

Ces règles sont toujours respectées chez Corneille et Racine à trois exceptions près.

En ce qui concerne le hiatus, aucune autre voyelle que le *e* muet n'est atone en français. Or comme le français est une langue oxytone, un V en fin de mot est forcément tonique, ce qui rend impossible son élision, et son maintien met alors en évidence l'absence d'attaque. Il y a cependant une différence entre une tendance à éviter le hiatus, ce qui devait être le cas avant le XVIIe siècle, et la mise en place d'une règle normative, comme celle qu'a imposée Malherbe. D'ailleurs, comme nous l'avons vu dans ses premiers écrits, Malherbe ne la respectait pas : dans « Larmes du Sieur Malherbe », sur 160 vers on ne trouve pas moins de 14 cas de hiatus. Si le hiatus externe est interdit, le hiatus interne, la diérèse, ne l'est pas. Or la seule différence entre ces deux cas tient à la présence d'une frontière de mot dans le premier cas. La distinction qui est faite dans le vers classique entre hiatus externe et diérèse donne un poids particulièrement important à la frontière de mot. Les expériences d'Anne Christophe [1993], dans le Laboratoire de sciences cognitives et psycholinguistique (CNRS-EHESS), montrent que dans des échantillons de parole où les mots du lexique ne sont pas reconnaissables, les frontières de mots sont perçues prosodiquement par les sujets. Puisque la frontière de mot est perçue par les locuteurs, diérèse et hiatus externe constituent des situations prosodiquement différentes. Mais cela n'explique pas pourquoi la diérèse n'est pas interdite.

Quant à la seconde règle, qui interdit les séquences Ve#C, c'est encore une fois Malherbe qui l'a imposée. Est-ce vraiment parce que cela souligne trop la contradiction entre la prose (qui ne prononce plus le *e* après voyelle) et le mètre qui impose que le *e* muet constitue à lui seul une position métrique ?

En revanche, comment expliquer que les séquences Ve#V soient quant à elles autorisées ? Le *e* muet est élidé devant voyelle, et l'on se retrouve alors devant un cas de hiatus, à moins de considérer que l'élision du *e* laisse une trace (d'où la préférence pour le terme de synalèphe).

Cornulier [1977, p. 108], partant du fait que la règle d'élision devant voyelle est un universel métrique, indique qu'une suite V#V est :

> Sujette à la règle métrique d'élision, dont la non-application constitue une irrégularité métrique, et dont l'application entraîne une inacceptabilité grammaticale. Ce qu'on appelle règle de l'hiatus n'est donc qu'une conséquence bâtarde de l'obligation métrique d'élider jointe à l'impossibilité grammaticale d'élider.
>
> [Cornulier, 1977, p. 108].

De même pour Cornulier, l'interdiction Ve#C tient à une interférence de la grammaire et de la métrique, la première exigeant une suppression du *e* que la métrique refuse. L'interdiction de ces situations serait un moyen d'éviter les contradictions entre l'ordre syntaxique et l'ordre métrique.

On pourrait poser que la composante rythmique de la langue fait tendre la syllabe vers un modèle simplifié CV, qui correspond à une courbe de sonorité ascendante. Comme le rappelle Angoujard [1992, p. 57-67], le rythme binaire (creux-sommet) y joue un rôle fondamental. L'enchaînement des courbes de sonorité d'un modèle CVCVCV... serait purement métrique. Dans le vers, d'une manière peut-être plus régulée que dans la prose, parce que la position métrique est constitutive du vers, le rythme binaire creux-sommet est plus recherché qu'ailleurs. Néanmoins, les variations des formes syllabiques introduisent une composante rythmique dans la séquence de langue. Ainsi la forme de la syllabe oscille-t-elle entre l'identité parfaite au modèle CV et d'autres formes moins métriques.

Rappelons enfin les règles phonologiques adoptées pour le métromètre :

- Le découpage syllabique correspond classiquement au découpage syllabique de langue : une consonne entre voyelle est rattachée à la voyelle qui suit, une consonne à l'initiale d'un mot aussi. Les groupes de consonnes sont toujours scindés sauf quand ils sont formés d'une obstruante et d'une liquide. Les consonnes finales sont rattachées à la voyelle précédente sauf en cas de liaison enchaînée ;
- Pour la diérèse, nous avons ajouté un trait « diérèse » ou « non diérèse » aux formes du lexique, en appliquant des règles ou en codant des exceptions. La présence du trait « diérèse » empêche la semi-vocalisation de la voyelle haute ;

- Pour le *e* muet, trois règles sont appliquées : élision du *e* avant voyelle, élision du *e* après voyelle et affaiblissement des *e* muets en fin de mot (sauf monosyllabes et règle des trois consonnes).
- Pour la liaison, devant voyelle, toutes les liaisons sont faites dans le vers : liaison avec enchaînement (la consonne latente est rattachée à la voyelle suivante) quand elle correspond à une liaison obligatoire en langue, sans enchaînement (la consonne latente est maintenue dans sa syllabe d'origine quand la liaison est interdite ou facultative en langue. Une consonne latente devant consonne s'élide.

V. MARQUAGES POSSIBLES, MARQUAGES CHOISIS

Nous avons explicité dans le chapitre précédent les règles qui permettent au métromètre de construire correctement, à partir du vers écrit, la séquence des positions métriques. Cela permet de découper les segments que nous délimitons par des parenthèses. Dans chacun de ces segment, le contenu des positions est transcrit dans l'alphabet phonétique international.

Allumé le flambeau d'un hymen odieux,
(a)(l y)(m e)(l ə)(f l ɑ̃)(b o)(d œ̃)(nɟ i)(m ɛ n)(ɔ)(d i)(ø)

Racine, *Bérénice*, vers 402.

Le métromètre nous permet donc déjà d'identifier le nombre de positions et le contenu phonétique des positions, en respectant les règles de dénombrement du vers classique.

Cependant, nous attendons d'un outil d'analyse systématique du vers qu'il nous offre une représentation plus riche du vers. Puisque, en effet, nous souhaitons construire la figure *rythmique* du vers, il nous faut introduire d'autres dimensions qui contribuent au rythme. Les positions métriques jouent bien évidemment un rôle dans l'élaboration du rythme, mais d'autres phénomènes s'élaborent à d'autres niveaux.

Nous partons de l'hypothèse que le rythme, en tant que forme, résulte de la mise en œuvre, « feuilletée » ou superposée, de figures d'alternance du même et du différent pour les composantes phonétique, lexicale, syntaxique, prosodique et même sémantique. Il existe évidemment des liens entre ces différentes composantes linguistiques, mais chacune jouit cependant d'une autonomie suffisante pour être envisagée séparément. Elles s'organisent toutes ensemble pour former le rythme.

Pour entendre cette unité du rythme, il faut pouvoir la décomposer, autrement dit, appréhender les parties pour comprendre le tout. Le système de

marquages proposé par Lusson [1974] consiste à se donner des propriétés et à indiquer si chaque propriété est réalisée pour un événement élémentaire donné (une position dans le vers par exemple). Ce système est un moyen simple et efficace de description de chacune des composantes qui constituent le rythme. A chacune correspondront des types de marquages spécifiques : les marquages phonétiques ne seront pas de même nature que les marquages syntaxiques ou prosodiques, puisque les propriétés étudiées varient.

Si les marquages portent sur des composantes linguistiques différentes, ils peuvent cependant tous s'exprimer dans un codage binaire de type présence/absence ou 1/0. Ceci permet de *normaliser* la description des différents niveaux d'analyse, ce qui ouvre la voie à une analyse globale des différentes composantes.

Lusson et Roubaud [Lusson et Roubaud, 1981 et 1983 ; Lusson, 1982] ont parfois travaillé sur des ensembles de marquages corrélés qu'ils ont représenté par une échelle de poids, construite en général sur des critères phono-syntaxiques. Une position donnée dans le vers acquiert la valeur 0 si elle correspond à un *e* muet, 1 si c'est une syllabe pleine interne, 2 si elle précède une frontière de mot... Cette échelle de poids est une manière de synthétiser différents marquages en un seul. Cela implique de pouvoir identifier finement les catégories syntaxiques, ce que notre outil n'a pas su faire.

Quand Meschonnic définit le rythme, en le considérant comme la convergence de différentes composantes, il ne s'éloigne pas tant qu'il le prétendait de l'approche de Lusson :

> Je définis le rythme dans le langage comme l'organisation des marques par lesquelles les signifiants linguistiques et extra-linguistiques (dans le cas de la communication orale surtout) produisent une sémantique spécifique, distincte du sens lexical, et que j'appelle la signifiance : c'est-à-dire les valeurs propres à un discours et à un seul. Ces marques peuvent se situer à tous les niveaux du langage : accentuelles, prosodiques, lexicales et syntaxiques. Elles constituent ensemble une paradigmatique et une syntagmatique qui neutralisent précisément la notion de niveau.
>
> [Meschonnic, 1982, p. 216-217].

Concrètement, cette « signifiance » du rythme naît de la convergence et du renforcement des effets liés aux différents niveaux d'appréhension. Dans une

visée d'analyse pragmatique de corpus, il nous faut décomposer les niveaux d'analyse, adopter un mode de codage unique (les marquages) pour ensuite faire converger les différents niveaux d'analyse et repérer les figures émergentes.

Pourtant, même si Meschonnic considère le rythme comme résultant de toutes les dimensions du discours, il refuse catégoriquement la notion de marquage, qui dit implicitement l'existence d'une mesure. Car pour Meschonnic, il n'y a pas de mesure dans le vers :

> Si l'intervalle [entre deux accents] compte autant que l'accent, (...) ; si la mesure est une abstraction inutile et nuisible ; si la prosodie, autant que les limites de groupe, sont parties intégrantes du rythme, — la rythmique ne peut plus compter, chiffrer. La notation doit être qualitative, non plus quantitative. Elle excède la statistique. Comme la prosodie fait rythme à l'intérieur du vers, non-métriquement, la notation rythmique-prosodique excède la scansion.
>
> [Meschonnic, 1982, p. 252].

Selon Meschonnic, on ne peut donc avoir une approche du rythme qui soit quantitative, seule l'approche qualitative a un sens. D'emblée, il se situe dans une position où il s'interdit de poser des propositions falsifiables, de soumettre à validation ses hypothèses. À travers ce refus, nous retrouvons un vieux débat stérile qui pollue les sciences sociales : le débat entre les approches qualitatives et quantitatives. Ce débat se retrouve en sociologie dans les oppositions entre les partisans des entretiens/analyse de contenu d'une part, et des enquêtes sur des échantillons représentatifs de quelque chose/traitements statistiques, de l'autre. Dans la pratique, les meilleures études sociologiques ne se privent pas de croiser les deux dimensions (*cf. La Distinction* de Pierre Bourdieu, pour ne citer que ce grand classique) : qui peut le plus peut le moins. De fait, les chiffres ne rebutent que ceux qui les trouvent difficiles d'accès. Si l'approche statistique ne permet pas une analyse locale extrêmement fine, elle a tout de même le mérite de dresser des figures globales et de relativiser la place de phénomènes tout à fait marginaux. Plus généralement, le débat a été résolu au niveau méthodologique, et l'on s'accorde maintenant sur l'avantage de l'approche complémentariste qui consiste à croiser les méthodes pour mieux cerner l'objet.

Nous avons donc pris comme structure de base le vers décomposé en douze positions et nous avons projeté sur cette séquence d'événements élémentaires des « marquages » spécifiques, qui vont être décrits dans ce chapitre. Par les marquages choisis, nous cherchons à mettre en évidence l'organisation du discours à l'intérieur du vers. Nous n'avons introduit, volontairement, que des marquages de langue ; aucun marquage métrique n'est entré en jeu. Nous cherchons à tester une hypothèse forte : *la structure métrique et rythmique du vers résulte d'un agencement particulier des éléments de discours et de cela seulement*. Pour prendre une métaphore, nous pensons que les effets de rythme sont construits par l'assemblage habile de propriétés du langage, un peu comme le marquetier produit des dessins par simple assemblage de bois ou de minéraux naturels, sans avoir besoin de changer leur couleur ou leur texture. Les « accents de vers » seront ainsi des « accents de langue » judicieusement placés par le poète. Ainsi, la césure n'existe que parce qu'un mot phonologique s'achève en sixième position, elle ne préexiste pas au placement de ce mot.

Si, pour l'identification des positions métriques, nous avons usé de règles propres au fonctionnement du vers, comme les règles du décompte du *e* muet ou de la diérèse, pour les marquages, nous n'avons utilisé *que des marquages de langue*, autrement dit que des informations phonético-syntaxiques produites par l'analyseur syntaxique du métromètre. Aucune marquage métrique (par exemple un marquage de fin de segment d'hémistiche) n'a été introduit ; aucune condition métrique n'intervient (comme celle qui consisterait à considérer par exemple que toutes les fins de segment métrique coïncident avec la fin d'un mot et avec un accent).

Dans ce chapitre, nous présentons le système de description multidimensionnel que nous avons adopté pour décrire les vers. Pour chaque composante linguistique (phonétique, lexicale, syntaxique, prosodique), nous décrirons l'ensemble des marquages qui paraissent envisageables, puis ceux que nous avons retenu, l'état de l'art actuel de la modélisation de la syntaxe des vers ne permettant pas de les mettre tous en œuvre. Nous envisagerons successivement les marquages phonétiques, les marquages lexicaux, les marquages syntaxiques et enfin les marquages prosodiques.

V.1. POSITIONS METRIQUES : VOYELLES ET SYLLABES PHONETIQUES

Comment se répartissent les voyelles et les syllabes dans le vers ? Telle est une des questions qui motive le marquage phonétique des positions. Ce marquage peut également permettre de répondre à des questions plus complexes sur l'architecture sonore des vers et « d'étudier l'aspect sonore du langage poétique comme un système formel capable de transmettre une information autre que celle des valeurs-sens (...) », comme le proposait Gauthier [1974, p. 40] dans son ouvrage, *Le Système euphonique et rythmique du vers français*. Les réseaux d'assonances et d'allitération, plus généralement les figures euphoniques[103], pour reprendre le terme de Gauthier [1974], sont une des composantes du rythme. Si le parallélisme est bien au cœur de la fonction poétique, comme projection de l'axe paradigmatique sur l'axe syntagmatique [Jakobson, 1963, p. 220], il nous faut pouvoir repérer ces parallélismes. Il n'existe pas encore de moyens pour explorer systématiquement les récurrences de structures : Gauthier [1974] a certes proposé la description de quelques figures euphoniques (oscillatoire et sérielles) faisant intervenir deux timbres, mais celles-ci ne couvrent qu'une partie des cas et ne sont pas aisément modélisables. En dernier lieu, les marquages phonétiques permettent d'explorer des questions spécifiques, comme celle du hiatus externe (absent dans le vers classique) : à partir de quand le hiatus disparaît-il du vers ; réapparaît-il au XIXe siècle?

Nous avons donc introduit deux marquages : un marquage syllabique qui indique la syllabe de chaque position et un marquage vocalique qui extrait de chaque syllabe son noyau vocalique. Plutôt que de proposer un marquage en 0/1 pour chaque syllabe ou voyelle (1 si la position est occupée par *a*, 0 sinon...), comme il n'y a qu'une syllabe et qu'une voyelle par position, on utilise une représentation multi-modale : le marquage syllabique consiste à indiquer le contenu de la syllabe sur chaque position, le marquage vocalique donne la

[103] Pour Gauthier [1974, p. 10] l'euphonie est « l'étude des phonèmes d'un poème, considérés en dehors des valeurs sémantiques qu'ils supportent par ailleurs, dans leur organisation et arrangement mutuel, et en liaison avec les cadres traditionnels de la prosodie, de la métrique et de la versification. »

voyelle métrique de chaque position. A partir de ces deux marquages, nous pouvons construire des marquages binaires (présence ou absence de telle syllabe, respectivement voyelle, sur telle position).

De cette nuit, Phénice, as-tu vu la splendeur ?

Racine, *Bérénice*, vers 302

	p. 1	p. 2	p. 3	p. 4	p. 5	p. 6	p. 7	p. 8	p. 9	p. 10	p. 11	p. 12
syllabe métrique	d ə	s ɛ	t əf	n ɥ i	f e	n i s	a	t y	v y	l a	sp l ɑ̃	d œ r
voyelle métrique	ə	ɛ	əf	ɥ i	e	i	a	y	y	a	ɑ̃	œ

Ces marquages phonétiques ne sont pas triviaux : ils sont dépendants des choix de découpage syllabique et de la définition des voyelles.

Commençons par le marquage syllabique. En introduction au chapitre IV et dans la partie consacrée à la liaison (section IV.3), nous avons présenté les choix de découpage syllabique choisis dans le métromètre : devant un mot à initiale vocalique, les consonnes finales et les consonnes de liaison indirecte (qui codent des liaisons facultatives) constituent la « rime » de leur syllabe d'origine, autrement dit elles sont syllabées avec leur mot d'origine. En revanche, pour les liaisons obligatoires, la Consonne de Liaison constitue l'attaque de la syllabe suivante.

Cet amour est ardent, il le faut confesser,

(s ɛ)(t a)(m u r)(ɛ)(t a r)(d ã t)(i l)(l ə)(f o)(k ɔ̃)(f e)(s e)

Racine, *Bérénice*, vers 421.

Le *r* final d'*amour*, consonne finale, est maintenu dans sa syllabe d'origine (si nous avions privilégié l'enchaînement systématique, il aurait pu être rattaché à la syllabe suivante) ; le *t* final d'*ardent* est également maintenu dans sa syllabe d'origine car il s'agit d'une liaison qui a été évaluée comme facultative en langue ; enfin le t de *cet* et de *est* ont tous les deux été rattachés à la syllabe suivante car ils correspondent à des cas de liaison obligatoire. Dans le métromètre, nous avons limité les cas d'enchaînements aux liaisons en préférant que les phonèmes soient rattachés à leurs mots d'origine.

En ce qui concerne les voyelles métriques, les noyaux vocaliques que nous avons conservés sont les suivants :

/ə/	(le, que, peupl**e#C**)
/əf/	(blêm**e#C**, flamm**e#C**, d**e**voir)
/a/ et /ɑ/	(courage, âme)
/ɑ̃/	(puissance, amant)
/ɛ/	(père, même)
/e/	(hyménée, parler)
/i/	(vie, empire)
/ɛ̃/	(rien, main)
/ɔ/	(mort, personne)
/o/	(autre, chose)
/ɔ̃/	(raison, honte)
/œ/	(cœur, pleurs)
/ø/	(yeux, lieu)
/œ̃/	(communs, importuns)
/u/	(vous, amour)
/y/	(vue, plus)
/wa/	(moi, gloire)
/wɛ̃/	(point, moins)
/ɥi/	(lui, ruine)

Clef de lecture : # indique une frontière de mot ; **C :** mot à initiale consonantique.

Tableau 20. Noyaux vocaliques des positions métriques

Cette liste de noyaux vocaliques n'est pas celle des phonèmes vocaliques du français, que l'on trouve habituellement dans les manuels. Elle est le *résultat* de tout le travail qui a été fait sur la rime (*cf.* chapitre III) et sur la diérèse (*cf.* chapitre IV.1).

Expliquons-nous : la démarche la plus naturelle, et c'est celle que nous avions adoptée dans un premier temps, consiste à considérer comme noyaux vocaliques les seize phonèmes vocaliques du français : /ə/, /a/, /ɑ/, /ɛ/, /ɑ̃/, /e/, /ø/, /i/, / ɛ̃/, /ɔ/, /œ/, / ɔ̃/, /o/, /u/, /ç/, /œ̃/, ceux que l'on trouve dans tous les manuels. Ainsi, le système vocalique du français est-il présenté dans *La grammaire d'aujourd'hui* [Arrivé et al. , 1986], en fonction de deux principes d'organisation : l'ouverture et l'antériorité (*cf.* Figure 20).

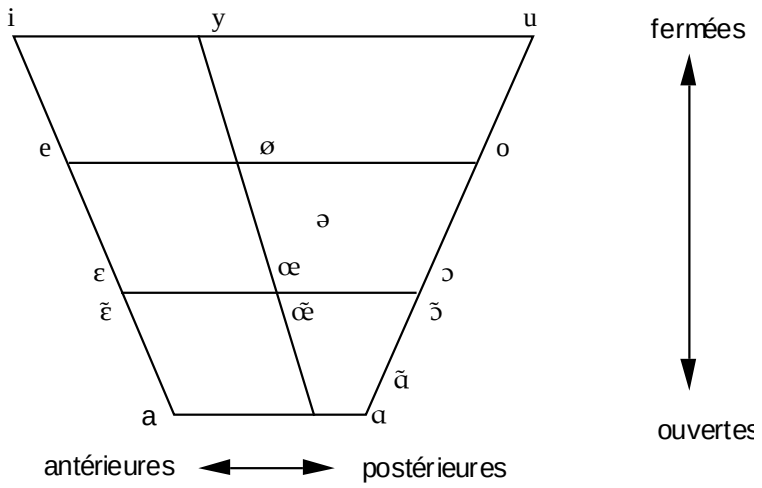


Figure 20. Positions respectives des différentes voyelles [Arrivé et al, 1986, p. 514]

Nous avons dans un premier temps retenu les seize voyelles du français habituellement présentées, la liste des voyelles du français ne paraissant pas problématique. Cependant, le travail sur la rime (*cf.* chapitre III) et l'examen de la diérèse (*cf.* chapitre IV.1) nous ont fait modifier la définition des voyelles : dans certains cas, nous avons attaché au noyau vocalique la semi-voyelle qui le précède. En effet, si la plupart des groupes vocaliques admettent selon le contexte la diérèse ou la synérèse, d'autres ne donnent jamais ou quasiment jamais lieu à diérèse. Ainsi en est-il des graphies *oi* et *oin*, qui, aujourd'hui, se réalisent phonétiquement par /w a/ et /w ɛ̃/ et de *ui* (/ɥi/), qui ne conduit à une diérèse, dans le vers classique, que dans les mots *assiduité* et *ruine*. *Ruine* est d'ailleurs orthographié dans les éditions d'époque en *ruïne* (le tréma étant un indice graphique de la diérèse) :

> Je sçay, que j'ay moy seule avancé leur ruïne,
>
> Racine, *Andromaque*, éd. P. Mélèse, vers 60.

Par ailleurs, placées à la rime, les diphtongues *oi* (respectivement *oin*) ne peuvent jamais rimer avec des finales autres que *oi* (ou *oin*). En dehors de rares exceptions (moins de vingt cas), les rimes dont le noyau vocalique est en *ui* riment avec d'autres mots en *ui*. Ces groupements vocaliques constituent des

unités insécables à la rime. Marc Klein [1991, p. 1-15] les identifiait comme des *diphtongues de fortes cohésion*. Elles ont comme particularité de violer *OLG, la règle qui interdit les séquences obstruantes-liquides-glissantes (semi-voyelles). Cette règle, qui est une reformulation de la loi des trois consonnes, pose que dans une séquence OLG, la glissante acquiert sa valeur vocalique pour éviter la succession de trois consonnes. Dans les diphtongues à forte cohésion cela n'a pas lieu : *fruit, droit, groin* constituent chacun une syllabe.

Dans ces trois cas, nous avons considéré que la semi-voyelle était partie intégrante du noyau vocalique. Quant aux autres cas de glissante (semi-voyelle) devant voyelle, nous ne les avons pas rattachés à la voyelle, car il y a vraiment alternance entre voyelle haute et semi-voyelle. Ainsi pour *lui*, *moi*, *point* les noyaux vocaliques sont respectivement /ɥ i/, /w a/, /w ɛ̃/, tandis que pour *lion*, *lier*, *louer*, les noyaux sont /i/-/ɔ/, /i/-/e/, /u/-/e/,

Parce que le *e* muet a un rôle métrique tout à fait fondamental dans le vers, nous en avons distingué deux types (*cf.* chapitre IV.2) :

- le *e* muet fort qui correspond au *e* des monosyllabes et aux *e* muets internes ou finaux qui, s'ils étaient élidés, créeraient une suite de trois consonnes. Par exemple, les *e* de *je*, *te*, *le*, *contre*, *brelans* sont forts ;

- le *e* muet faible qui correspond aux *e* en fin de polysyllabes ou aux *e* intérieurs si leur élision ne crée pas une suite de trois consonnes. Par exemple, dans les mots suivants, *Rom<u>e</u>, fê<u>te</u>, dang<u>e</u>reux, écout<u>e</u>rait*... les *e* soulignés sont faibles.

> Rome s'en fait d'avance en l'esprit une fête,
> (r o)(m əf)(s ã)(f ɛ)(d a)(v ã s)(ã)(l ɛ s)(p r i tl)(y)(n əf)(f ɛ t)
>
> Corneille, *Tite et Bérénice*, vers 8.

> Le dangereux soupçon de n'être pas aimé
> (l ə)(d ã)(ʒ əf)(r ø)(s u p)(s õ)(d ə)(n ɛ)(t r ə)(p ɑ zl)(e)(m e)
>
> Corneille, *Tite et Bérénice*, vers 18.

Les *e* de *avance*, *fête* sont élidés, le premier parce qu'il est suivi d'une voyelle, le second parce qu'il est en fin de vers. Les *e* de *Rome*, *une* sont faibles car ils correspondent à la fin d'un polysyllabe, le *e* de *dangereux* est faible parce qu'il est interne à un polysyllabe et que son élision n'entraîne pas une

suite de trois consonnes. Les *e* de *le* et *de* sont forts parce que ce sont des monosyllabes, celui de *être* aussi parce que son élision créerait une suite de trois consonnes.

Enfin, l'étude de la rime (*cf.* chapitre III et surtout l'annexe XI) montre que l'opposition entre *a* antérieur et *a* postérieur, au sens phonologique, n'est pas pertinente dans l'alexandrin classique, puisque que les deux peuvent rimer ensemble comme dans les vers suivants :

Fût-il la valeur même, et le dieu des combats,	/a/
Il verra ce que c'est que de n'obéir pas.	/ɑ/

Corneille, *Le Cid*, vers 567-568.

La peur d'un vain remords trouble cette grande âme :	/ɑ/
Elle flotte, elle hésite ; en un mot, elle est femme.	/a/

Racine, *Athalie*, vers 875-876.

En revanche, la distinction entre *o* ouvert et fermé s'avère pertinente, puisque l'on ne trouve pas de cas de *o* ouvert rimant avec un fermé *o*. Ainsi *trône* ne peut-il rimer avec *donne*.

Ces différentes raisons nous ont donc conduit à adopter les noyaux vocaliques présentés dans le Tableau 20. Ils sont cohérents avec la structure des vers et constituent les bases pour l'identification de figures phoniques. En effet, si *ui* ne rime que rarement avec *i*, c'est sans doute que les deux sons ne peuvent pas être facilement reconnus comme équivalents. Par conséquent, si l'on veut étudier les assonances, il sera peut-être peu fondé de considérer *i* et *ui* comme formant une assonance. C'est l'étude de la rime, et dans une moindre mesure de la diérèse, qui nous a permis de trancher sur les phonèmes ou groupements de phonèmes qui sont équivalents ou différents. Nous voyons ici encore, en passant, que l'approche empirique sur gros corpus peut aboutir à des aménagements assez importants des formalismes usuels.

Pour conclure sur les aspects phonétiques, le marquage syllabique donne la transcription phonétique de chaque position et le marquage vocalique extrait de chaque syllabe le noyau vocalique. Le marquage syllabique est bien évidemment dépendant des choix de découpage syllabique adoptés (dans notre

cas, non enchaînement des consonnes finales de mot devant voyelle, et enchaînement pour les consonnes latentes en cas de liaison forte) et des choix de représentation de certains phénomènes comme la diérèse (dans notre cas, pas de semi-voyelle dans les situations de diérèse). Le marquage vocalique dépend des types de noyaux vocaliques retenus. Nous avons retenu les dix-neuf qui figurent dans le Tableau 20 et sont assez différents des seize voyelles du français généralement identifiées. L'analyse du vers (rime et diérèse) nous a conduit à réexaminer la catégorisation classique.

V.2. MARQUAGE DES FINS DE MOT

Nous venons de décrire les marquages phonétiques. Nous allons maintenant nous intéresser aux marquages lexicaux. Dans la séquence continue que définit le vers, le mot est l'unité la plus évidente de découpage grâce à la présence de frontières de mots comme l'espace ou les signes de ponctuation.

Par ailleurs, les expériences d'Anne Christophe [Christophe, 1993 ; Mehler et Christophe, 1995] sur le rôle de la prosodie dans la reconnaissance des mots montrent que des adultes et des bébés sont capables de distinguer un bisyllabe (de type CVCV) contenant une frontière de mot (stimuli inter), d'un bisyllabe n'en contenant pas (stimuli intra)[104]. La présence d'une frontière de mots entre les deux syllabes entraîne un allongement de la première voyelle (celle qui se trouve avant la frontière) et un allongement de la consonne initiale de la syllabe suivante. La durée entre les deux voyelles est significativement plus élevée quand il y a une frontière entre les deux syllabes. Dans ces expériences, les

[104] Le protocole est le suivant : pour des bisyllabes donnés, comme *mati* ou *mãta*, ont été recherchés des mots qui les contenaient (par exemple mathématicien) et des paires de mots tels que le bisyllabe soit à cheval sur les deux mots (puma timoré). Des locutrices qui ne connaissaient pas le but de l'expérience ont lu des phrases porteuses contenant tantôt le mot, tantôt le couple de mots. Ensuite, le bisyllabe a été découpé dans le signal. Ce sont ces bisyllabes qui sont soumis aux sujets de l'expérimentation.

Les adultes, en général, ne connaissent pas le but de l'expérience et doivent uniquement classer dans deux catégories la production phonique des bisyllabes. La perception du changement est mesurée chez les bébés par le taux de succion. Les bébés perçoivent le changement quand après avoir émis des stimuli inter (bisyllabes à cheval sur deux mots), des stimuli intra (bisyllabes internes à un mot) sont produits.

indices prosodiques de fins de mots n'apparaissent pas systématiquement au cas par cas, mais les différences sont significatives sur les échantillons analysés. En français, l'accent porte sur la dernière position du mot. Il est difficile de savoir si les sujets reconnaissent une frontière de mot ou un accent. L'allongement de la consonne initiale après la frontière de mot a été repérée dans d'autres langues. De plus, le même protocole appliqué à la langue espagnole, où l'accent n'est pas en général sur la dernière position, conduit à des résultats similaires : distinction des stimuli intra et inter. Il semble bien exister des indices prosodiques de frontière de mot indépendants de l'accent.

Il aurait été intéressant que des expériences complémentaires soient menées. En particulier, que d'autres contextes syntaxiques soient explorés : pour les stimuli intra, seuls des noms ont été choisis et pour les stimuli inter des couples nom, adjectif. Par ailleurs, pour mieux différencier les effets de l'accent de mot et de la fin de mot, il aurait été intéressant de choisir des contextes avec un *e* muet à la fin du premier mot.

Quoiqu'il en soit, nous avons donc marqué la dernière syllabe métrique de chaque mot. Dans la tradition des marquages métrico-rythmiques de Lusson [1974] et de Roubaud [*1978*, 1988] qui se prêtent au traitement mathématique, le codage a été fait de manière binaire : 1 marque la syllabe métrique finale d'un mot, 0 les syllabes initiales et internes.

De cette nuit, Phénice, as-tu vu la splendeur ?

Racine, *Bérénice*, vers 302.

syllabes métriques	d ə	s ɛ	t əf	n ɥ i	f e	n i s	a	t y	v y	l a	sp l ɑ̃	d œ r
fins de mots	1	0	1	1	0	1	1	1	1	1	0	1

D'un inutile amour trop constante victime

Racine, *Bérénice*, vers 255.

sm	d œ̃	n i	n y	t i l	a	m u r	t r o	k ɔ̃ s	t ɑ̃	t əf	v i k	t i m
fdm	1	0	0	1	0	1	1	0	0	1	0	1

La difficulté majeure provient des noms composés, ou locutions, que les analyseurs syntaxiques un peu sophistiqués sont capables de reconnaître grâce à des dictionnaires de locutions et/ou à des règles syntaxiques. Doit-on considérer pour le marquage de fin de mot que la locution constitue un seul mot, ou autant

de mots que ceux qui la composent ? C'est cette dernière option que nous avons choisie, d'une part car dans la langue du XVIIe siècle les mots composés sont rares, y compris les locutions grammaticales, et d'autre part car le métromètre avait tendance à en repérer là où il n'y en avait pas. Par exemple, dans le vers suivant, *bien* est analysé à tort par le métromètre comme partie d'une locution conjonctive *bien que* au lieu d'être analysé comme nom :

> D'aspirer à ce bien que vous m'avez promis
>
> Corneille, *Mélite*, vers 1786.

Si nous n'avions pas forcé le métromètre à « traverser les locutions », pour beaucoup de vers, ces erreurs auraient entraîné l'absence d'un marquage de fin de mot en sixième position.

Dans ce marquage, ce sont les espaces et les signes de ponctuations qui permettent de repérer les frontières de mots. Seule l'apostrophe n'est pas considérée comme un séparateur. Ainsi, dans le décasyllabe suivant, *puisqu'aujourd'hui* ne constitue qu'un seul mot.

> Puisqu'aujourd'hui mon père est l'offensé,
>
> Corneille, *Le Cid*, vers 349

sm	p ɥ i s	k ɔ	ʒ u r	d ɥ i	m ɔ̃	p ɛ r	ɛ	l ɔ	f ɑ̃	s e
fdm	0	0	0	1	1	1	1	0	0	1

Ce choix est discutable, mais l'inverse l'est aussi.

Le marquage de fin de mot adopté, nous permettra donc d'identifier les positions finales de chaque mot dans le vers, qu'ils se finissent par un *e* muet ou par tout autre noyau vocalique, afin d'étudier la répartition des fin de mots sur les douze positions, du nombre de mots par vers et de repérer des vers extrêmes par rapport à ce marquage (vers composés de très peu ou de beaucoup de mots).

V.3. NATURE SYNTAXIQUE DES POSITIONS

Après avoir abordé les marquages phonétiques et de fin de mot, nous en venons maintenant aux marquages syntaxiques.

L'analyse syntaxique du vers, comme nous l'avons vu, intervient en premier lieu au moment de l'identification des positions métriques : elle permet de distinguer les homographes hétérophones, comme *content* (adj) /*content*

(verbe) ; *fier* (adjectif) / *fi-er* (verbe), et donc d'aboutir à un dénombrement correct, et elle intervient également pour déterminer l'intensité de la liaison.

Son rôle cependant ne s'arrête pas là. La structure syntaxique peut également être mobilisée pour l'analyse du rythme. En effet, la syntaxe configure en partie la prosodie. Les syntagmes coïncident le plus souvent, aux *e* muets finaux près, avec les groupes de souffle ou mots phonologiques, qui sont les composantes minimales du rythme. Ensuite, la structure syntaxique peut également permettre de traiter la question des relations entre mètre et syntaxe et en particulier les cas de discordance entre vers et phrase. Par exemple, dans le vers suivant déjà cité, le mot phonologique *venger un père* enjambe l'hémistiche :

Qu'ils semblent tous / venger un père, / comme vous.

Corneille, *Cinna*, vers 152.

La fin de segment métrique ne coïncide pas avec une fin de syntagme, ce qui est assez rare et mérite donc d'être observé.

V.3.1. Quels marquages syntaxiques pourraient être envisagés ?

Pour disposer d'un marquage syntaxique fin, qui rende compte de la structure des différents niveaux d'organisation syntaxique (phrases, syntagmes, groupes, mots...), il aurait été souhaitable d'avoir une représentation hiérarchisée, sous forme d'arbre, de la structure de la phrase inscrite dans le ou les vers. Il aurait été aisé de déduire de cet arbre des marquages hiérarchisés tels qu'ils sont envisagés par Lusson et Roubaud, à savoir des marquages qui définissent des groupements d'événements élémentaires (positions métriques, mots, syntagmes...) enchâssés. Le poids de chaque position dépend du nombre de parenthèses fermantes dont la position est suivie.

(((L')(effet)) (((montra) (soudain)) ((ce) (conseil) (salutaire)))).
1 2 0 1 0 2 1 0 1 0 0 4

Corneille, *Rodogune*, vers 52.

.

A supposer qu'il soit possible de construire cette représentation syntaxique hiérarchisée, les marquages suivants pourraient être mis en place :

- un marquage de phrase qui consisterait à indiquer la position des fins de phrase, ce qui nous permettrait, entre autres, de mesurer la coïncidence entre phrase et vers, comme l'a fait en partie Garrette [1995] sur la phrase de Racine ;
- un marquage de fin de syntagme qui permettrait d'évaluer les phénomènes d'enjambement de vers à vers, d'hémistiche à hémistiche et de dresser une typologie des relations entre syntagmes et segments métriques ;
- un marquage de catégorie syntaxique des mots ;
- un marquage syntaxique global reflétant la structure en arbre, où chaque position métrique hériterait du poids lié à sa position syntaxique.

Ces marquages sont chacun porteurs de difficultés théoriques et pratiques. Examinons à présent celles qui sont posées par chacun de ces marquages.

Un marquage de fin de phrase implique que l'on soit capable d'identifier clairement les phrases. Or la notion de phrase n'est pas simple : Garrette, par exemple, utilise tout au long de son ouvrage deux définitions de la phrase (phrase grammaticale et phrase stylistique) indiquant par là qu'il n'y a pas de définition unique. Deux marquages auraient pu être adoptés pour rendre compte des deux définitions. Nous verrons dans le chapitre VI concernant les limites du métromètre pourquoi ce marquage n'a pas été introduit : l'unité d'analyse du métromètre est en effet le vers et non la phrase. En raison de ce choix, nous n'aurions pu introduire un marquage de phrase grammaticale (unité de prédication), mais en revanche nous aurions pu introduire un marquage de phrase stylistique (unité d'énonciation) à partir de la ponctuation.

L'identification des syntagmes, qui conduirait au marquage de leurs frontières, elle non plus, ne va pas de soi, surtout dans le vers classique où l'ordre « naturel » n'est souvent pas respecté. Toujours dans le chapitre sur les limites du métromètre, nous verrons comment le phénomène de la métaposition (inversion de certains syntagmes) rend difficile l'identification systématique des syntagmes. Notons que le vers, en raison des distorsions qu'il inflige à la syntaxe, est un terrain particulièrement difficile pour les analyseurs syntaxiques

actuels. Celui que nous utilisons est un des meilleurs de la génération actuelle, pourtant il est souvent mis en défaut par le vers.

Le troisième marquage correspond à l'affectation de la catégorie syntaxique aux mots. Il est loin d'être évident car il dépend de la théorie grammaticale mobilisée. La syntaxe étant une discipline en évolution, les appellations et les contours des catégories changent. Par ailleurs, quelle que soit la théorie, les frontières entre les catégories ne sont pas toujours très nettes, comme entre un adjectif et un participe passé.

Examinons la contribution de Milner à la question des catégories syntaxiques dans *Introduction à une science du langage* [1989]. Il propose dans cet ouvrage les grandes lignes de l'architecture d'une nouvelle théorie du langage. Celle-ci serait constituée d'une part d'une théorie du lexique (théorie des termes), de l'autre d'une théorie de la syntaxe (théorie des positions) qui toutes deux contribuent à l'élaboration de la sémantique.

La théorie des termes doit définir la constitution d'un terme, qui est « l'ensemble de toutes ses propriétés distinctives et n'est que l'ensemble de ses propriétés distinctives. » [Milner, p. 324]. Pour chaque terme, il existe trois types de caractères qui peuvent entrer dans sa définition : 1) l'appartenance à une catégorie, 2) la forme phonologique, 3) la signification lexicale.

Milner opte pour une théorie positionnelle de la syntaxe. Celle-ci combine la notion de « sites » (la place occupée par le terme) et celle de « positions syntaxiques ». Chaque site-position a une étiquette catégorielle qui n'est pas forcément la même que la catégorie d'appartenance du terme qui occupe le site.

> Notre terminologie tente d'accentuer cette différence : quand il s'agit des sites-positions, elle parle d'*étiquetage* ; quand il s'agit des termes, elle parle d'*appartenance*.
>
> [Milner, 1989, p. 358].

Autrement dit, un terme hors contexte a une appartenance catégorielle donnée, un terme en contexte acquiert l'étiquette catégorielle du site occupé. Appartenance et étiquette puisent dans le même ensemble, les parties du discours, mais ne coïncident pas forcément.

Tesnière avait l'intuition de cette distinction, en développant l'idée de translation. Pour lui :

> La translation consiste donc à transférer un mot plein d'une catégorie grammaticale dans une autre catégorie grammaticale, c'est-à-dire à transformer une espèce de mot en une autre espèce de mot.
>
> [Tesnière, 1959, 1982, p. 364].

Ainsi dans *Le train de Paris*, la translation est l'opération qui transforme le nom *Paris* en adjectif grâce à la préposition *de*. Une partie des mots vides et des morphèmes grammaticaux ont pour fonction d'assurer ces translations.

On peut considérer que la translation est l'opération par laquelle un terme ayant une catégorie d'appartenance donnée acquiert les propriétés nécessaires pour occuper une position syntaxique spécifique dont l'étiquette ne correspond pas forcément à la catégorie du terme.

Si, théoriquement, la distinction entre catégorie d'appartenance et étiquette catégorielle devrait résoudre en partie les problèmes de cohérence dans l'étiquetage syntaxique, dans la pratique, les outils de marquage syntaxiques confondent communément les deux dimensions. Nous avons retenu un marquage de catégorie syntaxique, dont la qualité dépend de l'analyseur syntaxique utilisé.

Il nous reste enfin à examiner les conditions de possibilité du quatrième marquage, un marquage qui rendrait compte de la structure syntaxique globale. Quelque soit le type de théorie syntaxique utilisé pour représenter la structure de la phrase (représentation arborée ou structure des liens entre constituants syntaxiques), il faudrait pouvoir constituer à partir de la représentation syntaxique de la phrase un marquage syntaxique global qui rende compte de cette structure. Il faudrait, en particulier, pouvoir représenter des phénomènes d'intrication au sens large (enchâssement, concaténation, imbrication). Ces phénomènes, interprétés à l'époque chomskienne comme des effets des règles transformationnnelles qui s'appliquaient à la structure profonde de la phrase, ont donné lieu à des recherches en linguistique mathématique dont Roubaud rend compte dans la *Description du projet* [1979, p. 18-22]. Ils ont conduit à des analyses spécifiques dans le cadre de certaines lectures rythmiques, comme celle de trois sonnets d'Etienne Jodelle [Lusson et Roubaud, 1980 et 1981] :

Des astres, des forests, et d'Achéron l'honneur (situation d'enchâssement)

Luy moi, prend moy, tien moy, mais hélas ne me perd (situation de concaténation)

Ce type de transformation permet de rendre compte élégamment de la particularité des vers rapportés, comme le vers suivant :

Lune, Diane, Hecate, aux cieux, terre et enfer (situation d'imbrication : vers rapporté)

Ainsi, Lusson et Roubaud proposent de considérer que « l'opération caractéristique des « vers rapportés » est une *imbrication* qui transforme les peignes, initialement concaténés en peignes imbriqués » [1981, p. 50].

La formalisation de ces phénomènes d'intrication est particulièrement complexe. En l'absence d'une formalisation avancée, un marquage syntaxique global reflétant la structure de la phrase est difficilement envisageable pour toutes les phrases dans lesquelles des règles de transformation ou des phénomènes d'intrication modifient l'arbre syntaxique.

V.3.2. Le marquage syntaxique choisi

Nous venons d'envisager un certain nombre de marquages syntaxiques et les difficultés que chacun pose. Pour des raisons que nous présenterons dans la chapitre consacré au métromètre et à ses limites (*cf.* chapitre VI), seul le marquage des catégories syntaxiques a pu être mis en œuvre. Les catégories, imposées par l'analyseur syntaxique Sylex, sont critiquables, leurs frontières particulièrement floues, mais l'analyse a le mérite d'être à peu près cohérente de bout en bout[105].

[105] Un système de près de cent traits syntaxiques permet dans Sylex de raffiner cette typologie, en particulier de distinguer, quand cela est possible, les adjectifs des participes passés, les déterminants des pronoms... Nous n'avons pu les utiliser de manière satisfaisante dans la version actuelle du métromètre.

L'analyseur Sylex distingue neuf catégories syntaxiques, qui peuvent être regroupées en deux grandes catégories : les mots pleins (ou mots lexicaux) et les mots vides (ou mots-outils, mots grammaticaux...).

> Les mots pleins sont ceux qui sont chargés d'une fonction sémantique, c'est-à-dire ceux dont la forme est associée directement à une idée, qu'elle a pour fonction de représenter et d'évoquer.
>
> (...) Les mots vides (...) ne sont pas chargés d'une fonction sémantique. Ce sont de simples outils grammaticaux dont le rôle est uniquement d'indiquer, de préciser ou de transformer la catégorie des mots pleins et de régler leurs rapports entre eux.
>
> [Tesnière, 1959, 1982, p. 53].

Dans les mots pleins, on distingue les substantifs, les noms propres, les adjectifs, les verbes et les adverbes. Les mots vides comprennent les sous-catégories suivantes : déterminants et pronoms, prépositions, conjonctions et relatifs.

> De cette nuit, Phénice, as-tu vu la splendeur ?
>
> Racine, *Bérénice*, vers 302.

syllabes métriques	d ə	s ɛ	t əf	n ɥ i	f e	n i s	ɑ	t y	v y	l ɑ	s p l ɑ̃	d œ r
catégories syntaxiques	7	6	6	0	4	4	1	6	2	6	0	0

Catégories syntaxiques

0 = Nom	6 = Déterminant, pronom
1 = Verbe	7 = Préposition
2 = Adjectif	8 = Conjonction
3 = Adverbe	9 = Pronom relatif
4 = Nom propre	

Quelques choix de catégorisation adoptés par l'analyseur Sylex méritent d'être explicités. Tout d'abord, aucune distinction n'est faite entre les déterminants et les pronoms. Ainsi, *le* article et *le* pronom seront confondus. Ces deux catégories n'ont pas été distinguées car on y trouve une forte proportion d'homonymes (*le, la, les* sont tantôt des pronoms, tantôt des déterminants). Ensuite, la frontière entre participe passé et adjectif pose de nombreuses difficultés en français : l'option de l'analyseur syntaxique Sylex est de considérer tous les participes passés comme des adjectifs (exemple : *il est mort*) et non comme des verbes. Dans certains cas, cela est étrange. Ainsi, dans le vers suivant *conduit* sera-t-il catégorisé comme un adjectif :

Ce chemin à l'empire a conduit de tout temps,

Racine, *Bajazet*, vers 444.

Toutes les positions qui correspondent à un mot donné reçoivent comme marquage la catégorie syntaxique du mot. Ainsi, les trois dernières positions du vers suivant qui correspondent au mot *empereur* acquièrent la catégorie de nom (code = 0) :

J'ambitionne et crains l'hymen d'un empereur

sm	ʒ ɑ̃	b i	s i	ɔ n	e	k r ɛ̃	l i	m ɛ n	d œ̃	ɲ ɑ̃	p əf	r œ r
cs	1	1	1	1	8	1	0	0	6	0	0	0

Dont j'ai lieu de douter si j'aurai tout le cœur,

sm	d ɔ̃	ʒ ɛ	l j ø	d ə	d u	t e	s i	ʒ ɔ	r ɛ	t u	l ə	k œ r
cs	9	1	0	7	1	1	9	1	1	6	6	0

Corneille, *Tite et Bérénice*, vers 25-26.

Quand une position métrique est constituée d'éléments provenant de deux mots, elle acquiert la catégorie du mot auquel appartient le noyau vocalique. Par exemple, dans *j'ambitionne, j'ai* et *j'aurai*, la première position (ʒ ɑ̃), (ʒ ɛ)et (ʒ o), (qui est en partie un article, en partie un verbe, aura la catégorie de verbe. Cela conduit à une légère sous-évaluation des mots grammaticaux.

A partir du marquage syntaxique de chaque position, on pourra construire la figure de répartition d'une catégorie syntaxique donnée, et celle d'un regroupement de catégories, par exemple celle des mots-outils.

Nous avons retenu les catégories syntaxiques de base de l'analyseur syntaxique Sylex, mais celui-ci offre un système de plus de cent traits syntaxiques qui aurait permis de raffiner cette typologie (un trait en particulier permet de distinguer quand cela est possible l'adjectif du participe passé).

V.4. MARQUAGE "ACCENTUEL"

Nous avons défini les marquages permettant d'étudier la répartition des phonèmes vocaliques, des fins de mots et des catégories syntaxiques. Il nous reste à introduire des marquages pour l'étude des aspects prosodiques.

La prosodie nous renvoie immédiatement à la question de l'accent. Quel est le statut de l'accent en français ? Quelle est sa nature ? Quel rapport entretient-il avec le mot phonologique ? Comment se distribue-t-il par rapport aux constituants syntaxiques ? Autrement dit, comment coder les phénomènes accentuels à partir d'un texte écrit analysé syntaxiquement ?

Comme nous l'avons vu en introduction à ce chapitre, nous avons fait le choix de ne projeter sur la séquence des positions métriques que des marquages de langue. Nous faisons, une fois ces positions identifiées, comme si nos textes étaient en prose, sans faire l'hypothèse de l'existence d'un accent de vers autonome.

Que sait-on sur les aspects accentuels et prosodiques de la langue ? La prosodie désigne l'étude de phénomènes linguistiques ayant des corrélats acoustiques assez hétérogènes : la hauteur, l'intensité et la durée, qui sont « les composantes phoniques du rythme accentuel » [Meschonnic, 1982, p. 217]. Le point commun de ces phénomènes est qu'ils sont suprasegmentaux et que leur rattachement aux unités discrètes de la langue (phonème, syllabe, mot) n'est pas trivial. Pour la prose, F. Dell [1984] propose d'intégrer les données prosodiques à la représentation phonétique d'une phrase. Dans son formalisme, celle-ci contient la « chaîne syllabique, qui spécifie une suite de syllabes » et le « profil prosodique ».

> J'appelle [profil prosodique] un ensemble de trois représentations parallèles : le profil mélodique, qui spécifie la hauteur mélodique en chaque point de la chaîne syllabique, et les profils de durée et d'intensité, qui en font respectivement autant pour la durée et l'intensité.
>
> [Dell, 1984, p. 65].

L'appariement entre la chaîne syllabique et les trois dimensions du profil prosodique est loin d'être évidente. Dell rappelle, suite aux études de Liberman [1975], que le profil mélodique se déduit à partir de son motif, les motifs étant des stéréotypes mélodiques abstraits. Pour apparier chaîne syllabique et motif, il faut s'appuyer sur les syllabes pivots. La thèse de Dell consiste à considérer que :

> La description raisonnée du profil mélodique des phrases françaises requiert celle préalable de leur schéma accentuel, car les critères qui déterminent la position des syllabes clefs au sein des tronçons sont d'ordre exclusivement accentuel.
>
> [Dell, 1984, p. 67].

Pour Dell, l'accent, ou plutôt le schéma accentuel, est le fondement du profil prosodique en français.

Nous examinerons d'abord l'existence et la nature de ce schéma accentuel. Ensuite, nous présenterons le marquage « accentuel » que nous avons choisi, qui n'est qu'une approximation très insatisfaisante de ce que pourrait être une modélisation du schéma accentuel. Ce marquage accentuel, tout insatisfaisant qu'il soit, semble être une piste intéressante pour appréhender la question du rythme. Enfin, nous proposerons une approche alternative de ce marquage, qui s'appuie sur la métrique.

V.4.1. Existe-t-il un accent en français ?

L'existence même d'un système accentuel en français, les différents types d'accents que celui-ci distingue et les règles d'accentuation sont des sujets très controversés.

Si l'on revient aux origines de la langue française (Cerquiglini [1995, p. 11-36]), il semble bien que la particularité de l'accent en français tienne à une double influence : l'accent est importé du latin mais il se voit renforcé par l'influence germanique qui en fait un accent d'intensité. L'influence germanique a eu, entre autres effets, celui d'affaiblir les voyelles atones, en particulier les voyelles finales qui deviennent peu à peu des *e* muets. Il serait étonnant, alors qu'il a contribué à donner aux mots leur forme actuelle, que l'accent n'ait plus sa place dans l'analyse de la prosodie.

La tradition linguistique reconnaît donc, en langue, l'existence d'un accent de mot en français situé sur la dernière syllabe pleine des mots lexicaux, à savoir les noms, les verbes, les adverbes et les adjectifs [Vaissière, 1991]. Si la dernière syllabe est occupée par un *e* muet, la syllabe pleine est l'avant-

dernière. Ainsi dans les mots suivants, nous avons indiqué en gras la syllabe accentuée : para**dis**, paro**die**, pa**ra**de. Dell [1984] introduit la notion de mot (phonétiquement) féminin (polysyllabe dont le noyau de la dernière syllabe est un *e* muet) et de mot masculin : « traditionnellement, on dit que les mots féminins sont accentués sur l'avant-dernière syllabe et que les mots masculins sont accentués sur la dernière » [Dell, 1984, p. 68].

Dell ne fait pas de distinction entre mots-outils et mots pleins pour l'accent, même si dans ses exemples il apparaît que les mots grammaticaux ne sont pas accentués.

Pourquoi les mots vides, grammaticaux ou fonctionnels ne portent-ils pas d'accent ? Ce fait semble commun à plusieurs langues, comme le confirment les observations de Zipf [1934].

L'examen empirique de corpus de textes anglais et allemands l'amène à constater que « le degré d'accentuation des mots tend à être en relation inverse de la fréquence relative de leurs occurrences » [Zipf, 1934, p. 110]. Plus précisément, il pose :

> L'existence d'un état d'équilibre dans le système accentuel d'une langue. Dans cet état, il semble que l'on puisse isoler quatre facteurs déterminants :
> - la fréquence relative (F)
> - le degré d'intensité de l'accent (A)
> - le degré de cristallisation de la configuration C)
> - le degré de précision de la signification (M)
>
> Le rapport entre ces quatre facteurs semble être le suivant ; A et M sont en relation inverse de F et C.
>
> [Zipf, 1934, p. 130].

Zipf considère que « la fréquence relative est le facteur déterminant dont semblent dépendre l'accent, la signification et la cristallisation » [p. 131].

Pour dire vite, plus un mot est fréquent moins sa signification est précise et moins il a de chances de porter l'accent. Or dans tous les corpus de textes, les mots-outils sont les mots les plus fréquents. Ils sont donc les moins susceptibles de porter l'accent.

Pour le français, mais c'est sans doute vrai aussi pour les langues étudiées par Zipf, ces observations refléteraient plutôt des effets de structure. En français en tout cas, l'accent est un accent de *mot phonologique*, comme nous allons le

voir, qui tombe en fin de syntagme. Or il est rare qu'une fin de syntagme soit occupée par un mot outil. De par leurs positions dans les syntagmes, les mots-outils sont rarement accentués.

L'existence d'un accent de mot est généralement admise. Mais en discours, cet accent de mot est confronté à l'accent de mot phonologique, qui est le seul accent admis par la grande majorité des linguistes.

En effet, on considère généralement que l'accent en français est principalement un accent de *groupe de souffle* [Vaissière, 1991], ou un accent de *mot phonologique*[106] pour reprendre la terminologie de Milner et Regnault. C'est la dernière syllabe du groupe de souffle ou du mot phonologique qui porte l'accent. L'accent de mot tend à s'affaiblir si le mot est situé à l'intérieur du mot phonologique : « une syllabe accentuable est désaccentuée si elle est à l'intérieur d'un mot phonologique » [Milner et Regnault, 1987, p. 82].

L'effacement de l'accent dans le groupe est une spécificité du français absente de langues comme l'espagnol, l'anglais...

Plutôt que de considérer que les accents de mots à l'intérieur du mot phonologique s'effacent, on peut faire l'hypothèse qu'il reste une trace de ces accents de mots dans le schéma accentuel. C'est pour rendre compte de ce phénomène que Dell [1984, p72] introduit la notion de « degré d'accent ». Celle-ci conduit à une hiérarchisation des accents à l'intérieur du *tronçon*, « toute portion maximale de l'énoncé qui ne contient ni silence, ni rupture intonationnelle » [Dell, 1984, p. 68]. L'accent de mot est plus ou moins fort selon sa position syntaxique à l'intérieur du tronçon. « En français, chaque mot porte un accent dont la force est proportionnelle à l'importance de la coupe syntaxique qui suit ce mot » [Dell, 1984, p. 77]. Dell pose donc l'hypothèse d'une congruence entre la structure syntaxique et le schème accentuel, puisque l'intensité de l'accent de mot dépend de sa position syntaxique dans l'unité syntaxique que représente le tronçon.

---

[106] En général, un mot phonologique correspond à un groupe syntaxique majeur. Pour Milner et Regnault [1987, p. 29] est mot phonologique « ce qui constitue *à la fois* une unité de souffle, une unité grammaticale et une unité de signification ».

Bien entendu cette hypothèse sur la hiérarchisation des accents devrait être validée par l'étude de signaux acoustiques, en mesurant en chaque point (ou syllabe) hauteur, intensité et durée des syllabes. En effet, les trois indicateurs devraient être mesurés car il semble bien qu'une des spécificités de l'accent en français soit la diversité de ses modes de réalisation : la manifestation acoustique de l'accent « consiste la plupart du temps à augmenter l'intensité, ou la hauteur, ou la longueur de la voyelle qui constitue le centre de la syllabe accentuable » [Milner et Regnault, 1987, p. 79].

Dell propose en particulier d'observer les jugements de proéminence relative des syllabes. Pour lui, « ils restent invariants lorsqu'on fait varier le profil mélodique » [Dell, 1984, p. 74], c'est-à-dire l'intonation sur l'ensemble de l'énoncé (mélodie déclarative, interrogative, exclamative...).

Pour le moment, nous avons envisagé deux types d'accents, celui de fin de mot, l'accent lexical hors contexte, ou celui de fin de mot phonologique. Or, certains auteurs supposent l'existence d'un autre accent qui apparaîtrait sur la première syllabe d'un polysyllabe. Milner et Regnault [1987, p. 81-85] le désignent comme contre-accent. Pour eux, c'est également un accent de mot phonologique. Il disparaît quand il porte sur une syllabe interne à un mot phonologique. Cet accent est reconnu comme moins fort que l'accent de fin de mot, ce que devrait montrer le signal acoustique.

D'après Dell, l'intensité de l'accent est liée à la place de la syllabe à l'intérieur d'un tronçon intonatif, c'est-à-dire à la nature de la frontière syntaxique qui le suit. Mais comment délimiter ces tronçons ? Dans la séquence de graphèmes, une seule catégorie de signes peut nous y aider, ce sont les signes de ponctuation. Ils sont l'unique marque de la prosodie marquée par la graphie. Ils sont des indices du type d'intonation de phrase (assertive, interrogative ou exclamative), mais ils délimitent également des pauses, plus ou moins longues, dans le flux de parole, qui doivent coïncider avec des tronçons. Le rôle d'une analyse syntaxique complète serait de nous indiquer la force des frontières de groupements syntaxiques, ce qui nous permettrait ensuite de construire la hiérarchie des accents.

V.4.2. Marquages prosodiques dans le vers

Nous avons opté pour un marquage de « l'accent » de mot extrêmement simple. Il s'agit à peu de choses près de la « marque fondamentale » proposée par Roubaud [1988] dans la seconde partie du programme Dynastie.

> Définition : sont notés 1 pour la marque fondamentale :
> a) les dernières voyelles métriques pleines des mots de catégorie syntaxique majeure, c'est-à-dire verbes, substantifs, adjectifs et adverbes (mais ni les auxiliaires, ni les adjectifs possessifs ou numéraux).
> b) Les dernières voyelles pleines des groupements syntaxiques.
> (Les espèces a) et b) sont évidemment fortement corrélées)
>
> [Roubaud, 1988, p. 48].

Le marquage est binaire en « 0 » ou « 1 » : la syllabe qui porte l'accent de mot est marquée par un « 1 », les autres par « 0 ».

Le système n'implémente que la première partie du programme, la seconde nécessitant une analyse syntaxique satisfaisante. Est ainsi marquée comme accentuée la dernière syllabe pleine des mots appartenant aux catégories syntaxiques suivantes : substantif, nom propre, verbe, adjectif et adverbe. Autrement dit, l'accent porte sur la dernière position métrique sauf si le noyau vocalique est un *e* muet, dans ce cas, elle porte sur l'avant dernière. Deux situations fréquentes ont donné lieu à des traitements particuliers :

- Les auxiliaires n'ont pas été considérés comme des mots pleins ;

- Quand le pronom personnel est placé après le verbe, en position d'enclitique, relié à lui par un tiret, on considère que verbe et pronom ne forment qu'un mot au regard de l'accent.

On trouvera ci-dessous un exemple de marquage accentuel.

> De cette nuit, Phénice, as-tu vu la splendeur ?
>
> Racine, *Bérénice*, vers 302

syllabes métriques	d ə	s ɛ	t əf	n ɥ i	f e	n i s	a	t y	v y	l a	sp l ɑ̃	d œ r
marquage accentuel	0	0	0	1	0	1	0	1	1	0	0	1

Dans les cas de pronom enclitique après le verbe, l'accent est déplacé sur le pronom, mais en revanche, on analyse bien la catégorie syntaxique de chacune

des parties. Dans le vers suivant l'accent porte sur *la* (ma(P. 4)=1), mais *la* est analysé comme pronom (cs(P. 4)=6) :

Emportons-la d'ici pour la mieux secourir,

Corneille, *Suréna*, vers 1735

sm	ɑ̃	p ɔ r	t ɔ̃	l a	d i	s i	p u r	l a	m j ø	s ə	k u	r i r
ma	0	0	0	1	0	1	0	0	1	0	0	1
cs	1	1	1	6	3	3	7	6	0	1	1	1

(6 : pronom)

Les pronoms personnels à valeur substantive sont normalement analysés comme des substantifs, ils portent donc l'accent. Ainsi dans le vers suivant, *vous* est analysé comme nom (cs(P. 6)=0) et porte un accent (ma(P. 6)=1).

À cet effort pour vous qui pourrait le contraindre,

Corneille, *Tite et Bérénice*, vers 39

sm	a	s ɛ	ɟt e	f ɔ r	p u r	v u	k i	p u	r ɛ	l ə	k ɔ̃	t r ɛ̃ d r
ma	0	0	0	1	0	1	0	0	1	0	0	1
cs	7	6	0	0	7	0	9	1	1	6	1	1

(0 : nom)

Malheureusement, parfois ces pronoms toniques sont analysés comme mots-outils et ne sont pas marqués.

L'analyseur syntaxique Sylex distingue trois types d'unités polylexicales : le mot composé (*peut-être*), la locution grammaticale (*au dessus-de*) et l'expression semi-figée (*verser du sang*). En raison de certaines difficultés d'analyse, nous avons pris le parti, dans tous les cas, de donner la catégorie syntaxique de chacune des entités lexicales impliquées, de marquer les fins de mots intérieures et d'accentuer chacune des parties de l'expression en fonction de sa catégorie syntaxique.

Le tableau suivant montre quelques exemples de marquage de mots composés :

	peut-être	au devant de	hé bien	moi-même
Fin de mot	1 0 1	1 0 1 1	1 1	1 0 1
Catégorie syntaxique	1 1 1	7 7 7 7	0 2	6 3 3
Marque accentuelle	1 1 0	0 0 0 0	1 1	0 1 0

Tableau 21. Le traitement des mots composés

Ce choix est regrettable pour quelques mots composés, figés depuis longtemps dans la langue, dont on peut regretter qu'ils ne soient pas traités comme un seul mot et qu'ils aient donc plusieurs marques accentuelles. C'est très clairement le cas pour *peut-être*.

Ce marquage n'est que moyennement satisfaisant. Au lieu d'indiquer la place de *l'accent de mot phonologique*, il indique *l'accent de mot*. L'analyse syntaxique ne nous permet pas, à l'heure actuelle, de repérer correctement les mots phonologiques de langue à l'intérieur du vers. En effet, l'analyse syntaxique rencontre de nombreuses difficultés dans le vers classique et le découpage en syntagmes est trop entaché d'erreurs pour pouvoir s'y fier. Introduire une hiérarchisation d'accent, qui impliquait d'utiliser la structure syntaxique globale de la phrase, aurait conduit à trop d'erreurs.

Notre marquage accentuel est la combinaison des trois marquages précédemment vus. La marque « accentuelle » est conçue comme une combinaison des marquages vocaliques, syntaxiques et de fin de mot. Il nous faut en effet connaître la catégorie syntaxique du mot pour savoir s'il est accentuable ou non. Il nous faut ensuite l'indication de position de la dernière syllabe, et enfin, en fonction de la nature vocalique de cette dernière position on placera l'accent sur celle-ci (si elle n'est pas occupée par un *e* muet) ou sur la position qui la précède (s'il s'agit d'un *e* muet).

Dans l'idéal, le marquage accentuel aurait pu être enrichi par un marquage de ponctuation. On aurait pu affecter à chacun des signes de ponctuation un poids d'autant plus élevé que la ponctuation est forte. Ce marquage de ponctuation a été utilisé pour la première fois par Lusson dans « Isomorphismes rythmiques dans les récitatifs de Lulli » [1982] puis largement utilisé pour l'étude du vers ïambique anglais, russe, pour le décasyllabe français et pour l'alexandrin... Il a été utilisé pour étudier la figure du vers mais aussi celle du poème, celle du sonnet plus précisément. Ce marquage a été repris par Cornulier [1995] qui a désigné cette méthode par le terme « ponctuométrie ». Dans la version actuelle du métromètre, nous n'avons pas pu mettre en œuvre ce marquage.

V.4.3. Une autre approche de l'accent, le marquage métrico-syntaxique

Pour définir le marquage accentuel présenté dans la section ci-dessus, nous avons utilisé des théories linguistiques qui ne sont pas consensuelles, et surtout nous avons finalement adopté un marquage accentuel très rudimentaire reposant sur l'accent de mot plein. Rappelons que nous souhaitions mettre en œuvre une hiérarchisation des accents selon leurs positions dans les tronçons intonatifs, mais que les difficultés de l'analyse syntaxique sur le vers ne nous l'ont pas permis.

Cette situation nous paraissant insatisfaisante, nous avons cherché à la compléter par une autre, qui en soit indépendante. Pour cela, nous avons adopté une approche radicalement opposée. Supposons que nous ne sachions rien de l'accent. Quelle démarche pouvons-nous adopter pour décrire le rythme du vers ?

La dernière position du vers est la plus marquée, car elle est à la fois la fin du vers (passage à la ligne), la fin du second segment métrique, et elle porte la rime. Pour identifier les positions les plus marquées du vers, nous allons utiliser les caractéristiques de la fin du vers et les projeter sur les autres positions. Or la dernière position du vers coïncide avec la dernière syllabe d'un mot, ou l'avant-dernière si le mot finit par un *e* muet. Nous allons donc marquer la dernière position (ou avant-dernière) des mots en affectant à chaque mot un poids déterminé par le poids de la catégorie syntaxique auquel il appartient en position 12. Autrement dit, nous pouvons affirmer avec certitude que cette position marque un accent. Au lieu de marquer l'accent par la catégorie syntaxique, nous avons fait l'inverse, en considérant que le placement en dernière position marquait la catégorie syntaxique du mot correspondant. On obtient ainsi pour chaque catégorie syntaxique un taux de marquage accentuel qui est le taux d'occurrence de cette catégorie en dernière position.

Ainsi, si nous avons 40 % de noms en position 12 et 2 % de pronoms personnels, chaque fois que l'on rencontrera un nom on affectera à sa dernière ou avant-dernière position métrique le poids 0,4, et pour les pronoms, on affectera le poids 0,02.

Pour établir ce marquage, il nous faut la répartition des catégories syntaxiques en position 12. Nous l'avons car c'est un des résultats du marquage syntaxique. Nous verrons d'ailleurs que ce marquage donne des résultats cohérents avec la marque accentuelle, ce qui donne une certaine validité empirique au choix de marquage des mots pleins.

V.5. SYNTHESE SUR LES MARQUAGES ADOPTES

Nous n'avons pas pu mettre en œuvre tous les marquages que nous souhaitions, principalement en raison des limites de l'analyse syntaxique appliquée au vers. Finalement, nous avons retenu comme marquages :

- le marquage des syllabes métriques ;
- le marquage des noyaux vocaliques, avec un jeu de dix-neuf noyaux vocaliques tiré des observations empiriques sur la rime, la diérèse et le *e* muet ;
- le marquage des fins de mots ;
- le marquage des catégories syntaxiques (en neuf catégories) ;
- le marquage accentuel ou la « marque fondamentale » de Roubaud [1988] : marquage de la dernière (ou avant-dernière quand le mot finit par un *e* muet) position des mots pleins ;
- le marquage métrico-syntaxique (qui part de la structure de la douzième position pour marquer les autres positions).

Les marquages retenus sont simples et applicables de manière systématique. Leur qualité principale est la cohérence : ils ont été déterminés partout de la même manière, avec des règles explicites. On nous pardonnera sur ce point : cela paraît une condition trivialement nécessaire ; elle n'est, hélas, pas toujours remplie dans les marquages manuels réalisés par certains.

Il existe des relations entre les différents marquages. Certaines sont constitutives : par exemple le marquage d'accent se déduit à partir des marquages de catégories syntaxiques, de fins de mot et de *e* muet. D'autres sont de nature plus statistique, et proviennent de particularités phoniques de

certaines catégories. Par exemple, les mots-outils sont moins nombreux et plus fréquents que les mots pleins. Ils impriment des sonorités particulières aux positions qu'ils occupent majoritairement. Ainsi nous verrons que /e/ est typique des positions 1 et 6 à cause de la préposition *et*.

Le marquage accentuel est le plus problématique de notre liste, même si nous avons tenté de le conforter par une approche purement métrique. Nous voudrions apporter un dernier argument en faveur de ce marquage, par rapport au marquage de fin de syntagme dans le vers. A cause du *e* muet final compté, il existe des décalages d'une position entre le marquage de fin de syntagme et le marquage accentuel de mot phonologique, qui lui tombe sur la dernière position accentuée. Or pour le rythme du vers, c'est le marquage de groupe de souffle qui est plus important que le marquage de syntagme, car le vers est fait pour un œil-oreille et pas seulement pour un œil.

Les critères d'observation de la « métrico-métrie », méthode d'analyse métrique du vers proposée par Cornulier [1982b], peuvent être déduits directement à partir de nos marquages. Mais l'inverse n'est pas vrai.

Reprenons les critères F, M, C et P proposés dans le cadre de cette méthode. Fn indique que la n-ième position correspond à une voyelle féminine, autrement dit, si l'on admet l'idée d'accent, à une voyelle postérieure à la dernière voyelle accentué, soit à un *e* faible. Mn indique que la voyelle de la n-ième position « est préconclusive d'un « mot » (antérieure, dans un mot, à sa dernière masculine) » [Cornulier, 1995, p. 275-276], Cn que « sa n-ième voyelle appartient à un proclitique suivi de son appui », Pn que « sa n-ième voyelle appartient à une préposition d'une syllabe suivie de son complément » [Cornulier, 1982b, p. 140].

Le métromètre, à travers les marquages que nous avons choisis d'implémenter, permet de construire aisément les critères de la « métrico-métrie ». Le marquage des *e* muets faibles donne les Fn, celui des fins de mots les Mn, et l'identification des mots-outils permet de construire immédiatement CPn (Cn ou Pn). Comme dans le métromètre les marquages sont construits automatiquement, ils s'appliquent tous à toutes les positions et l'attention n'est

pas centrée sur quelques positions spéciales comme les 4, 6, 7 et 8. Le métromètre permet d'avoir une vision d'ensemble et relative du vers.

Nous disposons en plus dans le métromètre de marquages phonétiques, d'un marquage de catégories syntaxiques et de marque accentuelle. Les marquages de la « métrico-métrie », particulièrement adaptés pour rendre compte de déviations par rapport aux règles du vers classique (*e* muet en sixième position, mot à cheval sur l'hémistiche...), semblent moins opératoires pour construire une figure globale des vers.

Comme nous l'avons dit précédemment, seuls ont été introduits des marquages de discours qui valent pour la prose comme pour le vers. Si, dans le dénombrement et l'identification des positions métriques, nous nous situons exclusivement dans le cadre du vers, à aucun moment nos marquages ne font l'hypothèse qu'il s'agit d'un vers. Ainsi, fin de vers et fin d'hémistiche n'interviennent pas dans les marquages et ne constituent pas en tant que tels un marquage spécifique. *Nous n'avons pas introduit de marquage métrique*. Derrière ce choix, se situe évidemment le désir de tester une hypothèse forte : le rythme n'est-il qu'agencement habile de traits de la langue commune ?

Maintenant que nous avons défini un certain nombre de marquages de langue (contenu phonétique et vocalique des positions métriques, indication des fins de mots, catégories syntaxiques des positions métriques et positions des accents de mot), nous allons montrer qu'ils sont suffisants pour rendre compte des aspects rythmiques du vers, et que lorsqu'ils ne le sont pas, ce sont les marquages qu'il faut incriminer (nous n'avons pas toujours pu mettre en œuvre ceux que nous souhaitions) plutôt que notre hypothèse.

VI. LE METROMETRE : L'OUTIL ET SES LIMITES

Puisque nous cherchons à identifier la structure phonétique, syntaxique et prosodique du vers, nous souhaitons disposer d'un outil qui, d'une part, segmente la séquence graphique du vers en positions métriques et en donne une transcription phonétique et qui, d'autre part, décrive chaque position par un ensemble de marquages.

Un simple dénombrement des positions métriques ne nous suffit pas[107]. Nous souhaitons aussi disposer des caractéristiques concrètes de chacune des positions dans les dimensions discutées dans le chapitre V : les phonèmes de la syllabe, le noyau vocalique, la catégorie syntaxique (catégorie du mot qui contient la syllabe) et le marquage accentuel. Autrement dit, nous souhaitons que chaque vers soit décrit par sa séquence phonétique, syntaxique et par sa ligne prosodique. Bref, il fallait un programme pour appliquer, au corpus concret de vers que nous avions construit au chapitre III, les règles de découpage définies au chapitre IV et les marquages définis au chapitre V. Il s'agit, très concrètement de réaliser environs dix millions d'opérations individuelles de marquage. À l'évidence, il fallait faire appel à l'informatique.

Nous avons utilisé des outils de linguistique informatique déjà existants : le lexique du français, construit à l'origine par Frédéric Pigamo [1990], l'analyseur syntaxique par couches développé par Patrick Constant [1991] et le

[107] Nous avons dans un premier temps tenté de mettre en œuvre le programme Dynastie [Roubaud, 1986] de dénombrement du vers classique, avec l'aide de J. - G. Ganascia, professeur à Paris VI, et de deux de ses doctorants (V. Corruble et G. Ramalho). L'ambition du projet était de partir de la séquence graphique pour identifier les positions métriques, sans introduire ni syntaxe, ni phonologie. Nous avons finalement abandonné cette voie, non pas tant à cause des problèmes de dénombrement (qui concernaient principalement la question de la diérèse et qui auraient pu être résolus), mais parce que ce système ne nous permettait pas d'utiliser des marquages de langue. L'annexe II présente le programme Dynastie et sa mise en œuvre.

phonétiseur développé par François Yvon [1993 et 1996]. Tous ces outils ont été conçus et programmés dans le laboratoire Informatique de l'ENST. En mettant en œuvre conjointement un analyseur syntaxique et un phonétiseur, en y ajoutant notre ensemble de règles, nous sommes en mesure de proposer une représentation multidimensionnelle du vers qui contient les marquages que nous nous sommes proposés.

Nous présentons d'abord les outils de linguistique informatique utilisés et la manière dont ils ont été assemblés pour constituer le métromètre. Ensuite, nous passerons en revue les limites du métromètre.

VI.1. OUTILS DE LINGUISTIQUE INFORMATIQUE

L'analyseur Sylex a été développé par Patrick Constant, à l'origine à l'ENST. Il a été développé sous un environnement UNIX, mais il en existe actuellement une version pour les PC. Le phonétiseur a été développé sous UNIX par François Yvon, maître de conférence à l'ENST, dans le Département Informatique. L'analyseur syntaxique est complètement intégré dans le phonétiseur.

Le métromètre [Beaudouin et Yvon, 1994 et 1996] est constitué d'une couche de règles supplémentaire, rajoutée à l'édifice, qui permet au phonétiseur de fonctionner selon les normes de la métrique classique. Il n'existe pour le moment qu'une version du métromètre, à l'ENST, les briques composant l'outil étant la propriété de l'ENST.

Nous décrivons rapidement les principes généraux de fonctionnement des outils utilisés. Pour une présentation plus détaillée, on se reportera à Constant [1991 et 1995] et à Yvon [1993]. Ensuite, nous présentons la manière dont ces outils ont été adaptés au traitement du vers.

VI.1.1. L'analyse syntaxique par couche

L'analyseur syntaxique Sylex est un outil commercial, diffusé par la société INGENIA. Constant [1995] souhaitait construire un système d'analyse

performant en termes de vitesse de calcul. Dans ce but, il s'est appuyé sur la grammaire de Tesnière [*1959*, 1982] et les grammaires de liens ([Sleator et Temperley, 1991] cités par Constant, 1995]) et n'a pas adopté de représentations en arbres et de règles de dérivation qui ont une complexité trop importante.

L'objectif de Constant est de reproduire le processus d'analyse de textes que nous réalisons quotidiennement pour comprendre des énoncés. Il pose une distinction fondamentale entre syntaxe descriptive et syntaxe analytique : « En syntaxe descriptive, la compréhension de la phrase préexiste à sa structuration. En syntaxe analytique, c'est le contraire » [Constant, 1991, p. 95]. Nous nous situons donc dans un contexte de syntaxe analytique, et dans ce contexte : « les deux problèmes les plus importants (...) sont l'attachement et l'ambiguïté » [Constant, 1991, p. 95].

Autrement dit, il faut résoudre les ambiguïtés lexicales mais aussi les ambiguïtés syntaxiques ou structurales. Or les tests réalisés par Constant montrent qu'en moyenne 50 % des mots sont ambigus lexicalement dans un corpus. Partant du constat que l'ambiguïté est un constituant fondamental de la langue naturelle, qui se résout aisément en situation de compréhension, mais pas forcément simplement en syntaxe analytique, Constant propose un modèle d'analyse syntaxique qui prend acte de cette ambiguïté fondamentale et la résout pas à pas. Une autre caractéristique de l'analyseur syntaxique est qu'il est insensible à la notion de phrase grammaticalement correcte. Il est ainsi en mesure de donner une interprétation syntaxique de phrases incomplètes. Nous pouvons donc utiliser cet analyseur pour traiter des vers, même s'ils ne correspondent qu'à une partie d'une phrase. Il en résultera naturellement que l'analyse sera moins complète, en particulier que les liens de dépendance entre groupes syntaxiques situés sur des vers différents ne pourront être identifiés.

Concrètement, l'analyseur fonctionne de la manière suivante. La première étape est un accès lexical. Chaque mot ou locution (groupement de mots) identifié dans la chaîne de caractères est recherché dans le lexique. Ce lexique, développé par Frédéric Pigamo [1990], contient environ 480 000 formes, à savoir des entrées de dictionnaires mais aussi toutes les formes fléchies et une

quantité non négligeable de locutions, environ 20 000. Rappelons à titre de comparaison que le Grand-Robert en neuf volumes contient 80 000 entrées (sans formes fléchies).

Ensuite, à chaque forme est associé l'ensemble de ses catégories morpho-syntaxiques *potentielles*. On est alors dans une situation d'ambiguïté maximale.

Les étapes suivantes auront pour objectif de réduire cette ambiguïté. Chaque étape correspond à l'application d'un ensemble ou couche de règles, traitant d'un problème syntaxique particulier : ambiguïté nom/verbe, structuration des groupes verbaux... Il existe à peu près 70 couches de règles, qui regroupent environ 25 000 règles. Chacune de ces couches permet d'éliminer un certain nombre des interprétations de départ en raison de leur incohérence. En fin de course, il se peut que certaines ambiguïtés n'aient pas été levées. Dans ce cas, les différentes interprétations syntaxiques sont proposées.

A la fin de l'analyse, on dispose des informations suivantes : la catégorie morpho-syntaxique de chaque terme ; les groupements syntaxiques ; les liens éventuels entre ces groupements. Ces liens ont en général pour source le verbe (*cf.* les stemmas de Tesnière [1959, 1982] dont le sommet de l'édifice est toujours le verbe) et pointent vers des groupements : vers le sujet, l'objet direct, l'objet indirect...

VI.1.2. La transcription graphème-phonème

Nous avons choisi de représenter le vers écrit par une séquence de phonèmes, qui est censée représenter la forme orale. Nous ne cherchons pas vraiment à nous approcher de la réalisation orale, mais simplement à nous libérer des variations orthographiques. L'alphabet phonétique international dans lequel les vers sont transcrits permet de retrouver l'équivalence entre une lettre (un phonème) et un son, et cela quel que soit le contexte.

Le système de transcription graphème-phonème utilisé pour l'analyse du vers est celui qu'a développé François Yvon. Dans le cadre de son travail de phonologie celui-ci rappelle que :

> (...) Le problème de la conversion graphème-phonème consiste (...) à reconstituer, à partir d'une forme graphique (un mot, un syntagme, une phrase) exprimée dans un système [alphabétique], la forme orale qui lui correspond. Cette forme orale est exprimée sous l'aspect d'une séquence de phonèmes.
>
> [Yvon, 1996, p. 22].

Notons que représenter graphiquement la forme orale ne va pas de soi. Comment représenter en particulier les aspects prosodiques ? Si l'on vise la synthèse de la parole, cette étape de transcription est cependant indispensable.

Les difficultés posées par cette opération de transcription sont de deux ordres : des difficultés liées à la variabilité et d'autres liées à une déviation du système orthographique, par rapport à un modèle idéal de correspondance biunivoque entre lettre et son, entre graphèmes et phonèmes. Nous examinerons successivement ces deux catégories de difficultés.

En ce qui concerne la variabilité, il faut distinguer à la suite de Lacheret-Dujour [1990, cité par Yvon, 1996] la variabilité systématique et la variabilité libre. La variabilité systématique recouvre tous les cas où la correspondance entre graphème et phonème varie selon l'environnement graphématique immédiat (un *c* suivi de *a*, *o*, *u* sera un /k/ ; suivi de *e*, *i* un /s/). Aux frontières de mot, les phénomènes de liaison ont une incidence importante sur la réalisation phonétique de certains phonèmes tels que les consonnes latentes.

La variabilité libre peut provenir de variantes régionales (on ne prononce pas « rose » de la même manière à Paris et à Marseille), idiolectales et prosodiques. Ce sont principalement des variations liées au contexte d'énonciation. Ces variations sont particulièrement sensibles pour le *e* muet qui se maintient d'autant plus fréquemment que le niveau de langue est soutenu, pour la liaison dont l'usage varie selon le niveau de langue et dans une moindre mesure pour la diérèse. Ainsi selon les conditions générales d'énonciation, on aura de grandes variations dans la réalisation orale. Il suffit pour en avoir un aperçu immédiat de parcourir la bande FM. De France Culture à Fun Radio, tous les niveaux de langue y sont représentés.

> La transcription graphème-phonème n'a donc de sens qu'en référence à une norme de prononciation donnée, qui spécifie à la fois les alphabets de symboles valides en entrée comme en sortie, et le niveau stylistique de prononciation que ces symboles encodent.
>
> [Yvon, 1996, p. 25].

Il existe deux grandes déviations du système orthographique par rapport au modèle idéal. Premièrement, il n'y a plus de correspondance biunivoque entre lettre (ou graphème) et phonème. Une lettre peut correspondre à deux phonèmes (x =>ks), deux (ou plus) graphèmes à un seul phonème (ai=>ε).

Deuxièmement, une même séquence graphique peut avoir des réalisations différentes. Des séquences de lettres se prononcent parfois différemment selon le mot dans lequel elles se situent. Ainsi *oi* dans *oignon* correspond à /o/ et dans *poigne* à /wa/. De manière exemplaire, c'est aussi le cas des homographes hétérophones, des mots qui s'écrivent de la même manière mais qui ont des prononciations différentes. Par exemple, *ent* dans *divergent*, correspond à un *e* muet s'il s'agit de la troisième personne du pluriel du verbe *diverger* et un /ɑ̃/ s'il s'agit de l'adjectif. L'analyse syntaxique permet de lever ces ambiguïtés. Ainsi, elle distinguera l'adjectif *content*, du verbe *content* (*ils content*, 3ème personne du pluriel), qui seront respectivement transcrits en /k ɔ̃ t ɑ̃/ et /k ɔ̃ t əʃ/. Elle distinguera aussi l'adjectif *fier* du verbe *fier*, le verbe à l'imparfait *nous passions* du nom *passions*. Lever ces ambiguïtés est essentiel pour le décompte métrique puisque le verbe *fi-er* et le nom *passi-ons* portent une diérèse contrairement à leurs homographes respectifs (*cf.* chapitre IV.1). Le marquage morpho-syntaxique améliore sensiblement la transcription des mots. Certaines ambiguïtés ne peuvent cependant pas être levées, lorsque les deux homographes appartiennent à la même catégorie syntaxique comme entre *fils* (*bobines de fils multicolores*) et *fils* (*les crimes de mes fils*), mais elles n'ont pas d'incidence sur le décompte métrique.

Ces déviations du système orthographique sont principalement liées à l'évolution plus lente de l'écrit par rapport à l'oral. L'orthographe porte la trace d'états antérieurs de la prononciation. Comme la poésie [Roubaud, 1969], l'orthographe est une forme-mémoire de la langue. En s'écartant de la

prononciation orale, l'orthographe s'est elle-même constituée en système partiellement autonome.

Toutes ces difficultés impliquent que l'automatisation de la transcription graphème-phonème fasse simultanément appel à des systèmes de règles et à des listes ou lexiques d'exceptions. Ainsi, par exemple, certains cas de diérèse systématiques ont été traités par des règles, tandis que les cas qui ne pouvaient être formalisés par des règles ont été introduits dans le lexique.

Concrètement, une fois que l'analyse syntaxique a attribué à chaque mot son étiquette syntaxique et qu'elle a déterminé l'intensité des frontières de mot, la transcription graphème-phonème proprement dite peut s'appliquer. Elle comprend trois étapes. Dans un premier temps, le phonétiseur procède mot à mot à une vaste harmonisation orthographique pour normaliser les cas trop atypiques, c'est-à-dire les mots dont l'orthographe est peu cohérente avec la prononciation. Au cours d'une seconde étape, chaque mot, récrit dans une orthographe simplifiée, est transcrit dans l'alphabet phonétique à l'aide d'un petit jeu de règles. Enfin, la dernière étape consiste à appliquer un ensemble de règles phonologiques, paramétrables en fonction du type de transcription souhaité. Ces règles évaluent l'intensité des *e* muets, des liaisons et traitent aussi de la nasalisation des voyelles.

Pour répondre à la question de la variabilité, le phonétiseur peut produire autant de transcriptions phonétiques que de registres de langue définis. Si l'on fait le choix d'un registre de langue familier, tous les *e* muets vont disparaître, seules les liaisons obligatoires seront faites... Si, au contraire, on choisit le registre « vers classique », la transcription sera conforme au vers. C'est au cours de la troisième étape que se jouent les variations dans la transcription[108].

[108] L'évaluation de la qualité d'un phonétiseur ne va pas de soi. Nous renvoyons à Yvon et al. [1998] pour la description d'un système d'évaluation appliqué à huit systèmes de transcription graphème-phonème.

Voici son fonctionnement sur un énoncé simple : *le chien mange.* Cet exemple est tiré de [Beaudouin et Yvon, 1994].

Règles	Nouvelle représentation
Règles de réécriture	
("le", article) -> lə	lə | chien mange
("chien", nom) -> ʃiɛn	lə | ʃiɛn || mange
("mange", verbe) -> mɑnʒə	lə | ʃiɛn || mɑnʒə
Règles phonologiques	
Semi-vocalisation des voyelles hautes	lə | ʃjɛn || mɑnʒə
Nasalisation + chute de la consonne nasale	lə | ʃjɛ̃ || mɑ̃ʒə
Suppression du *e* muet final	lə | ʃjɛ̃ || mɑ̃ʒ

Tableau 22. Dérivations phonologiques

VI.1.3. Déroulement pratique

Dans la pratique, on active dans le phonétiseur le jeu de paramètres qui correspond à l'analyse du vers, dans lequel sont prises en compte les questions de diérèse, de *e* muet et de liaison. Le métromètre fonctionne alors de la manière suivante. Il prend comme entrée le vers, ou une fraction de vers si celui-ci est découpé. Le vers devient alors un énoncé à part entière dans l'analyse, même s'il ne coïncide pas toujours avec la phrase.

De cette nuit, Phénice, as-tu vu la splendeur ?

Racine, *Bérénice*, vers 302.

A chacun des mots du vers est associée sa catégorie morpho-syntaxique. Si le terme est ambigu, toutes les catégories potentielles sont affectées au mot. Par exemple *lui*, peut être un pronom, un nom (quand il a une valeur substantive après une préposition) ou le participe passé de *luire*. Les différentes couches de règles permettent ensuite de lever une partie des ambiguïtés et de déterminer les groupements syntaxiques. La nature et l'intensité des frontières de mots est alors définie :

de # cette # nuit ## Phénice ## as # tu # vu # #la# splendeurr#

Une frontière entre deux groupements syntaxiques est plus forte qu'entre deux mots à l'intérieur d'un groupement syntaxique. Il se peut que l'analyseur ne parvienne pas à désambiguiser complètement le vers, il propose alors toutes les interprétations compatibles avec les règles.

Ensuite, chaque mot (affecté de son étiquette catégorielle) est transcrit dans l'API[109] :

đə # s ɛ t ə # n y i t_l[110]## f e n i s ə ## a z_l # t y # v y # #l a# s p l a n dœ r#

Enfin, l'ensemble des règles phonologiques définies spécialement pour le vers est appliqué. Les voyelles hautes deviennent des semi-voyelles, sauf dans les cas de diérèse où la règle de semi-vocalisation de la voyelle haute est bloquée (*cf.* chapitre IV.1).

đə # s ɛ t ə # n **ɥ** i t_l ## f e n i s ə ## a z_l # t y # v y ## l a# s p l a n dœ r #

Certaines voyelles se nasalisent et font tomber la consonne nasale, les consonnes latentes devant consonnes s'effacent :

đə # s ɛ t ə # n ɥ i ## f e n i s ə ## a # t y # v y # #l a# s p l **ɑ̃** dœ r#

Chaque *e* muet est analysé : selon sa position dans le vers et dans le mot, il sera interprété comme un *e* muet fort, faible ou inexistant (comme nous l'avons vu dans le chapitre IV.2).

đ**ə** # s ɛ t **$ə_f$**[111] # n ɥ i ## f e n i s_## a # t y # v y # #l a# s p l ɑ̃ dœ r#

Enfin, les syllabes ou positions métriques sont délimitées et dénombrées, en respectant les règles de découpage syllabique présentées au début du chapitre IV. La question des consonnes de liaison est réglée à ce moment-là. Pour les liaisons potentielles, l'évaluation de l'intensité de la frontière syntaxique permettra de déterminer s'il faut ou non faire la liaison et dans ce dernier cas s'il s'agit de liaisons avec ou sans enchaînement (*cf.* chapitre IV.3). Dans l'exemple suivant, il n'y a pas de cas de liaison :

---

[109] Alphabet Phonétique International

[110] Un t_l : indique un *t* latent en fin de mot, qui peut devenir consonne de liaison en fonction de son environnement et de sa position syntaxique.

[111] e_f : *e* muet faible.

(đə)(s ɛ)(t əf)(n ɥ i)(f e)(n i s)(a)(t y)(v y)(l a)(s p l ɑ)(dœ r)

Voici un autre exemple décrivant les différentes étapes de transcription du vers.

Et mes coursiers oisifs ont oublié ma voix,

e # m e z_l [112]# k u r s i e z_l # # w a z i f z_l ## ɔ n t_l # u b l i e ##m a # v w a z_l #

- semi-vocalisation des voyelles hautes :
  e # m *e* z_l # k u r s j e z_l # # w a z i f z_l # #ɔ n t_l # u b l i e ## m a # v w a z_l #
- nasalisation +chute de la consonne nasale + chute des consonnes latentes devant consonne :
  e # m *e* # k u r s j e z_l # #w a z i f z_l # #ɔ̃ t_l # u b l i e ## m a # v w a z_l #
- syllabation et traitement des consonnes latentes :
  (e) (m *e*)(k u r)(s j e z_l)(w a)(z i f z_l)(ɔ̃)(t_l u)(b l i)(e)(m a)(v w a)

Racine, *Phèdre*, vers 552.

Une fois la transcription phonétique effectuée, les différents marquages présentés dans le chapitre V sont affectés à chacune des positions. Chaque position est décrite par la syllabe phonétique qui l'occupe et par le noyau vocalique de cette syllabe. La catégorie syntaxique de chaque mot est projetée sur chaque syllabe. Ensuite, grâce à la position des frontières de mot, les syllabes qui correspondent à la fin d'un mot sont marquées. Enfin, la position des « accents » de mots est indiquée : elle résulte de la confrontation du marquage de catégorie syntaxique (qui distingue les mots pleins et les mots vides), du marquage de fin de mot et de la position des *e* muets faibles. En effet, sont marquées les positions qui correspondent à la dernière syllabe, ou à l'avant-dernière si la dernière est un *e* muet, des mots pleins.

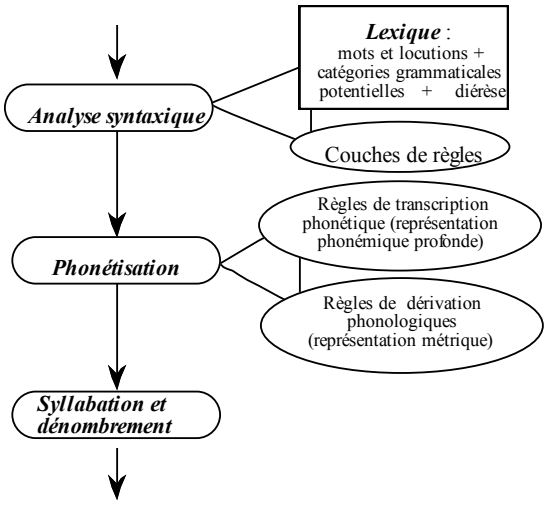
La Figure 21, qui apparaissait déjà en introduction, présente le fonctionnement du métromètre et la représentation du vers qui en résulte. Nous avons fourni dans l'annexe III, les résultats bruts que produit le métromètre appliqué à la première scène de *Tite et Bérénice* et aux stances de Don Rodigue dans *Le Cid*.

---

[112] z_l : consonne de liaison.

Ils sont sortis, Olympe ? Ah mortelles douleurs !

Racine, *La Thébaïde*, vers 1.



Syllabes métriques	i l	s ɔ̃	s ɔ r	t i zɩ	ɔ	l ɛ̃ p	a	m ɔ r	t ɛ	l əf	d u	l œ r
Voyelles métriques	i	ɔ̃	ɔ	i	ɔ	ɛ̃	a	ɔ	ɛ	əf	u	œ
Repérage des fins de mots	1	1	0	1	0	1	1	0	0	1	0	1
Catégories syntaxiques	6	1	2	2	4	4	0	2	2	2	0	0
Marquage accentuel	0	0	0	1	0	1	1	0	1	0	0	1
Nb syllabes	12											

Clef de lecture : La neuvième syllabe métrique correspond à la séquence phonémique / t ɛ/, elle n'est pas située en fin de mot (*cf.* repérage des fins de mots=0), appartient à un adjectif (*cf.* catégorie syntaxique=2) mais porte un accent de mot (marquage accentuel=1).

Repérage des fins de mot :

0 = syllabe interne	1 = syllabe finale

Catégories syntaxiques

0 = Nom	6 = Déterminant, pronom
1 = Verbe	7 = Préposition
2 = Adjectif	8 = Conjonction
3 = Adverbe	9 = Pronom relatif
4 = Nom propre	

Marquage accentuel

0 = syllabe ne porte pas d'accent	1 = syllabe porte un accent de mot

Figure 21. Fonctionnement du métromètre et représentation métrique

En résumé, l'analyse syntaxique intervient à différents niveaux :

- en amont de la transcription graphème-phonème pour distinguer les homographes hétérophones quand ils n'appartiennent pas à la même catégorie syntaxique ;
- au moment du traitement de la liaison : c'est l'évaluation de l'intensité des frontières de mots qui définit l'existence et la nature de la liaison. ;
- au cours du marquage accentuel. Nous avons vu (*cf.* chapitre V) que l'option a été prise de ne marquer que les mots pleins (noms, verbes, adjectifs et adverbes), il faut donc pour cela connaître la catégorie morpho-syntaxique de chaque mot ;
- enfin pour la construction de la figure syntaxique du vers, où le marquage morpho-syntaxique est affecté à chacune des positions métriques.

La représentation multidimensionnelle du vers (autant de dimensions que de marquages) a été conçue pour pouvoir être intégrée dans une base de données. Dans cette base (*cf.* annexe IV) chaque vers est décrit par cinq variables (syllabe métrique, voyelle, catégorie syntaxique, fin de mot et « accent ») multipliées par douze positions, soit 60 variables. S'y ajoutent un certain nombre de variables construites à partir des premières comme la séquence accentuelle des hémistiches, le nombre de phonèmes de chaque syllabe... Le travail sur la rime (*cf.* chapitre III) nous a également amené à construire de nouvelles variables comme le mot-rime, le genre de la rime, le groupe de mots-rimes, le rimème... Il faut enfin ajouter les variables paratextuelles (auteur du vers, pièce, genre de la pièce, acte, scène, nom du personnage qui prononce le vers...). Ces données ont été structurées dans une base SAS. La recherche des erreurs du métromètre, des cas exceptionnels, la construction de la figure du vers, les tests statistiques ont tous été faits à l'aide de ce progiciel[113]. S. Fisher [1995] a bien montré comment les approches linguistiques pouvaient utiliser des fragments de texte, attestés ou construits, transformés en « exemples », et convoqués pour décrire une pratique ou illustrer une théorie. Notre approche empirique veut extraire les régularités de

[113] Version 6.12 sous Unix, modules SAS de base et SAS Stat [SAS Institute Inc., 1989 et 1990].

l'ensemble même du texte, par des procédés mécaniques et fonder la théorie sur cette observation systématique. On verra d'ailleurs que les « exemples » sont choisis pour leur représentativité, et non pour leur qualité littéraire ou leur caractère exceptionnel.

L'outil fonctionne et produit des résultats inédits et remarquables sur de grands corpus de vers.

Maintenant que nous l'avons bien testé sur Corneille et Racine, mais aussi sur d'autres corpus, et que nous avons traité à maintes reprises les résultats qu'il produit, nous connaissons ses limites. La section suivante les décrit, en proposant un certain nombre d'améliorations qui permettront d'obtenir un système de marquage encore plus complet et cohérent

VI.2. LES LIMITES ACTUELLES DU METROMETRE

Les limites actuelles du métromètre sont de trois ordres. Nous examinerons d'abord celles qui sont liées à la phonétisation, puis celles qui proviennent des difficultés d'analyse syntaxique et enfin celles qui tiennent à l'insuffisance des marquages adoptés.

VI.2.1. De l'écrit à l'oral : problèmes liés à l'orthographe

Le métromètre, à partir de textes écrits, produit une représentation phonétique abstraite de ce que pourrait être la réalisation orale du vers : l'exécution, la diction.

Il aurait sans doute fallu donner une transcription phonétique conforme à la prononciation de l'époque où furent publiés ou représentés les textes. Et c'est là une difficulté du métromètre. Elle tient à l'incohérence entre le « réel » (les textes d'origine), les « données » (les éditions utilisées) et les fondements techniques du métromètre (outils conçus pour l'analyse de la prose actuelle).

Les textes sur lesquels nous travaillons sont de pièces conçues et écrites au XVIIe siècle et publiées dans une orthographe qui diffère sensiblement de la nôtre, comme en témoigne l'examen des éditions originales : pluriel en *-ans* et *-ens* (contre *-ants* et *-ents* aujourd'hui), le *s* avant consonne n'est pas encore transformé en accent circonflexe (*estre* plutôt que *être*), les accents graves n'apparaissent pas sur le *e*... Ces différences dans la graphie ne semblent cependant pas avoir une incidence importante sur la structure phonétique et le rythme des vers : par exemple l'opposition de longueur entre les voyelles a déjà disparu.

Dans notre corpus informatisé, nous ne disposons pas d'éditions d'origine mais d'éditions dont l'orthographe a été modernisée. Dans le métromètre, le lexique et la syntaxe utilisés sont de notre époque : les mots sont enregistrés dans leur orthographe d'aujourd'hui et la transcription phonétique reflète la prononciation actuelle. Il y a une certaine cohérence entre les éditions utilisées et le métromètre, mais il est vraiment regrettable de n'avoir pas pu travailler sur les éditions d'origine. Le métromètre tel qu'il a été développé s'adapte mieux à l'analyse des vers du XIXe siècle, puisque l'orthographe n'est pas très différente de la nôtre.

Notre corpus nous soutient dans notre démarche (transcription phonétique des vers dans une prononciation d'aujourd'hui et non pas dans une prononciation d'époque). En effet, il est constitué de pièces de théâtre, conçues pour la scène où les vers sont dits. De plus, il s'agit des pièces de Corneille et Racine, les auteurs les plus joués depuis le XVIIe siècle : la seule Comédie-Française leur a consacré depuis sa création plus de 16 000 représentations. Or la diction des vers au théâtre, sauf expériences particulières comme la représentation de *La Place Royale* à la Cartoucherie en 1995, a évolué avec la langue. Et si les comédiens, du moins aujourd'hui, respectent la mesure du vers, les douze positions, ils prononcent le français dans sa forme actuelle. L'expérience montrant que ces pièces sont toujours « jouables », c'est donc que la langue n'a pas évolué au point que les effets sonores (rimes, allitérations et assonances) disparaissent. Les outils utilisés, bien qu'ils ne prennent pas en compte un certain nombre d'évolutions subtiles (et qui sans doute nous resteront à jamais inconnues faute d'enregistrements sonores datant du XVIIe

siècle), permettent au moins d'appréhender les caractéristiques majeures qui nous intéressent pour l'étude du rythme.

VI.2.2. Les limites de l'analyse syntaxique

Sur les textes en vers, l'analyse syntaxique peine fréquemment et produit parfois des résultats aberrants. Ces difficultés tiennent à des choix que nous avons (mal) faits, à des défaillances de l'analyse syntaxique proprement dite et enfin à des particularités syntaxiques propres au vers.

VI.2.2.1. Le vers comme unité d'analyse

Par commodité et parce que l'objet d'étude est le vers, nous avons choisi comme unité d'analyse syntaxique le vers, et non la phrase. Or, Robert Garrette [1995] a montré que sur l'ensemble des pièces de Racine, seuls 36 % des « phrases grammaticales »[114] et 34 % des « phrases stylistiques » coïncident avec le vers. Dans près de deux tiers des cas, l'analyse porte sur des vers qui ne sont donc que des fragments de phrase.

Il aurait fallu faire porter l'analyse syntaxique sur la phrase entière et ensuite projeter sur le vers les résultats de l'étiquetage syntaxique et de l'analyse des syntagmes, puisque le vers est notre objet d'étude.

Prenons à titre d'exemple les deux vers de *Mélite* :

> Et je ménage en vain dans un éloignement
> Un peu de liberté pour mon ressentiment :
>
> Corneille, *Mélite*, vers 8-9.

A l'heure actuelle, le métromètre analyse un par un chaque vers : il ne peut donc identifier les liens syntaxiques entre les deux groupements du second vers et le verbe du premier (*ménage*). Il n'identifiera qu'un seul lien entre 'dans un

---

[114] Pour Garrette [1995], une phrase grammaticale coïncide avec une unité de prédication, et une phrase stylistique avec une unité d'énonciation ; cette dernière est définie par la ponctuation. Par exemple,

> [(Je ne me soutiens plus :) (ma force m'abandonne.)] (*Phèdre*, vers 154)

est constitué de deux phrases grammaticales et d'une seule phrase stylistique.

éloignement’ et ` je ménage en vain’, où le premier est complément circonstanciel du second.

` je ménage en vain’(4) ->- cc position head[115] ->- `dans un éloignement’(16)

En revanche, l’analyse de la phrase complète permet d’identifier trois liens : `pour mon ressentiment’ complément circonstanciel de but, `un peu de liberté’ complément d’objet et ‘dans un éloignement’ complément circonstanciel de position de ‘je ménage en vain’.

`je ménage en vain’(156) ->- cc but head[116] ->- `pour mon ressentiment’(169)
`je ménage en vain’(156) ->- a2 head ->- `un peu de liberté’(164)
`je ménage en vain’(156) ->- cc position head ->- `dans un éloignement’(160)

On notera cependant que les syntagmes sont correctement identifiés dans les deux cas. Dans le vers classique, les fins de syntagmes coïncident en général avec les fins de segments métriques. Pour cet exemple, choisir le vers comme unité d’analyse syntaxique ne pose pas de trop grandes difficultés.

En revanche, à partir de Hugo, qui fait beaucoup d’enjambement, le choix du vers comme unité d’analyse syntaxique devient beaucoup plus problématique. Ainsi, l’analyse des vers suivants est-elle *sensiblement* différente selon que l’on analyse chaque vers isolément ou la phrase complète.

Ils atteindront le fond de l'Asturie avant
Que la nuit ait couvert la sierra de ses ombres

Hugo, *La Légende des siècles*, Le petit roi de Galice, IV.

Dans l’analyse de la phrase présentée ci-dessous, *avant que* est reconnu comme « relatif », comme conjonction de subordination, tandis que dans l’analyse vers à vers, *avant* est traité comme une préposition.

---

[115] Complément circonstanciel de position de ‘je ménage en vain’.

[116] Cc but head : complément circonstanciel de but ; a2 head : complément d’objet ; cc position head : complément circonstanciel de position.

```
******************** Phrase 0-42 ****************************
"Ils atteindront le fond de l'Asturie avant"

1 Ils il pdet-a1-u
1 atteindront atteindre verbe <a2_head, fond, 2>
1 le le pdet
1 fond fond nom <cn_head, Asturie, 3>
1 de de pprep
1 l'le pdet
1 Asturie 0x0 nompropre
1 avant avant pprep
1 0x0 terminateur

******************** Phrase 0-47 ****************************
"Que la nuit ait couvert la sierra de ses ombres"
1 Que que rel
1 la le pdet
1 nuit nuit nom
1 ait couvrir verbe <a0_arel, Que, -3> <a1_head, nuit, -1>
<a2_head, sierra, 3>
1 couvert couvrir padj-n
1 la le pdet
1 sierra sierra nom <cn_head, ombres, 3>
1 de de pprep
1 ses son pdet
1 ombres ombre nom
1 0x0 terminateur

vb>******************** Phrase 0-90****************************
"Ils atteindront le fond de l'Asturie avant que la nuit ait
couvert la sierra de ses ombres"
1 Ils il pdet-a1-u
1 atteindront atteindre verbe <a2_head, fond, 2>
1 le le pdet
1 fond fond nom <cn_head, Asturie, 3>
1 de de pprep
1 l'le pdet
1 Asturie 0x0 nompropre
1 avant_que avant_que rel
1 la le pdet
1 nuit nuit nom
1 ait couvrir verbe <a1_head, nuit, -1> <a2_head, sierra, 3>
1 couvert couvrir padj-n
1 la le pdet
1 sierra sierra nom <cn_head, ombres, 3>
1 de de pprep
1 ses son pdet
1 ombres ombre nom
1 0x0 terminateur
```

Une partie non négligeable des erreurs est liée au choix du vers comme unité d'analyse syntaxique alors qu'il ne coïncide que le tiers du temps avec la phrase.

VI.2.2.2. Erreurs syntaxiques locales

Mais le métromètre produit aussi des erreurs d'analyse syntaxique qui ne sont pas liées au choix de l'unité d'analyse. Nous en avons repéré deux catégories qui reviennent d'une manière systématique.

La première correspond à des erreurs qui valent autant pour le vers que pour la prose. Celles-ci proviennent principalement de mots qui peuvent potentiellement être associés à plusieurs catégories grammaticales : nom ou pronom, adverbe ou préposition... Ainsi, les pronoms *lui, elle, nous, vous* peuvent, selon la position syntaxique, soit appartenir à la catégorie des mots pleins généraux, ils sont alors des *substantifs personnels* [Tesnière, 1959, 1982, p. 67], soit entrer dans la catégorie des pronoms.

Et le nouvel éclat de votre dignité
Lui doit enfler le cœur d'une autre vanité. pronom
Corneille, *Le Cid*, vers 171-172.

Si l'amour vit d'espoir, il périt avec lui : substantif personnel
Corneille, *Le Cid*, vers 108.

En reprenant une terminologie plus traditionnelle, ces pronoms ont tantôt une valeur prédicative comme dans *il se bat contre lui*, tantôt une valeur non prédicative comme dans *il lui donne*. Ces pronoms, quand ils ont une valeur de substantif personnel devraient toujours avoir la même catégorie, de préférence celle de substantif. Or ils sont tantôt considérés comme des noms, tantôt comme des pronoms. Dans le vers suivant, *lui* est analysé comme pronom, tandis que *vous* est analysé comme nom :

Sans en avoir **pour lui**, sans qu'il en ait **pour vous**,
Corneille, *Tite et Bérénice*, vers 58.

L'analyse rencontre les mêmes difficultés pour des mots qui peuvent appartenir tantôt à la catégorie d'adverbe, tantôt à celle de préposition ou de conjonction de subordination.

Que cette reine après se choisisse un époux.
(après : adverbe, analysé comme préposition par le métromètre)

Corneille, *Nicomède*, vers 350.

Lui, votre père ? Après son horrible dessein,
(après : préposition)

Racine, *Iphigénie*, vers 999.

Dans les deux vers précédents, *après* est analysé comme préposition, alors que dans le premier il aurait dû être analysé comme adverbe.

Globalement, l'analyseur peine sur certains mots qui selon les contextes peuvent correspondre à des catégories syntaxiques différentes, soit parce qu'il ne parvient pas à désambiguïser (comme dans le pronom *lui*), soit parce que toutes les catégories potentielles de la forme ne sont pas intégrées dans le lexique. Ainsi, *voici* et *voilà* n'admettent dans Sylex que la valeur de préposition, alors qu'ils ont en réalité gardé, de leur origine verbale, une valeur syntaxique particulière (certains les appellent «présentatifs»).

En second lieu, beaucoup d'erreurs proviennent d'une mauvaise détection des groupes verbaux interrogatifs ou exclamatifs. Comme le repérage de la bande verbale constitue une des premières étapes de l'analyse syntaxique, toute la phrase est réinterprétée en fonction d'une mauvaise analyse de départ qui conduit à des résultats aberrants. Ainsi chaque fois qu'une forme graphique ambiguë (nom ou verbe) est suivie d'un pronom personnel, l'analyseur syntaxique considère que le mot ambigu est un verbe et qu'il s'agit d'une phrase interrogative ou exclamative, dont une des caractéristiques peut effectivement être l'inversion du pronom sujet. Voici quelques exemples :

Il arrive, et / surpris il / change de visage :

Corneille, *Médée*, vers 1441.

Surpris il est identifié comme bande verbale, avec *surpris* comme verbe, et *il* comme pronom enclitique. Dans le vers suivant, *l'offense on* est interprété comme bande verbale exclamative, *offense* comme verbe avec *on* sujet postposé.

Si pourtant à /l'offense on /mesure la peine,

Racine, *Phèdre*, vers 605.

C'est la présence du pronom personnel qui entraîne les erreurs d'interprétations. En effet, dans cet autre vers, *offense* est bien analysé comme nom.

Mais il faut à l'offense opposer les bienfaits.

Racine, *Athalie*, vers 973.

Ces mauvaises analyses syntaxiques ont évidemment des répercussions sur les marquages que nous examinerons dans la section sur la limite des marquages.

VI.2.2.3. La métaposition

Les erreurs d'analyse que nous venons d'identifier sont souvent corrélées à une particularité de la syntaxe du vers classique, qui est, finalement, la source principale des erreurs syntaxiques du métromètre. En effet, le métromètre est démuni face à la *métaposition*, pour reprendre le terme proposé par Mitsou Ronat [1975]. Qu'est ce que la métaposition ? Le français est considéré comme une langue à ordre relativement fixe où la position des syntagmes et des catégories indique les relations de dépendance, les liens syntaxiques entre syntagmes, contrairement aux langues à déclinaisons où les relations de dépendance sont marquées morphologiquement, par exemple par des marques casuelles. Or, dans le vers classique, cet ordre est parfois bouleversé, créant des figures syntaxiques très rares dans la prose. On appelle métaposition le déplacement à gauche d'un syntagme prépositionnel (PP) qui peut être complément de phrase (S), de verbe (syntagme verbal : VP), de nom (syntagme nominal : NP) ou d'adjectif (syntagme adjectival : AP)[117]. Nous ne désignerons pas comme métaposition l'antéposition d'un adjectif par rapport au nom (*l'immonde objet* au lieu de *l'objet immonde*) car ce phénomène n'est pas propre au vers.

Ainsi sur les cinquante premiers vers de *Phèdre*, on en recense huit cas :

[117] Les symboles anglais ont été conservés pour la notation.

Qui sait même, qui sait si le roi votre père
Veut que [de son absence] on sache le mystère ?
PP NP

Racine, *Phèdre*, vers 17-18.

PP NA
[De ses jeunes erreurs] désormais revenu
[Par un indigne obstacle] il n'est point retenu ;
PP VP

Racine, *Phèdre*, vers 23-24.

Et fixant [de ses vœux] l'inconstance fatale,
PP NP
Phèdre [depuis longtemps] ne craint plus de rivale.
PP S

Racine, *Phèdre*, vers 25-26.

Cet heureux temps n'est plus. Tout a changé de face,
Depuis que [sur ces bords] les dieux ont envoyé
PP VP
La fille de Minos et de Pasiphaé.

Racine, *Phèdre*, vers 35.

Mais sa haine [sur vous] autrefois attachée,
PP AP

Racine, *Phèdre*, vers 41.

Remarquons au passage que cette particularité syntaxique du vers, si elle n'est pas expliquée, peut rendre la lecture des pièces classiques ardue. Il va sans dire que la pédagogie de la métaposition doit s'accompagner de celle de la métrique, qui seule permet de lever les ambiguïtés de sens, comme nous allons le voir.

Mitsou Ronat [1975b], dans un second article consacré à la métaposition, montre la nécessité de construire une règle de métaposition spécialement conçue pour l'alexandrin classique. En effet, selon Ronat, dans la théorie générative, aucune règle ne traite précisément de cette question, et la règle de SCRAMBLING, proposée par J. R. Ross [1967, cité par Ronat, 1975b], qui décrit les inversions en latin, est mal adaptée aux inversions dans le vers classique. Cette règle, SCRAMBLING, a pour effet, comme son nom l'indique, de permettre l'inversion des constituants mais aussi des éléments lexicaux. Syntagmes et mots peuvent être déplacés par rapport à leur position dans la structure

profonde[118]. Or le champ d'action de la métaposition en français est beaucoup plus étroit qu'il ne l'est en latin. Ronat relève trois différences entre les inversions dans le vers classique (métaposition) et les inversions latines (SCRAMBLING) qui justifient l'écriture d'une règle spécifique pour la métaposition. Tout d'abord, la direction : alors que dans SCRAMBLING le déplacement peut se faire vers la droite et vers la gauche, dans la métaposition, le déplacement ne peut s'effectuer que vers la gauche[119]. Ensuite, SCRAMBLING modifie d'une manière plus radicale l'arbre de la structure profonde parce que les éléments déplacés sont rattachés à la phrase ou à son nœud supérieur. Dans la métaposition, le syntagme prépositionnel, qui, en structure profonde, dépend de la tête de syntagme, est rattaché au même nœud que la tête : le déplacement entraîne une remontée dans l'arbre d'un cran. Enfin, la métaposition ne concerne dans le vers classique que les syntagmes prépositionnels, alors que dans SCRAMBLING, tous les syntagmes et éléments lexicaux sont concernés.

La métaposition est propre à l'alexandrin classique français et mérite d'être formalisée par une règle particulière. Mais quel est le rôle de cette bizarrerie syntaxique ? Pour Ronat, la métaposition a pour « rôle de renforcer la structure métrique » [1975, p. 11] : soit en renforçant la cohésion entre les deux hémistiches, soit en marquant davantage la fin de chaque hémistiche, les positions 6 et 12 dans l'alexandrin.

Mitsou Ronat [1975] distingue trois types de métapositions. Le premier est une inversion qui porte sur les deux hémistiches. Dans le cas le plus « pur », les deux éléments syntaxiques disjoints, syntagme prépositionnel et tête de syntagme, coïncident respectivement avec le premier et avec le second hémistiche.

---

[118] Ronat utilise la terminologie chomskienne de l'époque (années 70) qui est aujourd'hui un peu dépassée.

[119] Le PP déplacé laisse des traces là où il était et peut être le lieu d'une interprétation anaphorique. Pour que celle-ci puisse avoir lieu, il faut que l'antécédent (le PP) précède ou commande l'élément anaphorique [*cf.* Ronat, 1975b, p. 70].

[Par un indigne obstacle] il n'est point retenu ;

Racine, *Phèdre*, vers 24.

Sinon, il suffit que la fin du PP coïncide avec la fin du segment métrique et la fin de la tête de syntagme avec la fin du vers :

Qui sait même, qui sait si le roi votre père
Veut que [de son absence] on sache le mystère ?

Racine, *Phèdre*, vers 17-18.

Et fixant [de ses vœux] l'inconstance fatale,
Phèdre [depuis longtemps] ne craint plus de rivale.

Racine, *Phèdre*, vers 25-26.

.

Le second type correspond à des inversions à l'intérieur d'un hémistiche :

Pensez-vous qu'Hermione, [à Sparte] inexorable

Racine, *Andromaque*, vers 33.

C'est Vénus toute entière [à sa proie] attachée.

Racine, *Phèdre*, vers 306.

Enfin dans le troisième cas de métaposition la tête de syntagme est sur un hémistiche tandis que le complément est sur l'autre. Si les deux coïncident avec les fins d'hémistiches, on est ramené au premier cas, sinon « l'un des deux membres du syntagme disjoint doit se trouver à la fin d'un des deux segments métriques. » [p. 19]. Ronat écrit tenir cette hypothèse de Roubaud.

[De son fils], qu'il lui cache, il menace la tête,

Racine, *Andromaque*, vers 113.

Il est aussi des cas plus complexes de métaposition où plusieurs compléments sont antéposés :

Hélène [à ses parents] [dans Sparte] dérobée ;

Racine, *Phèdre*, vers 85.

Ronat a défini deux principes d'explication du rôle de renforcement métrique de la métaposition :

1) Le Principe de redondance M-P-S (métrico-phono-syntaxique). Pour Ronat [1975, p. 37], la structure métrique se trouve renforcée s'il y a concordance entre des marquages métriques et phono-syntaxiques. La syntaxe chomskienne distingue dans le syntagme la tête (X) et le complément (Comp). Ce dernier est dans une relation de dépendance par rapport à la tête. Ronat reprend la distinction syntaxique entre tête et complément pour l'appliquer à la

structure métrique. Partant du constat que le second hémistiche est plus « important » que le premier - parce qu'il porte la rime[120] -, elle considère par analogie le second hémistiche comme la tête de syntagme et le premier comme complément, et pose l'existence de dépendance similaire entre le premier et le second hémistiche. Dans la structure métrique, il y a une inversion systématique entre tête (second hémistiche) et complément (premier hémistiche) toujours placé avant. La structure métrique peut être considérée comme une forme de métaposition métrique permanente. Quand, dans la structure syntaxique, le complément est placé avant la tête de syntagme, autrement dit quand il y a métaposition, il y a redondance entre la structure métrique et syntaxique, et donc renforcement de la structure métrique.

2) La Règle des trois # et la règle des traces. Ces deux principes permettent de rendre compte du rôle de renforcement tonique des inversions dans le vers classique.

Ronat propose la règle des trois #[121] qui dit que si, dans une structure syntaxique, se trouvent trois # (indice de frontière devant et après catégorie de mot, de syntagme, de phrase) consécutifs, la voyelle pleine qui les précède immédiatement reçoit un renforcement tonique. La convention des traces proposée par Selkirk [1972, cité par Ronat 1975] indique que les transformations qui suppriment ou déplacent des constituants ne déplacent ni ne suppriment les frontières de mots associées à ces constituants. De plus, le syntagme déplacé s'insère avec ses frontières en sa nouvelle position.

Cette convention des traces entraîne un renforcement de la fin de la tête de syntagme, puisqu'en se déplaçant, le syntagme prépositionnel laisse les deux signes de frontières (##) qui l'entouraient, qui rajoutés à la frontière de la tête forment trois signes de frontière (###). On obtient donc un renforcement de la position qui les précède.

---

[120] Et nous pouvons ajouter qu'il est rythmiquement plus régulier que le premier (*cf.* chapitre VII.4).

[121] Dans la tradition chomskienne, le signe # est une marque de frontière : tous les mots de catégories majeures et les syntagmes sont encadrés de #. Par exemple :

(#(ses)(#longs#)(#mugissements#)#)

NP Det A N NP

Ronat montre, à partir de l'examen d'un corpus de 25000 vers, que dans 70 % des cas de métaposition la fin du syntagme prépositionnel coïncide avec la fin du premier hémistiche et la fin de la tête de syntagme avec la fin du second hémistiche. Dans ce cas, il y a renforcement du mètre par redondance entre les marquages syntaxiques et métriques et par renforcement tonique à la rime (convention des traces et 3 #).

Les seconds cas de métaposition, qui ont lieu à l'intérieur d'un hémistiche, correspondent à 29,7 % des cas : 29,5 % à l'intérieur du second hémistiche, ce qui entraîne un renforcement tonique de la rime ; 0,2 % à l'intérieur du premier hémistiche.

Enfin, dans 0,3 % des cas, on a simplement une dépendance à la césure : tête de syntagme sur le second hémistiche et syntagme prépositionnel sur le premier, la fin de l'un des deux coïncidant avec la fin d'un des deux hémistiches.

Dans tous les cas de métaposition, au moins l'un des deux éléments disjoints coïncide avec la fin d'un des deux segments métriques. La métaposition renforce le mètre et elle « ne peut se produire que dans un mètre qu'elle renforce » [1975, p. 31]. En effet, « les contraintes métriques assurent l'interprétabilité des séquences, elles seules permettent la remise en ordre de la lecture faite par le lecteur-auditeur » [1975, p. 31]. La métaposition est favorisée dans le vers, mais pas dans la prose, parce que la structure métrique du vers permet de reconstituer l'ordre « normal ». D'où, à l'inverse, la contrainte qui impose la coïncidence avec la fin d'un des deux segments métriques.

Nous n'avons pas étudié la question de la métaposition de manière approfondie, mais il semble que le principe de redondance métrico-phono-syntaxique, qui dit que l'inversion syntaxique renforce l'inversion métrique des hémistiches, donne un appui à la théorie de la congruence entre prosodie et syntaxe. Pourtant ce principe est loin d'aller de soi : pourquoi la rime donnerait un statut de tête au second hémistiche, pourquoi considérer qu'il y a inversion dans la structure métrique ? La théorie des trois # et la convention des traces

permettent à elles seules de montrer que la structure inversée renforce la structure métrique.

Il nous reste à comprendre pourquoi la métaposition n'est pas systématiquement appliquée. Ainsi, toujours dans les cinquante premiers vers de *Phèdre*, on trouve de nombreux cas de syntagmes prépositionnels qui auraient pu être déplacés, sans la moindre incidence sur le décompte (avec des effets sur la rime évidemment) :

> Je commence à rougir [de mon oisiveté]
> Depuis plus de six mois éloigné [de mon père],
>
> Racine, *Phèdre*, vers 4-5.

Pourquoi dans des contextes métriques qui le permettent, la métaposition n'est-elle pas systématiquement choisie ? Ceci nous fait relativiser l'hypothèse de renforcement métrique. Une hypothèse affaiblie peut être proposée pour rendre compte du caractère non-systématique de la métaposition : la métaposition est tolérée si elle renforce la structure métrique, c'est-à-dire si au moins l'un des segments coïncide avec la fin d'un des hémistiches.

D'autres facteurs président sans doute au choix de la métaposition. Il nous faudrait mener une étude de tous les syntagmes prépositionnels, et de leur position par rapport à la tête de syntagme pour éclairer cette question. Cela nous permettrait de mieux comprendre dans quels contextes elle est appliquée. Osons une hypothèse un peu triviale : il ne s'agirait pas ici d'une question de réception mais bien de production. Cette licence syntaxique n'offre-t-elle pas simplement un degré de liberté (par rapport au dénombrement et à la rime) dans l'écriture des vers ? Puisqu'en effet, il faut qu'il y ait concordance entre segments métriques et syntagmes de langue, ne faut-il pas un brin de jeu dans le système ?

Il se peut aussi, moins prosaïquement, que des effets stylistiques soient à l'œuvre : y a-t-il le même taux de métaposition dans toutes les pièces ? chez tous les auteurs ? Il est regrettable que dans son étude stylistique fouillée de la phrase de Racine, Robert Garrette [1995] n'ait pas traité la question de la métaposition.

Mitsou Ronat nous offre un cadre d'analyse de la métaposition et une règle syntaxique intéressants qu'il nous faudra adapter à notre formalisme syntaxique. La difficulté provient du caractère facultatif de la règle, sans que l'on connaisse exactement les principes de cette variation. Pour le moment, comme l'ordre des mots intervient dans l'analyseur syntaxique du métromètre, celui-ci, quand l'ordre est inversé, se retrouve dans l'incapacité d'identifier les liens entre syntagmes, et il se trompe même dans l'identification des syntagmes. Si nous reprenons le vers suivant de *Phèdre* :

C'est Vénus toute entière à sa proie attachée

Racine, *Phèdre*, vers 363.

Aucun lien n'est identifié entre *attachée* et *Vénus*, alors qu'il suffit de rétablir l'ordre linguistique pour que le lien soit identifié :

C'est Vénus toute entière attachée à sa proie
`Vénus toute entière'(1) -<- a2 head -<- `attachée'(10)

Les erreurs peuvent être plus graves encore, quand la métaposition conduit à une mauvaise analyse de la bande verbale, comme dans le vers suivant :

[De ce qu'il fait pour /vous] dois-je/ être le salaire,

Corneille, *Le Cid*, vers 1810.

Où l'analyseur identifie deux bandes verbales "il fait" (correcte) et "vous dois-je" (incorrecte). Ce même vers mis « dans l'ordre », il cesse alors d'être un alexandrin, est analysé correctement :

*Dois-je être le salaire de ce qu'il fait pour vous.

Deux bandes verbales sont correctement identifiées : "Dois-je" et "il fait".

La métaposition devra donc être analysée de façon systématique pour introduire dans le métromètre une couche de règles spécialement dédiée à cette question.

Les difficultés d'analyse syntaxique posées par la métaposition auraient pu être en partie levées si nous avions pris en compte les marquages métriques, c'est-à-dire la position des fins de segments métriques. Comme en effet, l'enjambement d'un hémistiche à l'autre et plus encore d'un vers à l'autre est très rare dans le vers classique, comme il y a effectivement concordance entre

les fins de syntagmes et les fins d'hémistiches, la connaissance de la position de la césure permettrait d'introduire une marque de frontière qui aiderait l'analyseur syntaxique dans son interprétation. En particulier la question des bandes verbales mal détectées devrait disparaître. Si nous reprenons l'exemple précédent :

> De ce qu'il fait pour vous # dois-je être le salaire,
>
> Corneille, *Le Cid*, vers 1810.

l'introduction d'une frontière métrique après *vous* empêcherait le métromètre de considérer *vous dois-je* comme une bande verbale, puisque le pronom rattaché indûment au verbe en serait séparé par un signe de frontière métrique. Comme toutefoisnous voulons montrer que le rythme ne résulte que d'une habile organisation des phénomènes de langue, nous n'avons pas introduit ce type de marquage de segment métrique.

VI.2.3. Les limites des marquages

Les erreurs d'analyse syntaxique n'ont qu'un faible impact sur la représentation phonétique du vers : les cas d'ambiguïté du type *fier* (adjectif ou verbe), *content* (verbe ou adjectif) étant finalement assez rares. Cependant, les difficultés d'analyse des syntagmes rendent fragile la représentation de la liaison, plus précisément la distinction entre liaison forte (« directe ») et faible (« indirecte »).

En dehors du niveau phonétique, ces erreurs d'analyse syntaxique ont plusieurs effets négatifs sur le métromètre. Elles entraînent des erreurs de marquage, et nous avons même dû renoncer à certains marquages comme ceux de syntagmes et de groupes accentuels en raison du taux d'erreur.

VI.2.3.1. Des erreurs de marquage

Les erreurs d'analyse syntaxiques peuvent se répercuter sur les marquages de catégorie syntaxique, de fin de mot et par conséquent « d'accent », puisque ce dernier marquage dépend des deux premiers (et de la place du *e* muet).

Ainsi, quand des adverbes sont analysés comme prépositions, ou des pronoms à valeur prédicative comme des pronoms et non comme des noms,

cela entraîne des erreurs de marquage syntaxique qui se répercutent sur le marquage accentuel.

Voici un exemple :

Sans en avoir **pour lui**, sans qu'il en ait **pour vous**,

Corneille, *Tite et Bérénice*, vers 58.

cs	7	7	0	0	7	**6**	7	6	6	1	7	**0**
fdm	1	1	0	1	1	1	1	1	1	1	1	1
ma	0	0	0	1	0	**0**	0	0	0	1	0	**1**

(6= pronom, déterminant ; 0=nom)

Lui est codé comme déterminant, *vous* comme nom. Du coup *lui* n'est pas accentué (à tort).

Ensuite, la mauvaise analyse syntaxique de certaines « bandes verbales » (mauvaise reconnaissance des formes impératives ou exclamatives) conduit à des erreurs de marquage syntaxique, et à une insuffisance de marquage de fin de mot (seule la fin de la bande verbale est marquée) et des erreurs de marquages accentuels. Dans le vers suivant, on n'a pas de fin de mot pour la dernière syllabe de *surpris*, un accent en trop sur *il*.

Il arrive, et / surpris il / change de visage :

Corneille, *Médée*, vers 1441.

cs	6	1	1	8	1	1	6	1	1	7	0	0
fdm	1	0	1	1	0	**0**	1	0	1	1	0	1
ma	0	0	1	0	0	1	**1**	1	0	0	0	1

Ce vers offre une situation similaire au précédent :

Si pourtant à /l'offense on /mesure la peine,

Racine, *Phèdre*, vers 605.

cs	9	3	3	7	**1**	**1**	6	1	1	1	6	0
fdm	1	0	1	1	0	**0**	**1**	0	0	1	1	1
ma	0	0	1	0	0	1	**1**	0	1	0	0	1

(1 : verbe)

Ici, la bande verbale *voulez-vous désormais* est considérée comme un seul bloc, il n'y a pas de fin de mot indiquée sur *vous :*

Ciel, à qui /voulez-vous désormais/ que je fie

Corneille, *Cinna*, vers 1121.

cs	0	7	9	1	1	6	3	3	3	9	6	1
fdm	1	1	1	0	0	0	0	**0**	1	1	1	1
ma	1	0	0	0	0	1	0	0	1	0	0	1

De même dans le vers suivant, le métromètre identifie comme bande verbale interrogative *vous a-t-elle tant* ce qui entraîne une mauvaise analyse du pronom *vous* qui est dans le vers rattaché à *pour* et un mauvais marquage de fin de mot : aucun marquage de fin de mot n'est fait à l'intérieur de la bande verbale. De plus, pour la marque accentuelle, il y en a une d'office en fin de bande verbale, ce qui dans ce cas précis est incorrect (accent sur *tant*) *:*

La couronne pour /vous a-t-elle tant/ de charmes ?

Racine, *La Thébaïde*, vers 72

cs	6	0	0	0	7	/6	1	6	6	3/	7	0
fdm	1	0	0	1	1	/0	0	0	0	1/	1	1
ma	0	0	1	0	0	/0	0	1	0	1/	0	1

C'est en particulier pour ce genre de raisons que nous n'atteignons pas dans le vers classique un taux de marquage de la dernière position de 100 %.

A-t-on les moyens d'évaluer le taux d'erreur du métromètre ? Est-il rédhibitoire ?

L'analyse des positions 6 et 12 nous permet d'évaluer en partie le taux d'erreur. Normalement dans le vers classique, les positions de fin de segments métriques devraient systématiquement admettre un marquage de fin de mot, puisqu'aucun mot ne traverse l'hémistiche, ni n'enjambe la fin du vers. C'est le cas pour la douzième position, d'une manière mécanique puisque l'unité d'analyse est le vers. Sur la sixième position, le taux de marquage de fin de mot, nous le verrons dans le chapitre VII, est de 99,7 % pour Corneille et Racine. Le taux d'erreur (0,3 %) est donc très faible.

Nous avons également pu évaluer le taux d'erreur en matière de catégorisation syntaxique. Sur la douzième position, le métromètre a repéré 950 mots-outils sur le corpus CORRAC, alors qu'il n'y en a en réalité que 139, dont

131 sont des pronoms conjoints post-posés, autorisés parce qu'ils portent l'accent en douzième position. Sur la dernière position on compte donc 811 erreurs sur un total de 74916[122] vers, soit un taux d'erreur de 1 %. Sur la sixième position, le taux d'erreur est sensiblement plus élevé : il approche des 3 % : 3243 mots-outils sont repérés par le métromètre, alors que, vérification faite, il n'y en a que 935 qui sont effectivement des mots-outils.

Dans ces erreurs, sont comptés les pronoms personnels à valeur substantive qui auraient dû être analysés comme des noms. Ce sont eux qui sont à l'origine de l'essentiel des erreurs. Ces pronoms personnels semblent être plus fréquents en sixième qu'en douzième position, à moins que les erreurs d'analyse soient plus fréquentes en sixième. Dans de très nombreux cas, ces pronoms à l'hémistiche sont dans des situations de métaposition :

Fasse pleuvoir [sur elle] un déluge de feux !

Corneille, *Horace*, vers 1314.

Il faudrait explorer plus précisément cette question des substantifs personnels en fin d'hémistiche et de vers.

A l'heure actuelle, à supposer que le taux d'erreur sur l'hémistiche se retrouve sur les autres positions (hormis la douzième), on aurait un taux d'erreur de catégorisation syntaxique de 3 %, qui est loin d'être négligeable. C'est ici l'analyse syntaxique qui est en cause et qui devra être améliorée. Il faudra en particulier rajouter une couche de règles pour traiter les cas de métaposition.

Comme le marquage accentuel est construit à partir du marquage de fin de mot et de la catégorisation syntaxique, le taux d'erreur pour le marquage accentuel est au moins du même ordre de grandeur.

VI.2.3.2. Des marquages qui n'ont pu être mis en œuvre

Comme il nous était impossible d'avoir une représentation complète de la structure syntaxique, nous avons dû renoncer à un marquage des syntagmes et de leur hiérarchie et, simultanément, à un marquage des groupes accentuels

[122] Vers reconnus comme alexandrins par le métromètre. L'écart avec les effectifs totaux tient aux vers découpés de Corneille qui n'ont pas été reconstitués.

avec hiérarchie des accents. A condition de supposer une congruence complète entre syntaxe et prosodie, la position et l'intensité des accents peuvent être déduits de la structure syntagmatique et de la position des *e* muets finaux.

Cela est regrettable à plus d'un titre. Les marquages de syntagmes nous auraient permis d'identifier clairement les phénomènes d'enjambement, c'est-à-dire le fait que certains syntagmes traversent le vers et l'hémistiche, et de dresser une typologie des phénomènes de discordance entre segment métrique et syntagme. Avec le marquage de catégorie syntaxique, nous pouvons certes repérer une partie des enjambements. En effet, quand la sixième ou douzième positions sont occupées par un mot grammatical, ces derniers correspondent généralement à un début de syntagme. Mais cette heuristique ne permet pas de repérer tous les cas d'enjambements, comme celui de ces vers :

> Serait-ce déjà lui ? C'est bien à l'escalier
> Dérobé. Vite, ouvrons.
>
> Bonjour, beau cavalier.
>
> Hugo, *Hernani*, vers 1-2.

Ensuite, si nous avions pu introduire des hiérarchies d'accents dans les marquages, nous aurions des représentations des vers aux reliefs plus contrastés qu'ils ne le sont avec un simple marquage d'accent de mot. En effet quand pour marquage accentuel, on a une suite de deux ou trois « 1 », très souvent ces « 1 » pourraient avoir des valeurs différenciées. Ainsi, dans le vers suivant, les accents auraient gagné à être hiérarchisés : l'accent sur *mis* devrait être le plus fort (car suivi d'une frontière syntaxique forte), suivi de l'accent sur *prêts*, puis de l'accent sur *tout*, qui devrait être quasiment inexistant :

> De ses amis, tout prêts à prendre sa querelle,
>
> Racine, *Britannicus*, vers 1348

sm	d ə	s e	ɹz a	m i	t u	p r ɛ ɹz	a	p r ɑ̃	d r ə	s a	k ə	r ɛ l
cs	7	6	0	0	3	2	7	1	1	6	0	0
fdm	1	1	0	1	1	1	1	0	1	1	0	1
ma	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	0	1	0	0	0	1

Certes le métromètre produit des analyses erronées de certains vers, mais la limite principale de l'outil, plus précisément du module d'analyse syntaxique, est de ne pas nous avoir permis de procéder à des marquages plus fins que ceux que nous avons adoptés.

VI.2.3.3. Des marquages oubliés : la ponctuation et les *e* muets élidés

Avec du recul, puisque nous n'avions pas accès au marquage de syntagme, il aurait été souhaitable d'introduire un marquage de ponctuation, comme approximation grossière des marquages syntagmatiques. Nous y avons pensé trop tard et l'introduire aurait demandé d'importantes modifications dans le métromètre : il aurait fallu prendre la phrase réelle comme objet d'analyse et projeter ensuite les résultats dans l'espace défini par le vers. Le fait que le vers ait été choisi comme unité d'analyse n'est pas une grande source d'erreur, comparée à ce que peuvent être la mauvaise analyse des pronoms personnels à valeur substantive et de la métaposition. Comme le taux d'erreur est suffisamment faible pour que l'on puisse faire des analyses statistiques robustes, nous avons préféré renvoyer ces modifications à la prochaine version du métromètre, pour nous concentrer sur l'exploitation des résultats.

Pour le traitement du *e* muet nous avons choisi de ne marquer que deux catégories de *e* muets : le *e* muet fort (qui correspond principalement aux monosyllabes) et le *e* muet faible (qui correspond à des *e* muets comptés, qui pourraient s'élider en langue). En revanche, nous n'avons, dans le métromètre, gardé aucune trace des *e* muets élidés. Or le *e* muet que nous avons élidé en fin de vers joue un rôle métrique puisqu'il distingue les rimes féminines des rimes masculines et qu'il est un marqueur de l'alternance en genre. Pour l'étude de la rime, il nous a fallu d'ailleurs le réintroduire.

Dans les traités jusqu'au XVe siècle, comme le montrent les citations de traités dans l'article de Roubaud [1977], on employait le terme de synalèphe, plutôt que celui d'élision, qui semble indiquer qu'il restait une trace de cet *e*.

Par ailleurs, si l'on pose comme hypothèse que le *e* muet même s'il n'est pas compté a des effets phonologiques (allongement de la voyelle qui les précède, transformation de sa prononciation...), il aurait été intéressant de garder la position de ces *e* muets pour mesurer leur rôle dans la construction du rythme.

Supposons que les syllabes saillantes du premier hémistiche soient la troisième et la sixième. Si les positions 3 et 6 sont suivies de *e* muets élidés, la structure rythmique du vers se trouve renforcée par l'allongement de la voyelle induit par la présence du *e* muet. Garder la trace des *e* muets nous aurait permis de voir si ce procédé de renforcement métrique est constitutif de l'écriture de Racine[123].

VI.2.4. Pas de règles pour l'identification d'un « bon » alexandrin

Pour le moment, le métromètre découpe en positions métriques selon les règles classiques tout paragraphe de texte qui lui est proposé. Ainsi un vers libre de Supervielle est analysé comme une séquence de 19 positions métriques :

Comme s'il y avait alentour une grande plaine, la liberté,

Supervielle, *Gravitations*, Le survivant

sm	kɔ	məf	sil	i	a	vɛlt	a	lɑ̃	tur	y	nəf	grɑ̃	dəf	plɛ	nəf	la	li	bɛr	te
nb	19																		
cs	9	9	6	6	1	1	3	3	3	6	6	2	2	0	0	6	0	0	0
fdm	0	1	1	1	0	1	0	0	1	0	1	0	1	0	1	1	0	0	1
ma	0	0	0	0	0	1	0	0	1	0	0	1	0	1	0	0	0	0	1

Le métromètre sait nous indiquer le nombre de positions métriques que contient la séquence qu'il analyse, mais il ne peut indiquer s'il s'agit d'un véritable alexandrin. Nous aimerions que l'outil soit capable d'identifier qu'un alexandrin est conforme au modèle classique, ou qu'au contraire il ne respecte pas certaines règles. Et dans ce dernier cas, nous aimerions qu'il nous indique quelles sont les règles qui ne sont pas respectées. Récapitulons ces règles :

- absence de *e* muet en positions 6 et 12
- absence de *e* muet faible en positions 1 et 7
- absence de hiatus externe dans le vers
- absence de séquence Ve#C dans le vers

[123] Cette hypothèse nous a été proposée par Madame Sipos, professeur de Khâgne au lycée Henri IV.

- les positions 6 et 12 correspondent à la dernière syllabe accentuée d'un mot
- la position 6 peut être suivie d'un *e* muet faible élidé ou synaléphé
- la position 12 peut être suivie d'un *e* muet faible même s'il ne peut être élidé
- la fin du premier segment métrique et la fin du vers coïncident avec une frontière syntaxique majeure

Dans ce travail, nous avons exploré tous les vers qui ne respectent pas l'une de ces règles à partir de requêtes dans la base de données des vers. Il serait plus simple que ces mesures soient faites en amont et que les violations de règles apparaissent comme des marquages spécifiques, ou au moins qu'il y ait un indice de conformité du vers au modèle. On pourrait, comme le proposaient Halle et Keyser dans leur ouvrage sur Chaucer [1966], fabriquer un indice de complexité du vers ou, comme le proposait Youmans [1989], une distance au prototype.

On souhaiterait, par exemple, que le métromètre nous indique que le vers suivant

> Dieux ! que ne suis-je assise à l'ombre des forêts !
>
> Racine, *Phèdre*, vers176.

est parfaitement conforme au modèle classique, tandis que celui de Corneille :

> Les sœurs crient miracle, et chacune ravie
>
> Corneille, *Médée*, vers 73.

s'en écarte en introduisant une séquence Ve devant consonne ; et celui de Malherbe :

> <u>Pendu ores</u> au croc, triste et muet se deut,
>
> Malherbe, *Poésies*, CVI Larmes du Sieur Malherbe.

s'en écarte en introduisant un hiatus. Et que l'absence de frontière de mot à la césure (*Péninsules* enjambe les deux hémistiches) est une violation plus forte encore du modèle classique.

> Je courus ! Et les Péninsules démarrées
>
> Rimbaud, *Poésies*, « Le Bateau ivre ».

Même si ces trois vers sont effectivement constitués de 12 positions métriques selon les canons classiques, ils sont plus ou moins conformes au

modèle. Et c'est cette variation que nous souhaiterions prendre en compte dans le métromètre.

Par ailleurs, nous aimerions aussi (dans un avenir plus lointain) disposer d'un outil qui soit capable d'identifier des alexandrins classiques dans la prose. Ce repérage d'alexandrins classiques ne pourrait être résolu que par l'intégration de l'ensemble des règles de constitution de l'alexandrin au métromètre. Mais cela ne serait pas suffisant. Nous avons précédemment vu que le métromètre, et c'est une de ses limites, analyse la ligne qui lui est présentée, autrement dit, il analyse un paragraphe de prose comme s'il était un unique vers. Pour reconnaître des vers classiques dans de la prose, il nous faudrait, sur la base de l'analyse du paragraphe faire glisser une fenêtre de 12 positions accompagnée de l'ensemble des règles le long du paragraphe, cette fenêtre s'arrêterait pour marquer un alexandrin, là où toutes les règles sont remplies. Nous verrions sans doute que les vers blancs sont très significativement absents de la prose du XVIIe siècle, et que les auteurs prennent grand soin de les éviter.

Récapitulons : le métromètre s'appuie sur un analyseur syntaxique et un phonétiseur. Ces outils ont été adaptés aux particularités du décompte du vers classique. Le métromètre analyse chaque vers, propose une ou plusieurs interprétations syntaxiques, transcrit le vers dans l'alphabet phonétique international selon les positions métriques. Ensuite, le métromètre attribue à chacune des positions des marquages phonétiques (syllabe et noyau vocalique), lexical (présence ou non d'une frontière de mot après la syllabe), morpho-syntaxique (catégorie syntaxique de la position) et accentuel (présence ou non d'un accent lexical sur la position).

Si le métromètre est déjà un outil extrêmement précieux, il n'est pas encore parfait. L'analyseur syntaxique, très performant pour l'analyse de la prose, rencontre des difficultés quand il s'agit du vers, à cause de la métaposition (antéposition d'un syntagme prépositionnel), mais aussi parce que l'analyse porte sur le vers au lieu de porter sur la phrase entière. Ces difficultés d'analyse limitent les possibilités de marquage et rendent difficile par exemple la mise au

point d'un marquage hiérarchisé des accents selon leur position dans les groupes de souffle. Des erreurs plus ponctuelles, comme le traitement erroné de certains pronoms à valeur substantive, entraînent par leur fréquence des biais dans les marquages syntaxiques.

La transcription phonétique réalisée par le phonétiseur est quant à elle parfaitement adaptée au cas du vers puisqu'elle a été conçue, testée et modifiée en fonction du vers. Cependant deux améliorations pourront être apportées. La première consisterait à conserver, à l'aide d'un nouveau marquage, la trace des *e* muets élidés. Ceci permettrait de distinguer immédiatement les rimes féminines des rimes masculines et d'évaluer quelles sont les positions qui privilégient l'élision du *e* (on pense bien évidemment à la sixième position). La seconde amélioration porterait sur la modélisation des cas de diérèse et de synérèse. Les suites de deux voyelles HV dont la première est une voyelle haute H (*i, u, ou*) sont actuellement représentées en (...H)(V...) en cas de diérèse et (...GV...) en cas de synérèse, où G est la glissante qui résulte de la semi-vocalisation de la voyelle haute (/j/, /ɥ/, /w/). Or l'analyse des couples de mots-rimes attestés montre que les cas de terminaison avec diérèse qui riment avec des terminaisons sans diérèse ne sont pas rares, et qu'en revanche les cas de terminaisons HV qui riment avec des V sont très rares. Introduire pour les cas de diérèse la semi-voyelle dans la dernière position métrique, (...H)(GV...), permettrait d'enrichir la description de la rime avec la semi-voyelle qui précède le noyau vocalique et de maintenir l'hypothèse, valable pour le vers classique, que la rime s'étend rarement au-delà de la dernière position métrique.

L'introduction de nouveaux corpus, nous avons pu en faire l'expérience avec Hugo, Baudelaire, Rimbaud et Mallarmé (*cf.* Annexe VI), amène un vocabulaire nouveau et des arrangements spécifiques de termes qui nécessitent des modifications du métromètre, en particulier dans le lexique. Cela concerne principalement les mots nouveaux avec diérèse (en premier lieu les noms propres) et certains mots commençant par un *h* aspiré non encore répertoriés dans les lexiques du métromètre.

La question de la disponibilité de l'outil mérite quelques remarques. Le métromètre s'appuie sur un analyseur syntaxique, propriété intellectuelle de l'ENST, commercialisé par la société Ingénia. Le phonétiseur, développé dans le même laboratoire de l'ENST, est un outil de travail interne, mais il n'est pas

diffusé dans la communauté scientifique, en premier lieu parce qu'il s'appuie sur l'analyseur syntaxique. Quand aux couches de règles correspondant au métromètre, nous les avons totalement intégrées dans l'outil. Autrement dit, aujourd'hui, il n'existe qu'une version du métromètre dans le laboratoire informatique de l'ENST. Aux problèmes liés à la nature composite de l'outil et aux enjeux stratégiques de l'analyseur syntaxique, produit industriel, s'ajoute la question des droits sur les corpus. Les textes provenant de l'INaLF sont sa propriété et ne sont pas pour le moment en libre accès. Ce n'est que grâce à la courtoisie de Robert Martin que nous avons pu en disposer pour ce travail.

Cependant, les « briques logicielles » permettant de fabriquer un tel outil seront de plus en plus faciles à trouver, de même que les corpus informatisés. Nul doute que d'autres pourront reproduire et continuer le type de travail que nous avons réalisé en construisant des métromètres un peu différents. Nous pensons en tout cas avoir décrit la manière dont nous avons adapté les outils de linguistique informatique au cas spécifique du vers, ainsi que les choix qui devraient être faits pour que l'outil soit plus performant, de manière suffisamment explicite pour être directement réutilisable dans tout autre projet analogue. Nous espérons que les règles formalisées pour la diérèse, le *e* muet et la liaison seront utiles. Ce travail fournit donc les bases pour améliorer le métromètre actuel ou pour en créer un nouveau.

VII. SON, SYNTAXE ET PROSODIE

DE L'ALEXANDRIN « ORDINAIRE »

Nous avons décrit dans le chapitre IV les procédures permettant d'identifier les positions métriques, dans le chapitre V les marquages choisis pour décrire le vers et dans le chapitre VI le fonctionnement du métromètre. Nous pouvons donc maintenant enfin explorer les caractéristiques des alexandrins avec ces instruments.

Quelles sont les différentes composantes rythmiques de la « figure » de l'alexandrin chez Corneille et Racine ? Nous appelons figure la représentation multidimensionnelle du vers construite grâce à la projection sur ce vers de catégories de la langue (phonèmes vocaliques, catégories syntaxiques, accents...). Cette démarche est conforme à l'hypothèse selon laquelle le vers puise dans la langue toutes ses ressources et qu'il informe la matière de la langue pour tendre vers son propre modèle métrico-rythmique.

Nous examinerons chacun des aspects de la figure du vers (phonétique, lexical, syntaxique et prosodique), lesquels sont autant de composantes du rythme total. On pourrait même dire, à l'image de la formule de Mauss en sociologie, que le rythme est un fait linguistique total qui mobilise tous les aspects de la langue (lexique, syntaxe, prosodie, sémantique...) et qui s'en nourrit.

Sur l'ensemble des alexandrins du corpus, près de 75 000, nous explorerons la répartition sur chaque position de toutes les modalités qu'admettent les marquages. Nous profiterons de l'outil pour examiner également les phénomènes rares sur notre corpus. La construction de ces figures moyennes du

vers (obtenues par des représentations statistiques d'un ensemble de vers) permettra ensuite (comme nous en donnons l'exemple au chapitre VIII) d'explorer les caractéristiques formelles d'un auteur, d'un genre ou d'une pièce par rapport au modèle moyen.

VII.1. LES VOYELLES METRIQUES DANS LE VERS

La voyelle métrique est le cœur de la position métrique. Elle est son élément constitutif indispensable[124], voire suffisant. Sont dites « métriques » les voyelles qui comptent dans le vers. Par exemple, dans le vers suivant,

> Et je perdais ma gloire à demeurer ici.
> (e)(ɜ ə)(p ɛ r)(d ɛ)(m a)(g l w a r)(a)(d ə)(m œ)(r ɛ rɺ)(i)(s i)
>
> Racine, *La Thébaïde*, vers 58.

les douze voyelles métriques sont /e/, /ə/, /ɛ /, /ɛ/, /a/, /w a/, /a/, /ə/, /œ/, /ɛ/, /i/, /i/.

Comment ces voyelles occupent-elles les différentes positions métriques ? Sont-elles réparties au hasard ? Nous savons bien que non puisque déjà le *e* muet est « interdit » en sixième et en douzième positions dans l'alexandrin[125]. Mais qu'en est-il des autres voyelles métriques ?

La répartition des phonèmes vocaliques est-elle guidée par la structure syntaxico-métrique du vers, ou par d'autres motifs, comme par exemple un souci d'euphonie, en tant que configuration spécifique des phonèmes ?

Nous nous appuierons ici encore sur le corpus constitué des œuvres théâtrales complètes de Corneille et Racine pour construire la figure *phonétique* de l'alexandrin ordinaire[126]. L'analyse portera sur tous les vers de douze

---

[124] Cornulier [1995, p. 24] cite cependant un vers de *Cyrano de Bergerac*, où la première position métrique n'a pas de noyau vocalique :

« Pst !... Petits !... Rendez-moi le sonnet à Phillis. »

[125] Et il nous appartiendra de vérifier si cette interdiction est bien respectée dans la réalité.

[126] Lorsque le vers est complexe, l'analyseur syntaxique propose plusieurs interprétations et autant de transcriptions phonétiques. Nous ne retenons ici que la première, car les variations d'interprétation syntaxique n'ont qu'une faible incidence sur les phonèmes vocaliques. Cela

positions[127] (pour Corneille les vers coupés par des tours de parole sont exclus). Le corpus est constitué de près de 75 000 alexandrins.

Comme nous l'avons justifié dans le chapitre V.1, nous avons retenu dix-neuf types de voyelles métriques : /ə/, /əɟ/, /a/ ou /ɑ/, /ã/, /ɛ/, /e/, /i/, /ɛ̃/, /ɔ/, /o/, /ɔ̃/, /œ/, /ø/, /œ̃/, /u/, /ç/, /wa/, /wɛ̃/, /ɥi/.

Il s'agit d'abord des seize voyelles que l'on considère classiquement comme étant celles du français, en ne faisant toutefois pas la distinction entre le *a* antérieur et le *a* postérieur et en distinguant deux types de *e* muet :

- le *e muet fort*, toujours prononcé tant qu'il est graphiquement présent (le *e* des mots outils : *je*, *te*, *le*, noté /ə/ et le *e* obligatoire pour éviter la succession de trois consonnes) :

Où l'on voit l'Achéron se perdre chez les morts,

(u)(l ɔ̃)(v w a)(l a)(k e)(r ɔ̃)(s ə)(p ɛ r)(d r ə)(ʃ e)(l e)(m ɔ r)

Racine, *Phèdre*, vers 12.

- le *e muet faible*, toujours présent graphiquement mais susceptible d'élision en fonction de sa position (élidé devant une voyelle et en fin de vers), noté /əɟ/ :

---

arrive dans des cas peu fréquents, par exemple quand une terminaison en *-ent* peut être interprétée comme /ə/ ou /ã/. Ainsi, *content* peut-il être un adjectif ou un verbe (troisième personne du pluriel de *conter*). Dans l'exemple suivant, et c'est pratiquement le seul, *content* est mal analysé :

Un roi qui non content d'effrayer les mortels, (Racine, *Iphigénie*, vers 872.)

(œ̃)(r w a)(k i)(n ɔ̃)(k ɔ̃)(t əf)(d e)(f r e)(j e)(l e)(m ɔ r)(t ɛ l)

Si le métromètre était enrichi de quelques règles métriques, il n'aurait jamais produit cette analyse puisqu'un *e* muet ne peut jamais occuper la sixième position dans le vers classique. Mais c'est justement l'objet du métromètre que de découvrir de telles règles.

[127] Plus exactement tous les vers qui, d'après le métromètre, ont douze positions. Sur le corpus CORRAC, ces deux propositions sont pratiquement équivalentes. Comme le métromètre n'a cessé d'être amélioré sur la base de ce corpus, il n'y a qu'un pourcentage infime d'alexandrins non reconnus.

Phèdre, atteinte d'un mal qu'elle s'obstine à taire,
(f ɛ d r)(a)(t ɛ̃)(t ə̲ʃ)(d œ̃)(m a l)(k ɛ)(l ə̲ʃ)(s ɔ p)(s t i n)(a)(t ɛ r)

Racine, *Phèdre*, vers 45.

Nous avons également considéré comme des « voyelles métriques » les groupements [w ɑ], [w ɛ̃] et [ɥ i], dont les composants ne sont jamais dissociés. La rime porte, en effet, toujours sur le groupement complet et ces groupements ne donnent jamais lieu à une diérèse.

Notons enfin que la transcription phonétique distingue des phonèmes qui dans la langue courante d'aujourd'hui ne sont plus distingués : /œ̃/ et /ɛ̃/. Ils sont cependant clairement distincts dans le vers classique, qui ne les associe jamais à la rime.

Ces dix-neuf voyelles métriques retenues ont été déterminées empiriquement à partir de l'examen de la rime et de la diérèse. Il ne peut y avoir équivalence à la rime entre des voyelles métriques différentes (sauf d'une manière très exceptionnelle entre /i/ et /ɥ i/).

Les trois vers suivants, qui ont la particularité d'avoir pratiquement autant de voyelles différentes que de positions, présentent des exemples des dix-neuf voyelles métriques (la première occurrence de chaque voyelle est soulignée).

Et toutefois, O dieux, un crime involontaire
(e̲)(t u̲)(t ə̲ʃ)(f w̲ ɑ̲ zl)(o̲)(d j ø̲ zl)(œ̲̃)(k r i̲ m)(ɛ̲̃)(v ɔ̲)(l ɔ̲̃)(t ɛ̲ r)

Racine, *La Thébaïde*, vers 603.

Dont l'âme généreuse avait tant de douceur,
(d ɔ̃)(l a̲)(m əʃ)(ʒ e)(n e)(r ø z)(a)(v ɛ)(t ɑ̲̃)(d ə̲)(d u)(s œ̲ r)

Racine, *La Thébaïde*, vers 513.

Ou du moins pour jamais j'irai lui dire adieu,
(u)(d y̲)(m w̲ ɛ̲̃)(p u r)(ʒ a)(m ɛ)(ʒ i)(r ɛ)(l ɥ̲ i̲)(d i r)(a)(d j ø)

Racine, *La Thébaïde*, vers 150.

VII.1.1. Répartition des phonèmes vocaliques

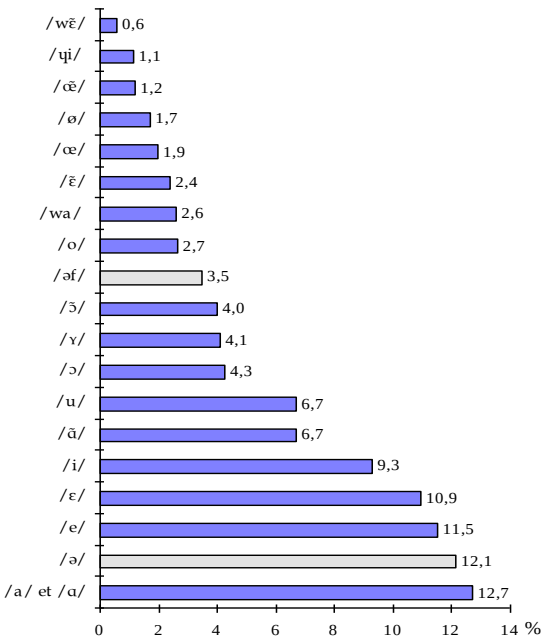
Quelle est la répartition globale des différentes voyelles métriques ? Ont-elles toutes la même fréquence ? La répartition est-elle la même dans la langue et dans le vers ? Sont-elles réparties au hasard sur les différentes positions ?

À ces questions voici quelques éléments de réponse. D'abord, la fréquence d'occurrence des différents phonèmes vocaliques varie dans une proportion de 1 à 21, entre le moins fréquent /w ɛ̃/ et le plus fréquent /a/ (*cf.* Figure 22).

Si l'on ne distingue pas les deux types de *e* muet, celui-ci est la voyelle la plus fréquente, puisqu'il représente 15,6 % des voyelles métriques. Notons que le *e* muet faible, celui qui compte ou ne compte pas selon le contexte, ne représente que 3,5 %[128]. Viennent ensuite le /a/, le /e/, le /ɛ/ et le /i/ qui représentent chacun plus de 9 % des phonèmes vocaliques. Les voyelles nasales représentent 14,8 % des voyelles et les voyelles métriques avec diphtongue 7,2 % des voyelles métriques.

[128] Il faudrait pouvoir évaluer la quantité de *e* muets élidés dans le vers.

Phonème vocalique		(%)
/ə/	(le, que)	12,1
/əf/	(blême, flamme)	3,5
/a/, /ɑ/	(courage, âme)	12,7
/ɑ̃/	(puissance, amant)	6,7
/ɛ/	(père, même)	10,9
/e/	(hyménée, parler)	11,5
/i/	(vie, empire)	9,3
/ɛ̃/	(rien, main)	2,4
/ɔ/	(mort, personne)	4,3
/o/	(autre, chose)	2,7
/ɔ̃/	(raison, honte)	4,0
/œ/	(cœur, pleurs)	1,9
/ø/	(yeux, lieu)	1,7
/œ̃/	(communs, importuns)	1,2
/u/	(vous, amour)	6,7
/y/	(vue, plus)	4,1
/wa/	(moi, gloire)	2,6
/wɛ̃/	(point, moins)	0,6
/ɥi/	(lui, ruine)	1,1
Total		100



Clef de lecture : 12,7 % des positions métriques ont comme voyelle métrique un /a/ ou un /ɑ/. A droite, les voyelles sont classées par ordre de fréquence, à gauche dans un ordre que nous retiendrons tout au long de ce travail.

Figure 22. Répartition des positions métriques selon les phonèmes vocaliques

Comparons maintenant la répartition des phonèmes vocaliques en vers et en prose. Pour la prose, les variations selon le locuteur et le mode d'énonciation sont sans doute plus importantes que pour le vers, dans la mesure où aucun cadre, tel que le modèle métrique, n'est présent. Pour établir une comparaison de la répartition des phonèmes vocaliques entre vers et prose, nous avons repris les travaux de F. Wioland [1985] qui a transcrit lui-même vingt échantillons de 10 000 phonèmes tirés d'émissions à la télévision. Dix échantillons proviennent d'émissions-débats, où plusieurs personnes dialoguent ensemble, dix autres de pièces de théâtre en prose diffusées à la télévision.

Cette comparaison est loin d'être pertinente, puisque pour la prose la transcription s'appuie sur des productions orales concrètes, tandis que pour le vers la transcription reflète une représentation abstraite de ce que pourrait être

la diction du vers. Or, entre cette représentation abstraite et la réalité de la diction, il peut y avoir de sérieux décalages. Ainsi, la répartition des phonèmes dans *Horace* doit-elle être sensiblement différente si l'on considère la mise en scène de Jean-Pierre Vincent [1971-1972], où les *e* muets comptés sont souvent élidés et les diérèses à peine sensibles, ou celle de Marion Bierry [1998], très proche du modèle abstrait que nous avons construit.

Nous avons adopté, pour cette comparaison, la catégorisation des phonèmes vocaliques de F. Wioland [1985]. Nous avons donc regroupé les deux types de *e* muets, rattaché les /w a/ à /a/, les /w ɛ̃/ à /ɛ̃/ et les /ɥ i/ à /i/. Ces regroupements expliquent les écarts de pourcentages entre la Figure 22 et la Figure 23.

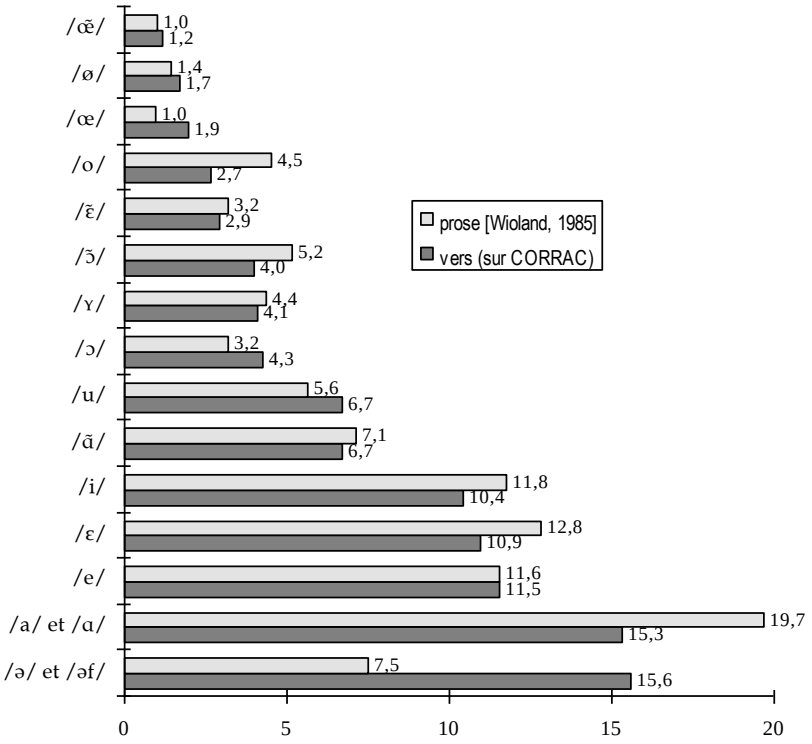


Figure 23. Répartition des voyelles entre prose et vers

De cette comparaison entre prose et vers, il ressort que la hiérarchie entre les voyelles est bouleversée par le rôle des *e* muets. Dans les corpus de prose,

ils ne représentent que 7,5 % des phonèmes vocaliques, tandis que dans les vers, ils représentent 15,6 % des voyelles. La différence ne tient pas seulement aux *e* muets faibles, qui ne représentent que 3,5 % des phonèmes.

Venons-en maintenant à la répartition des voyelles métriques sur les douze positions. Elle est présentée dans le Tableau 23. Les cases grisées signalent les voyelles surepresentées dans des positions données. Le tableau est divisé en deux sous-parties : les voyelles du haut du tableau sont particulièrement absentes des positions 6 et 12, à l'inverse des voyelles du bas du tableau, surepresentées en positions 6 et 12. Ensuite, dans chaque catégorie, les voyelles sont classées en fonction de la somme des écarts à la moyenne. Autrement dit, le haut et le bas du tableau présentent des voyelles au profil très différencié selon les positions métriques et l'on trouve au milieu les voyelles au profil le plus « plat ». Ces profils établis sur 75 000 vers peuvent être considérés comme solides.

Voyelle	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7	P. 8	P. 9	P. 10	P. 11	P. 12	Moy
/ə/	18,2	16,2	13,6	14,4	11,3	**0**	18,3	14,3	12,7	15,4	11,1	**0**	12,1
/əf/	**0**	3,6	5,8	5,5	5,0	**0**	0,0	4,4	6,5	6,5	4,5	**0**	3,5
/a/,/ɑ/	12,0	11,5	11,0	13,1	15,9	8,4	15,9	12,0	10,8	14,3	17,2	10,3	12,7
/ɔ/	2,8	4,5	4,0	3,7	5,7	4,2	2,6	5,3	4,3	3,9	6,8	3,3	4,3
/u/	8,2	8,0	6,9	5,9	7,0	6,6	7,6	7,4	5,9	4,9	6,8	5,2	6,7
/ɔ̃/	4,1	4,1	3,7	4,6	4,6	3,5	4,4	4,7	3,0	4,5	4,3	2,4	4,0
/œ̃/	1,7	1,2	1,0	1,6	1,1	0,2	3,3	1,2	1,0	1,0	0,7	0,1	1,2
/o/	2,1	2,2	2,6	3,1	3,0	2,8	3,5	2,7	2,8	2,7	2,5	1,9	2,7
/wɛ̃/	0,1	0,4	0,6	0,9	0,1	2,2	0,1	0,5	0,7	0,7	0,1	0,5	0,6
/ɥi/	0,9	1,7	1,3	0,8	0,6	1,8	0,6	1,4	1,3	0,5	0,8	2,1	1,1
/ç/	3,0	3,2	3,6	4,5	4,6	4,1	4,5	4,1	4,2	4,0	5,5	4,1	4,1
/ɛ̃/	1,3	2,4	2,5	2,2	1,4	4,7	1,1	2,4	2,9	2,7	1,2	3,6	2,4
/ɑ̃/	4,8	6,1	7,2	6,7	6,4	8,7	6,6	6,1	6,6	6,2	6,0	9,0	6,7
/ø/	0,7	2,2	1,8	1,7	0,7	4,3	0,5	1,9	1,8	1,5	0,4	3,2	1,7
/œ/	0,4	1,9	2,0	1,4	0,8	6,7	0,5	1,8	2,4	1,5	0,8	3,1	1,9
/wa/	2,0	3,1	2,9	1,8	1,4	6,3	0,7	2,3	2,6	1,4	0,9	5,8	2,6
/i/	10,0	8,0	8,2	7,0	9,2	10,7	8,0	7,6	8,6	7,4	11,6	15,0	9,3
/e/	15,8	7,8	9,1	10,8	13,0	10,0	14,2	8,1	9,9	12,7	12,7	14,2	11,5
/ɛ/	11,8	11,9	12,0	10,4	8,0	14,8	7,6	11,8	12,2	8,4	6,2	16,3	10,9

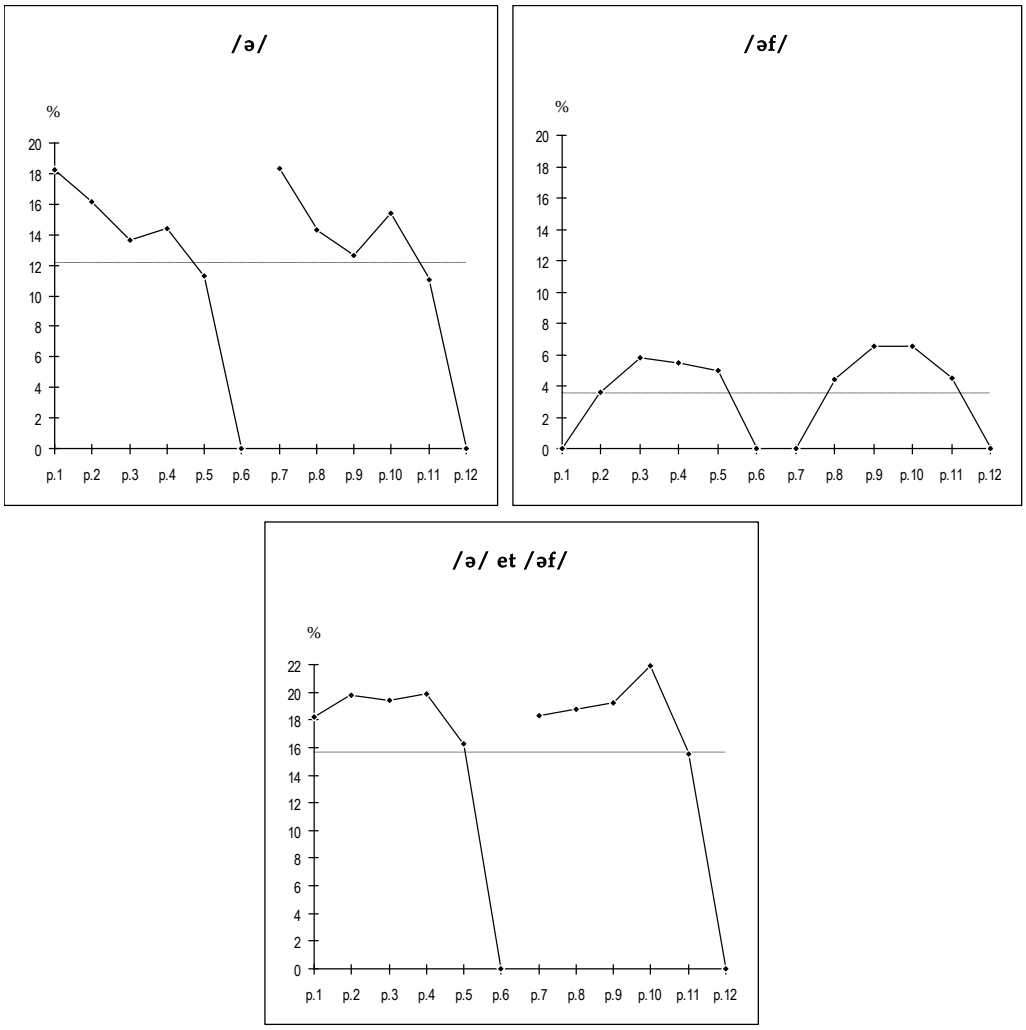
Clef de lecture : P. 1 désigne la position 1 dans le vers ; 18,2 % des vers ont pour première voyelle métrique un /ə/, cette proportion est grisée, car elle

est très significativement supérieure à la répartition moyenne de **/ə/** : 12,1 %[129].

Tableau 23. Répartition des voyelles en fonction des positions métriques

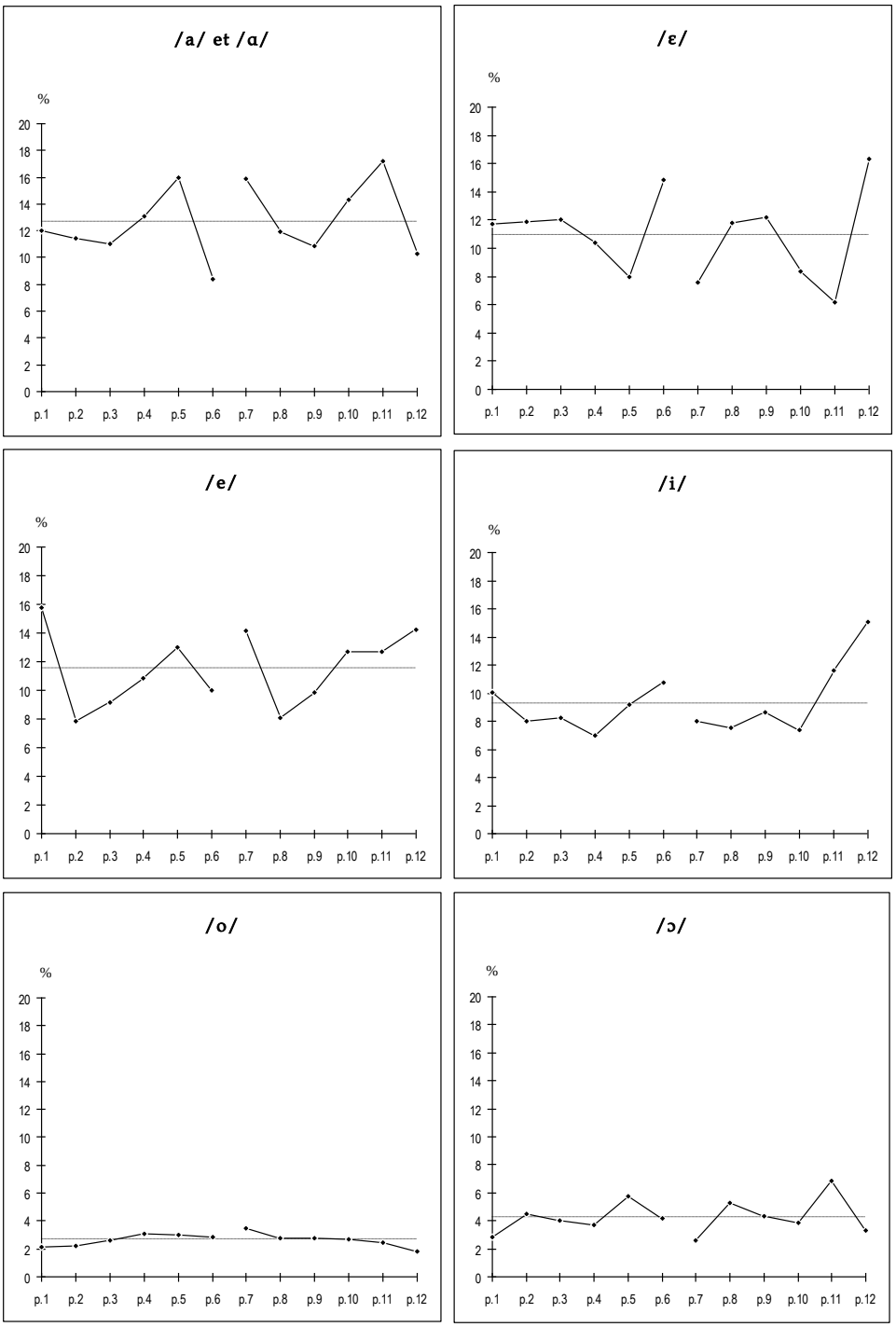
Les graphes des pages suivantes (Figures 24 et 25) montrent comment se répartissent sur chaque position les différentes voyelles métriques. La répartition moyenne du phonème est indiquée par une ligne en pointillé, afin de mieux mettre en évidence les positions typiques d'une voyelle.

[129] Pour effectuer ces tests statistiques, nous avons utilisé une procédure statistique développée par Derquenne [1992]. Celle-ci permet de savoir si une modalité, ici un noyau vocalique, est significativement plus abondant dans la position i que sur l'ensemble des positions. L'annexe V présente les résultats des tests.

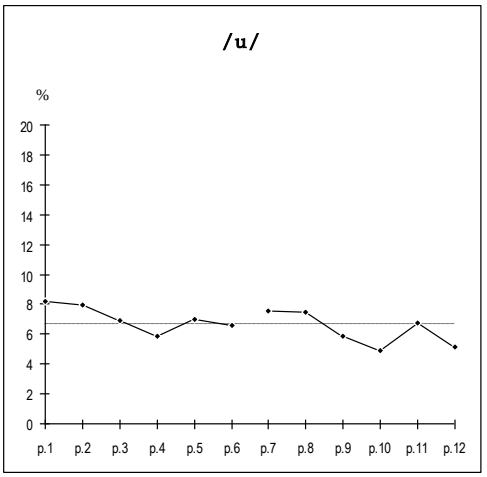


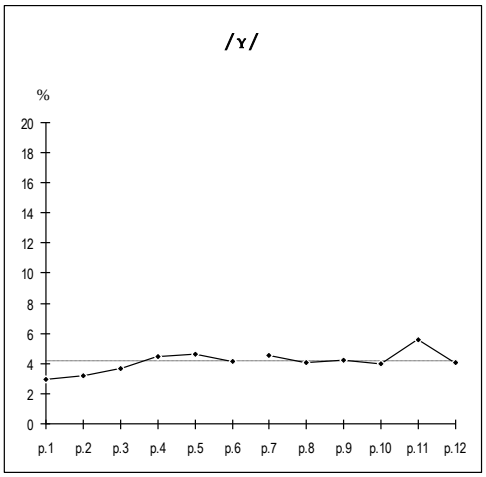
Clef de lecture : 18 % des vers ont un /ə/, faible ou fort, en première position, aucun n'en a en sixième, ni en douzième.

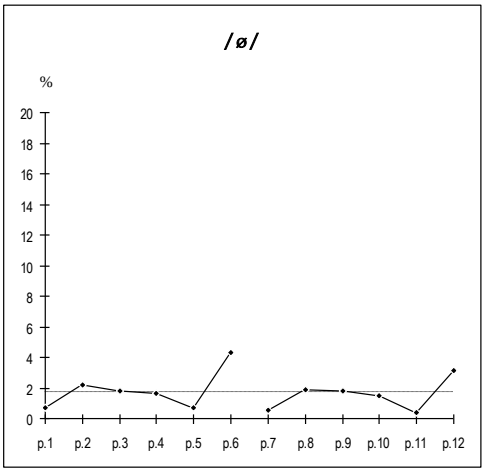
Figure 24. Répartition des* e *muets sur les douze positions

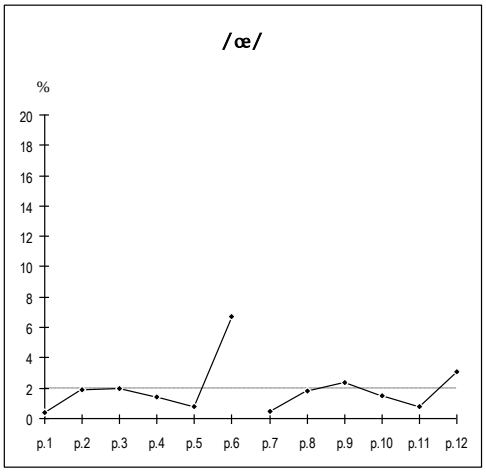


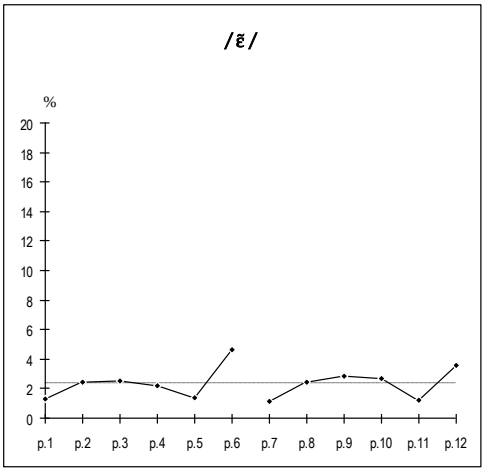
Clef de lecture : Près de 12 % des vers ont un /a/ ou un /ɑ/ en première position, 8 % en sixième et 10 % en douzième.

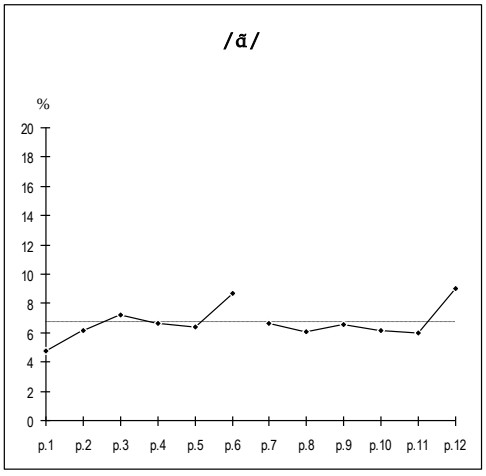

/u/
%
20
18
16
14
12
10
8
6
4
2
0
p.1 p.2 p.3 p.4 p.5 p.6 p.7 p.8 p.9 p.10 p.11 p.12


/ʏ/
%
20
18
16
14
12
10
8
6
4
2
0
p.1 p.2 p.3 p.4 p.5 p.6 p.7 p.8 p.9 p.10 p.11 p.12


/ø/
%
20
18
16
14
12
10
8
6
4
2
0
p.1 p.2 p.3 p.4 p.5 p.6 p.7 p.8 p.9 p.10 p.11 p.12


/œ/
%
20
18
16
14
12
10
8
6
4
2
0
p.1 p.2 p.3 p.4 p.5 p.6 p.7 p.8 p.9 p.10 p.11 p.12


/ɛ̃/
%
20
18
16
14
12
10
8
6
4
2
0
p.1 p.2 p.3 p.4 p.5 p.6 p.7 p.8 p.9 p.10 p.11 p.12


/ɑ̃/
%
20
18
16
14
12
10
8
6
4
2
0
p.1 p.2 p.3 p.4 p.5 p.6 p.7 p.8 p.9 p.10 p.11 p.12

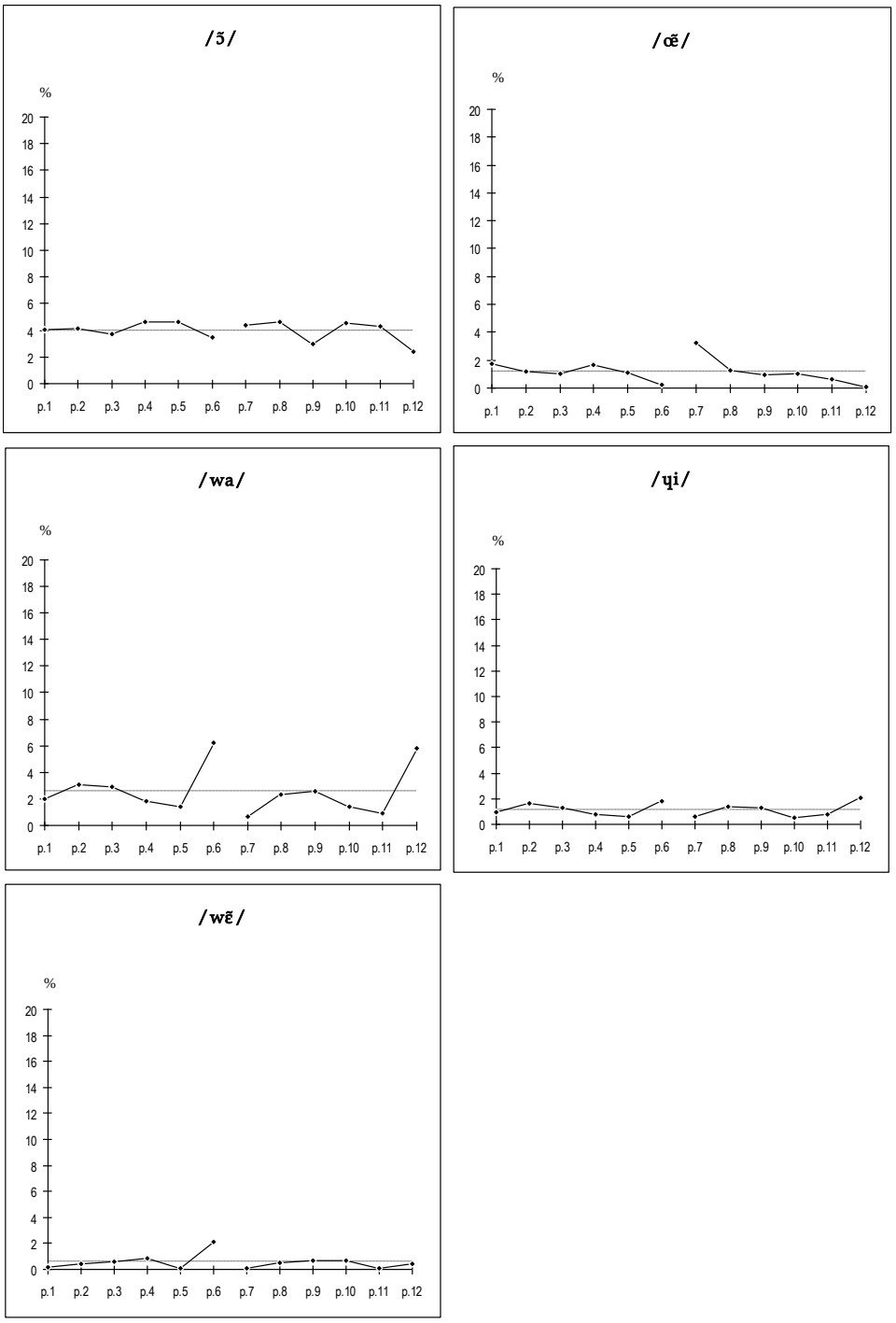


Figure 25. Répartition des voyelles métriques sur les douze positions

VII.1.1.1. Des répartitions contrastées

L'examen rapide de ces dix-neuf graphiques conduit à quelques résultats remarquables. Les voyelles ne sont pas réparties aléatoirement entre les différentes positions. Les écarts par rapport à la moyenne sont significatifs pour la grande majorité des voyelles métriques. Il existe des voyelles plus « métriques » que d'autres, en ce qu'elles ont un profil plus différencié selon les positions. Le *e* muet est évidemment la voyelle la plus métrique, dans la mesure où elle est absente de certaines positions. Qu'il soit fort ou faible, il n'apparaît **jamais** en sixième ni en douzième position. Le *e* muet fort est très typique de la première ou septième position, alors que le *e* muet faible en est complètement absent. Ce dernier, beaucoup moins fréquent puisque souvent élidé, est cependant assez typique des positions centrales.

Dans le vers classique, il n'y a jamais de *e* muet (faible ou fort) en position 6 , ni en position 12. De plus, les *e* muets faibles sont absents des positions 1 et 7.

Les autres voyelles qui connaissent les écarts les plus forts selon les positions, par rapport à leur répartition moyenne, sont /ɛ/, /e/, /a/ et /ɑ̃/, et /i/.

Les voyelles les plus spécifiques des sixième et douzième positions (elles y sont mieux représentées qu'en moyenne) sont le /ɛ/, les voyelles nasales (à l'exception de /ɔ̃/ et /œ̃/), les groupements semi-voyelle-voyelle et enfin /œ/ et /ø/. Autrement dit, les noyaux vocaliques complexes composés graphiquement de plusieurs lettres, digraphes ou trigraphes, sont mieux représentés en fin de segment métrique. Les voyelles /e/ et /i/ sont toutes deux très spécifiques de la douzième position mais pas de la sixième.

Inversement, les voyelles les plus typiques des première et septième positions sont : le *e* muet fort, le /e/, le /œ̃/ et le /u/. Il semble que ces voyelles appartiennent à des mots-outils comme les pronoms personnels, déterminants, conjonctions... : *de, je, que, ne, et, un, pour*... Nous aurons ultérieurement l'occasion de vérifier cette hypothèse en mettant en parallèle les informations phonétiques et syntaxiques. Ce sont surtout sur les premières et dernières positions que les écarts par rapport à la moyenne sont les plus élevés, à l'exception notable du /a/ qui est beaucoup mieux représenté en cinquième et onzième positions que sur les autres positions.

A l'opposé de ces voyelles au profil contrasté, il existe des voyelles pour lesquelles la répartition selon les positions est homogène comme celle du *o* fermé.

Il apparaît de manière très claire que les phonèmes vocaliques ne se répartissent pas au hasard. Ils jouent un rôle dans la construction métrico-rythmique du vers, d'où la nécessité de les faire intervenir comme paramètres de l'analyse rythmique.

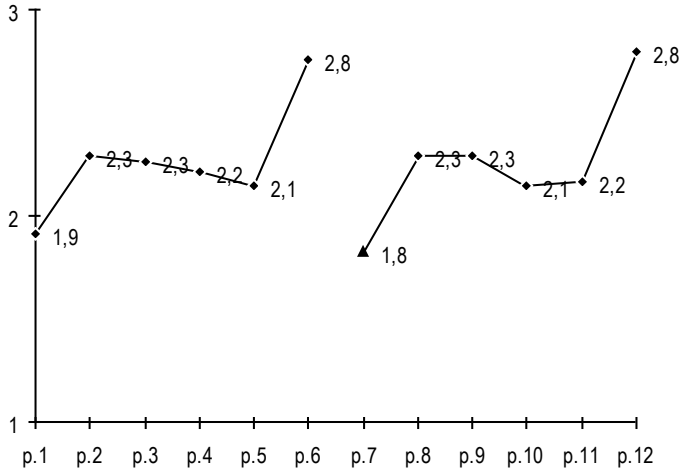
Si les raisons d'une telle répartition nous échappent en grande partie, nous venons cependant de montrer que la plupart des *voyelles* ont une répartition contrastée sur les positions. Qu'en est-il des *syllabes* métriques (noyau vocalique entouré des consonnes qui s'y rattachent) ? La diversité des syllabes rencontrées nous empêche de mener une analyse détaillée, mais il apparaît de manière assez claire que, sur les positions 6 et 12, les syllabes très riches en nombre de phonèmes sont relativement plus fréquentes qu'ailleurs. Il s'agit de syllabes avec attaque et coda, avec une semi-voyelle accompagnant la voyelle... Inversement, sur les premières positions de segments métriques, les syllabes les plus pauvres, en particulier celles qui sont réduites à leur noyau vocalique, sont mieux représentées qu'ailleurs. Dans le vers suivant, alors que la première position est constituée de deux phonèmes, la sixième en compte quatre. Sur le second hémistiche, alors que la première position se réduit à un phonème, la dernière en compte cinq.

Il crut que sans prétendre une plus haute gloire,
(i l)(k r y)(k ə)(s ã)(p r e)(t ã d r)(y)(n əf)(p l y)(ɔ)(t əf)(g l w a r)

Racine, *Mithridate*, vers 51.

La Figure 26 présente le nombre moyen de phonèmes par syllabe métrique selon les positions et montre un écart très net entre la première et la dernière position de segment métrique. Tout se passe comme si l'avancée vers la fin du segment métrique s'accompagnait d'un enrichissement graphico-phonétique des positions. Nous avons également vu que les noyaux vocaliques les plus complexes (digraphes ou trigraphes, diphtongues, nasales) étaient plus spécifiques des positions finales de segment. La dernière position métrique d'un segment métrique se distingue, par sa richesse phonétique, de la première

position du segment suivant, mais aussi de la position qui la précède ; le contraste avec les deux positions adjacentes renforce la saillance des positions 6 et 12. L'épaisseur graphico-phonétique de la position joue donc dans le vers un rôle rythmique.



Clef de lecture : la première position est en moyenne constituée de 1,9 phonèmes, la sixième de 2,8.

Figure 26. Nombre moyen de phonèmes par position métrique

VII.1.1.2. L'hémistiche s'affirme comme une unité autonome

L'allure des courbes de chaque voyelle montre d'une manière frappante le parallélisme quasi-parfait dans la répartition des voyelles sur les deux hémistiches (*cf.* Figure 24 et Figure 25). La même allure de courbe est reproduite d'un hémistiche à l'autre. On peut donc analyser la séquence de vers comme une courbe dont la période serait l'hémistiche. Même s'il y a parfois des écarts notables, les courbes vont toujours dans le même sens. Par exemple, l'écart est important pour le /i/ entre la sixième et la douzième position, mais dans les deux cas la proportion de /i/ y est plus élevée que dans la position précédente.

Nous savons depuis toujours que l'alexandrin classique est constitué de deux hémistiches de six syllabes et que la césure est marquée. Mais sans doute

ne savions-nous pas à quel point le parallélisme des structures vocaliques était grand entre les deux hémistiches —par-delà le nombre identique de syllabes métriques—, au point que le second hémistiche semble être une copie du premier, en tout cas en ce qui concerne sa structure vocalique. La répartition des phonèmes vocaliques montre à son niveau que l'alexandrin est bien un « segment métrique, concaténation de deux segments métriques équivalents » [Roubaud, *1978*, 1988, p. 86]. Chaque segment jouit d'une grande autonomie, dans la mesure où il n'y a pas de débordement d'un segment sur l'autre comme dans le vers à l'italienne.

VII.1.1.3. Quelques différences entre hémistiches

On observe donc des répartitions de voyelles très semblables sur les deux hémistiches. Il existe cependant quelques variations visibles à l'œil et confirmées par des tests statistiques (*cf.* Annexe V. B). Nous avons déjà vu qu'il y avait beaucoup plus de /i/ et de /e/ en position 12 qu'en position 6. Nous rencontrons sans doute ici un effet de la rime, *i* et *e* étant les voyelles des terminaisons de nombreux infinitifs en *er* et *ir* et de nombreux participes passés (*é*, *és*, *ée*, *ées*, *i*, *is*, *ie*, *ies*...) qui offrent un grand nombre de possibilités de rimes. Ils seront donc plus facilement placés en fin de vers.

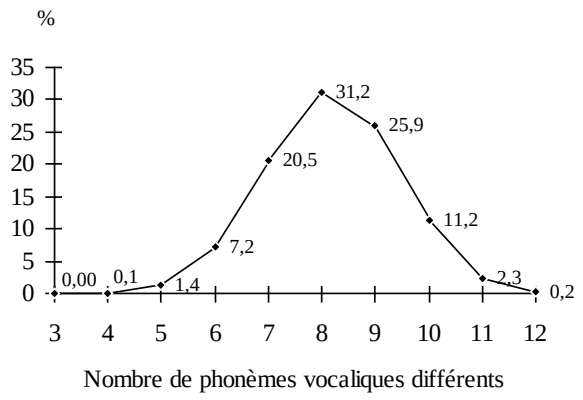
Ces différences entre les deux hémistiches ne sont pas facilement interprétables. Elles sont sans doute liées à des contraintes de rime, à des effets syntaxiques (certains mots-outils, peu nombreux mais très fréquents, placés plutôt sur l'un des deux hémistiches donnent un poids différencié à des voyelles particulières) et à des effets lexicaux. Ainsi, /œ/ est très typique de la position 6, grâce à la syllabe (ɲ œ r), qui correspond à la fin de *Seigneur*. Ce dernier mot est très fréquemment placé sur les positions 5 et 6 du vers. De même, /w ɛ̃/ est typique de la position 6 grâce à *point* qui, en revanche, n'apparaît que rarement à la rime.

Le métromètre nous permet donc de construire globalement la figure vocalique du vers, c'est-à-dire de voir comment se répartissent les voyelles sur les différentes positions. Il rend possible l'exploration d'autres phénomènes vocaliques dans le vers, comme le nombre de voyelles différentes par vers, mais aussi la recherche des cas rares au regard de certains critères.

VII.1.2. Variété vocalique dans les vers

Comment se combinent ces phonèmes vocaliques à l'intérieur du vers ? La courbe de répartition des vers selon le nombre de phonèmes vocaliques distincts est présentée dans la Figure 27. Nous avons considéré les semi-voyelles comme faisant partie de la voyelle.

nombre de phonèmes différents	Effectifs	%
3	5	0
4	81	0,1
5	1023	1,4
6	5401	7,2
7	15362	20,5
8	23365	31,2
9	19415	25,9
10	8383	11,2
11	1744	2,3
12	138	0,2
Total	74917	100



Clef de lecture : 31,2 % des alexandrins de CORRAC ont huit phonèmes vocaliques différents

Figure 27. Répartition des vers selon le nombre de phonèmes différents

Statistiquement, la courbe paraît symétrique, mais d'un point de vue stylistique, il n'y a pas de symétrie dans cette courbe de répartition. La variété vocalique (12 voyelles différentes dans un alexandrin) en tant que telle n'a rien de remarquable.

À chercher dans vos yeux une mort qui me fuit.

Racine, *Andromaque*, vers 496.

Au contraire, la combinaison dans l'alexandrin d'un faible nombre de voyelles constitue des cas d'assonance appuyée, stylistiquement remarquables, polyptotes et/ou parallélismes syntaxiques, comme le montrent les exemples suivants :

Il y va de ma charge, il y va de ma vie : (i, a, ə)

Corneille, *Polyeucte*, vers 1018.

Que ce que je l'ai fait il est digne de l'être, (ə, ɛ, i)

Corneille, *Don Sanche d'Aragon*, vers 926.

Croyez-moi généreuse, et soyez généreux, (wa, e, ø)

Corneille, *Tite et Bérénice*, vers 1229.

Toujours aimer, toujours souffrir, toujours mourir. (u, e, i)

Corneille, *Suréna*, vers 268.

Me négliger toujours ! Et pour qui vous néglige ! (ə, e, i, u)

Corneille, *Clitandre*, vers 921.

La majesté d'Aegée, et le sceptre d'Athènes, (a, ɛ, e, ə)

Corneille, *Médée*, vers 114.

Ou de ce grand amour, ou de ce grand courage ; (u, ə, ɑ̃, a)

Corneille, *L'Illusion comique*, vers 1550.

J'ai fait ce que j'ai dû, je fais ce que je dois. (e, ə, y, wa)

Corneille, *Le Cid*, vers 900.

Ou laissez-moi périr, ou laissez-moi régner. (u, e, wa, i)

Corneille, *Cinna*, vers 1192.

Si vous l'avez prisé, si vous l'avez chéri... (i, u, a, e)

Corneille, *Polyeucte*, vers 963.

Je sais ce que je suis ; je sais ce que vous êtes. (ə, ɛ, ɥi, u)

Corneille, *Pompée*, vers 1286.

Il n'est ni roi, ni père, il n'est prière, empire, (i, ɛ, wa, ɑ̃)

Corneille, *Suréna*, vers 561.

Je ne sais où je vais, je ne sais où je suis. (ə, ɛ, u, ɥi)

Racine, *Phèdre*, vers 1004.

Ces assonances coïncident quasiment toujours avec des structures construites sur des parallélismes : parallélisme entre les deux hémistiches quasi identiques au mot près. Ces structures plus fréquentes chez Corneille que chez Racine sont révélatrices d'un type de rhétorique caractéristique : envisager toujours les deux branches de l'alternative... Dans ces hémistiches construits sur des modèles lexicaux et syntaxiques quasi-identiques, la variation est faible, mais elle porte au moins sur le dernier mot du segment métrique, le mot le plus marqué, porteur de la rime à la fin du second segment métrique. Remarquons que l'on ne pourrait pas imaginer des hémistiches aux structures lexico-

syntaxiques quasi-identiques où la variation ne porterait pas sur la dernière position. Il y aurait alors une rime intérieure et un risque de confusion dans la reconnaissance des vers, ce qui est contraire à l'esthétique classique, soucieuse de clarté formelle.

Les vers qui mobilisent un faible nombre de voyelles correspondent à des structures syntaxiques fondées sur le parallélisme.

Jacques Jouet et Jacques Roubaud [1993] ont été intrigués par l'allure de ce vers de Baudelaire qui contient quatre *e* muets faibles comptés :

> Laissent piteusement leurs grandes ailes blanches
>
> Baudelaire, *Les Fleurs du Mal*, « L'albatros ».

Ils ont exploré le champ des alexandrins classiques, à savoir les 64 espèces d'alexandrins qui ne contiennent pas de *e* muets sur les positions 1, 6, 7 et 12. Parmi ces espèces, les plus remarquables sont les alexandrins jouetiens *maximaux* qui contiennent quatre *e* muets, comme le vers de « *L'albatros* ». Un alexandrin classique ne peut en contenir davantage, puisqu'il ne peut y avoir d'*e* muet aux positions 1, 6, 7 ni 12, et que deux *e* muets ne peuvent se succéder. Les exemples de vers sont donnés dans *[ə]*, paru dans la *Bibliothèque Oulipienne.*

Nous avons recherché sur CORRAC ces alexandrins saturés d'*e* muets. Voici les vers avec quatre *e* muets faibles que trouve le métromètre :

Cette belle, de vrai, quoique toute de glace,
ɛ - əf - ɛ - əf - ə - ɛ - wa - əf - u - əf - ə - a -

Corneille, *La Galerie du palais*, vers 429.

Elle vous traite mal, mais elle n'aime rien,
ɛ - əf - u - ɛ - əf - a - ɛ - ɛ - əf - ɛ - əf - ɛ̃

Corneille, *La Galerie du palais*, vers 868.

Passe une bonne part d'une si belle nuit, °[130]
a - y - əf - ɔ - əf - a - y - əf - i - ɛ - əf - ɥi

Corneille, *Le Cid*, vers 1268.

D'une seule maison brave toutes les nôtres,
y - əf - œ - əf - ɛ - ɔ̃ - a - əf - u - əf - *e* - o -

Corneille, *Horace*, vers 352.

Rendent grâces au ciel d'une commune voix, °
ɑ̃ - əf - ɑ - əf - o - ɛ - u - əf - ɔ - y - əf - wa

Corneille, *Œdipe*, vers 1960.

Qu'elle ne pousse point cette même tendresse,
ɛ - əf - ə - u - əf - wɛ̃ - ɛ - əf - ɛ - əf - ɑ̃ - ɛ

Racine, *Mithridate*, vers 632.

Une profonde nuit enveloppe sa race, °
y - əf - ɔ - ɔ̃ - əf - ɥi - ɑ̃ - əf - ɔ - əf - ɑ

Racine, *Athalie*, vers 1008.

Le métromètre trouve cinq vers avec cinq *e* muets faibles chez Corneille et deux chez Racine, alors que Roubaud en a repéré respectivement quatorze et trois. La différence tient à notre définition du *e* muet : il est fort /ə/, et non faible /əf/, dans le formalisme du métromètre, lorsqu'il suit deux consonnes (consonne + liquide), contrairement à l'option prise dans *[ə]*. Par exemple, dans le vers suivant les *e* de *contre*, *votre* sont forts :

Et contre votre frère, et même contre vous,
(e)(k ɔ̃)(t r ə)(v ɔ)(t r ə)(f r ɛ r)(e)(m ɛ)(m əf)(k ɔ̃)(t r ə)(v u)

Racine, *Mithridate*, vers 121.

Ce choix a été fait parce que ces *e* muets se maintiennent mieux en langue que les autres, leur chute entraînant une suite de trois consonnes.

---

[130] Sur 80 000 vers, seuls les trois alexandrins suivis d'un petit rond ont échappé à l'œil de J. Roubaud.

Résumons les résultats de cette section. Chaque phonème participe à la construction du rythme par les variations dans sa distribution selon les positions. Le *e* muet joue un rôle *métrique* par son absence systématique en positions 6 et 12, renforcé, pour le *e* muet faible, par son absence en positions 1 et 7. Si chaque voyelle a une allure de courbe qui lui est propre, il existe pour chacune un fort parallélisme entre les courbes des deux hémistiches. Chaque voyelle délimite l'hémistiche comme unité périodique. A leur manière, sans doute peu perceptible, les phonèmes vocaliques renforcent la structure métrique par le parallélisme des figures des deux hémistiches, mais ils contribuent aussi à l'élaboration du rythme par les variations de leur présence selon les positions.

Un autre élément phonétique contribue au rythme, c'est le nombre de phonèmes par position. L'épaisseur phonétique des positions 6 et 12 est nettement plus importante que celle des positions adjacentes, ce qui entraîne un contraste important entre la fin d'un segment métrique et le début du suivant et ce qui renforce la dernière position.

VII.2. FRONTIERES DE MOTS ET POSITIONS METRIQUES

Après les aspects phonétiques, venons-en aux aspects lexicaux. Nous avons mis en place un marquage de fin de mot qui attribue la valeur « 1 » à la dernière position métrique de chaque mot, et « 0 » aux autres positions.

Chez Corneille et Racine, la distribution des fins de mots sur les douze positions (*cf.* Figure 28) montre que la fin de vers et la fin d'hémistiche[131] coïncident toujours avec une fin de mot. Cela ne constitue pas un résultat nouveau, dans la mesure où l'autonomie des deux hémistiches implique qu'il n'y ait pas de mot enjambant les deux segments.

[131] Pour la fin d'hémistiche, nous atteignons 99,7 % de marquage de fin de mot, les 0,3 % restant étant dus à des erreurs (*cf.* chapitre VII).

Dans l'alexandrin « classique », les sixième et douzième positions coïncident toujours avec la dernière syllabe comptée d'un mot. Autrement dit, aucun mot ne déborde sur le segment métrique suivant.

Fins de mot	**P. 1**	**P. 2**	**P. 3**	**P. 4**	**P. 5**	**P. 6**	**P. 7**	**P. 8**	**P. 9**	**P. 10**	**P. 11**	**P. 12**
	73,8	64,0	67,7	72,8	38,5	99,7	73,9	60,3	67,3	68,4	32,1	100

Les courbes de répartition des frontières de mots sont pratiquement identiques sur les deux hémistiches. Une fois encore est vérifiée l'hypothèse selon laquelle l'alexandrin est la concaténation de deux segments métriques.

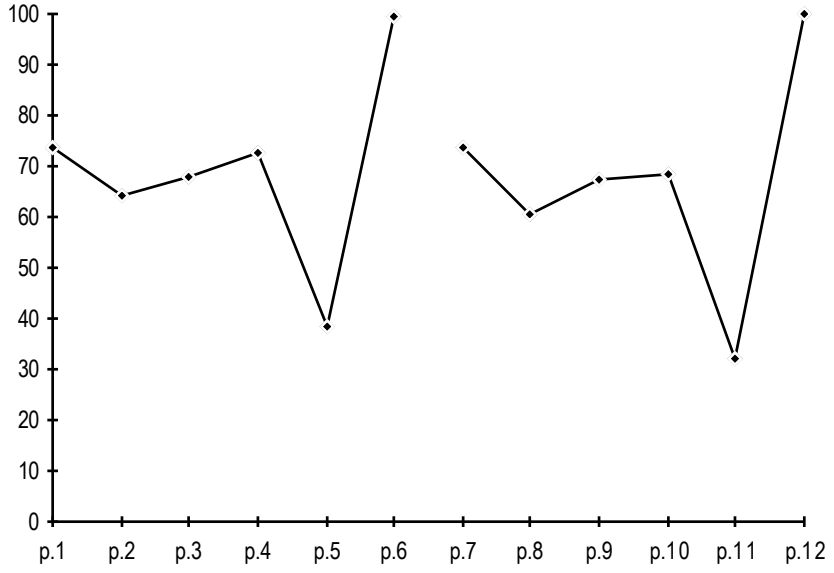


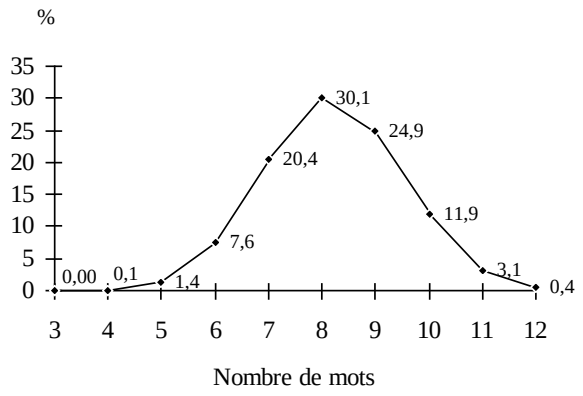
Figure 28. Distribution des frontières de mots sur les douze positions

Il faut de plus noter que la proportion de mots monosyllabiques en dernière position sur les deux hémistiches est relativement faible, comme en témoignent les « creux » de la courbe en cinquième et onzième positions. Alors que l'on a 74 % de monosyllabes en première et septième positions, on n'en a que 39 % en sixième position et 32 % en douzième position. Le marquage de la position

finale du segment est renforcé par l'absence de marquage sur la position précédente.

Ce marquage permet aussi de construire la répartition des alexandrins en fonction du nombre de mots qu'ils contiennent[132] (*cf.* Figure 29).

Nombre de mots	Effectifs	%
3	3	0,0
4	103	0,1
5	1073	1,4
6	5879	7,6
7	15716	20,4
8	23220	30,1
9	19249	24,9
10	9214	11,9
11	2411	3,1
12	318	0,4
Total	77186	100



Clef de lecture : sur CORRAC, 103 alexandrins sont composés de 4 mots

Figure 29. Répartition des alexandrins selon le nombre de mots

Prenons les vers les plus atypiques, ceux qui sont composés d'un très faible nombre de mots ou au contraire d'un nombre très élevé. Comme pour la répartition en fonction du nombre de phonèmes vocaliques différents, les vers comptant peu de mots sont plus remarquables que les vers avec beaucoup de mots. Le nombre minimal de mots dans un vers de Corneille et Racine est de trois et on en trouve trois exemples. A chaque fois, nécessairement, un mot occupe tout un hémistiche.

[132] Pour la construction de ce tableau, le marquage de fin de mot n'a en fait pas été utilisé. On est parti du corpus des vers entièrement reconstitués, y compris ceux de Corneille, et l'on a compté le nombre de mots par vers. Le tiret et l'apostrophe ne sont pas considérés comme des séparateurs de mots.

PAULINE
Imaginations !
POLYEUCTE
Célestes vérités !

Corneille, *Polyeucte*, vers 1285.

Laissez-en ralentir l'impétuosité

Corneille, *Othon*, vers 1564.

Compendieusement énoncer, expliquer,

Racine, *Les Plaideurs*, vers 794.

Il existe par ailleurs d'autres cas de mots-hémistiches, parmi lesquels revient souvent le terme *impétuosité*.

Je ne me pique point d'insensibilité

Corneille, *La Suivante*, vers 420.

L'impétuosité d'un premier mouvement

Corneille, *Rodogune*, vers 333.

Ne te figure point d'impossibilité

Corneille, *Théodore*, vers 1565.

Ma haine n'aura plus d'impétuosité

Corneille, *Héraclius*, vers 894.

L'impétuosité d'un si prompt changement

Corneille, *Andromède*, vers 1161.

Il en faut réprimer l'impétuosité

Corneille, *Othon*, vers 1572.

L'irrésolution doit-elle être éternelle,

Corneille, *Tite et Bérénice*, vers 1427.

L'impétuosité s'en devrait retenir

Corneille, *Psyché*, vers 1878.

L'impétuosité de cette grandeur d'âme,

Corneille, *Suréna*, vers1466.

En revanche, les vers avec douze mots (un par syllabe) sont beaucoup plus nombreux. A titre d'exemple :

C'est en vain que l'on fuit, tôt ou tard on s'y brûle,

Corneille, *Mélite*, vers 106.

Cependant, le sentiment de durée varie. A nombre de mots égal, les rythmes diffèrent en fonction de la structure des syllabes et du nombre de *e* muets élidés. Dans les deux premiers vers, les syllabes métriques sont toutes de type CV[133].

Je sais ce que je suis et ce que je vous dois

Corneille, *La Suivante*, vers 988.

J'ai fait ce que j'ai dû, je fais ce que je dois

Corneille, *Le Cid*, vers 900.

Dans les vers suivants, les syllabes CVC (*cœur*, *jour*, *pur*...) augmentent le sentiment de durée du vers :

Tu n'y vois que mon cœur, qui n'a plus un seul trait

Corneille, *Mélite*, vers 307.

Le jour n'est pas plus pur que le fond de mon cœur

Racine, *Phèdre*, vers 1112.

Et dans ces derniers vers, chaque syllabe métrique est allongée par la présence d'un *e* muet élidé.

On me joue, on me brave, on me tue, on s'en rit

Corneille, *La Place royale*, vers 639.

Et la terre, et le fleuve, et leur flotte, et le port

Corneille, *Le Cid*, vers 1299.

La fin d'un segment métrique coïncide toujours avec la fin d'un mot. Ce mot, dans plus des deux tiers des cas, est composé d'au moins deux syllabes. Au contraire, dans presque les trois-quarts des cas, le mot à l'initiale d'un segment métrique est un monosyllabe. Le contraste entre la longueur moyenne du dernier mot d'un segment et celle du premier mot du segment suivant est donc important et souligne cette position de frontière.

VII.3. REPARTITION DES CATEGORIES SYNTAXIQUES

Nous en venons maintenant aux marquages syntaxiques. Comme nous l'avons vu dans le chapitre V qui leur était consacré, le seul que nous ayons

[133] C : phonème consonantique, V : phonème vocalique.

finalement retenu est un marquage morpho-syntaxique. Nous cherchons à voir comment se répartissent, chez Corneille et Racine, les différentes catégories sur chacune des positions métriques.

L'analyseur syntaxique Sylex développé par P. Constant, qui est intégré au métromètre, distingue neuf parties du discours : les noms, les verbes, les adjectifs, les adverbes, les patronymes (ou noms propres), les déterminants et pronoms, les prépositions, les conjonctions et les relatifs. Rappelons qu'il y a des erreurs dans la catégorisation syntaxique. Par exemple les pronoms disjoints, qui théoriquement devraient être rattachés à des noms, sont parfois traités comme des pronoms. De même, dans la catégorie « relatif », on trouve des mots qui pourraient être analysés comme noms (*quoi* par exemple). Enfin, tous les participes passés sont interprétés comme étant des adjectifs.

Nous présenterons la répartition des catégories sur les douze positions, puis nous examinerons en détail la question des mots-outils en fin de segment métrique.

VII.3.1. Catégories syntaxiques et positions métriques

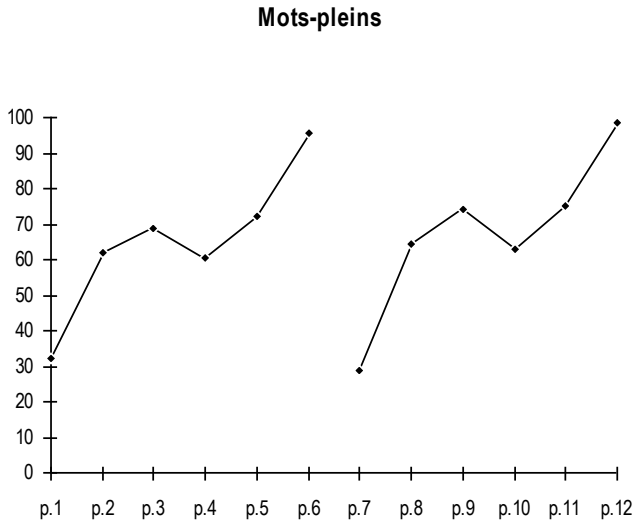
Les neuf catégories syntaxiques identifiées par Sylex peuvent être regroupées en deux grandes catégories : les mots pleins (ou mots lexicaux) et les mots vides (ou mots-outils, mots grammaticaux, mots charnière). Le Tableau 24 présente le pourcentage de positions métriques occupées par chacune des catégories.

Nom commun	23,3
Verbe	24,0
Adjectif	10,2
Adverbe	6,1
Patronyme	2,6
Total mots pleins	**66,3**
Déterminant, pronom	19,4
Préposition	8,0
Conjonction	2,8
Relatifs	3,5
Total mots vides	**33,7**

Clef de lecture : 8 % des positions sont occupées par des syllabes métriques appartenant à des prépositions.

Tableau 24. Répartition globale des catégories syntaxiques

Les deux tiers des positions métriques sont occupées par des mots pleins, un tiers par des mots-outils. La répartition sur les douze positions métriques des mots pleins apparaît sur la Figure 30.



Clef de lecture : 32 % des vers ont un mot plein en première position (68 % un mot-outil), et 99 % ont un mot plein en douzième position.

Figure 30. Mots pleins dans le vers

La douzième position opère un découpage remarquable au sein du lexique : elle n'est pratiquement occupée que par des mots pleins (noms, verbes, adjectifs, adverbes, patronymes). Elle permet donc une classification du lexique, si on ne dispose pas de l'information *a priori*.

Le même phénomène est vrai pour la sixième position, avec cependant un taux légèrement plus élevé de mots-outils. Les mots pleins sont donc typiques des fins de segments métriques ou d'hémistiches, tandis que les mots-outils se répartissent plutôt en début de segment, même si leur prépondérance n'exclut pas la présence de mots pleins. Ainsi la première position, dans près d'un tiers des vers, est occupée par un mot-plein. En tendance globale, la part des mots pleins augmente de la première à la sixième position sur chaque hémistiche.

Peut-on poser comme hypothèse que les mots-outils sont interdits en fin de segment métrique ? L'examen des mots-outils en positions 6 et 12 nous apportera quelques éléments de réponse.

Examinons à présent plus précisément la répartition des différentes catégories syntaxiques. Ces répartitions sont fragilisées par les erreurs d'analyse syntaxique (en particulier les pronoms disjoints - par exemple *lui* dans *pour lui*- qui auraient dû être analysés comme noms) et par le flou entre les catégories (les participes passés sont classés comme adjectifs, il n'y a pas de distinction entre déterminants et pronoms...).

Catégorie syntaxique	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7	P. 8	P. 9	P. 10	P. 11	P. 12	Moy
Nom commun	6,4	16,6	18,3	13,4	26,9	46,8	3,3	18,2	23,9	18,1	33,1	54,0	23,3
Adjectif	1,8	4,8	7,4	10,1	10,2	10,9	2,7	8,4	12,7	16,9	18,4	18,0	10,2
Adverbe	4,1	5,4	6,4	7,6	8,4	12,1	3,7	6,5	6,5	5,8	3,1	3,9	6,1
Patronyme	3,5	3,2	2,4	2,3	4,2	4,6	1,2	1,6	1,3	1,4	2,6	2,9	2,6
Verbe	16,5	32,0	34,0	27,1	22,2	21,2	17,8	29,7	29,8	20,6	17,7	19,9	24,0
Déterminant, pronom	32,0	25,7	17,7	21,9	18,2	3,8	31,5	27,4	14,6	21,4	17,4	1,2	19,4
Préposition	9,9	6,4	8,2	12,5	8,2	0,3	19,9	5,5	7,1	11,5	6,3	0,1	8,0
Conjonction	15,9	0,4	1,3	1,4	0,5	0,1	10,6	0,2	1,4	1,7	0,3	0,0	2,8
Relatif	10,0	5,4	4,1	3,6	1,1	0,1	9,3	2,5	2,7	2,5	1,1	0,0	3,5

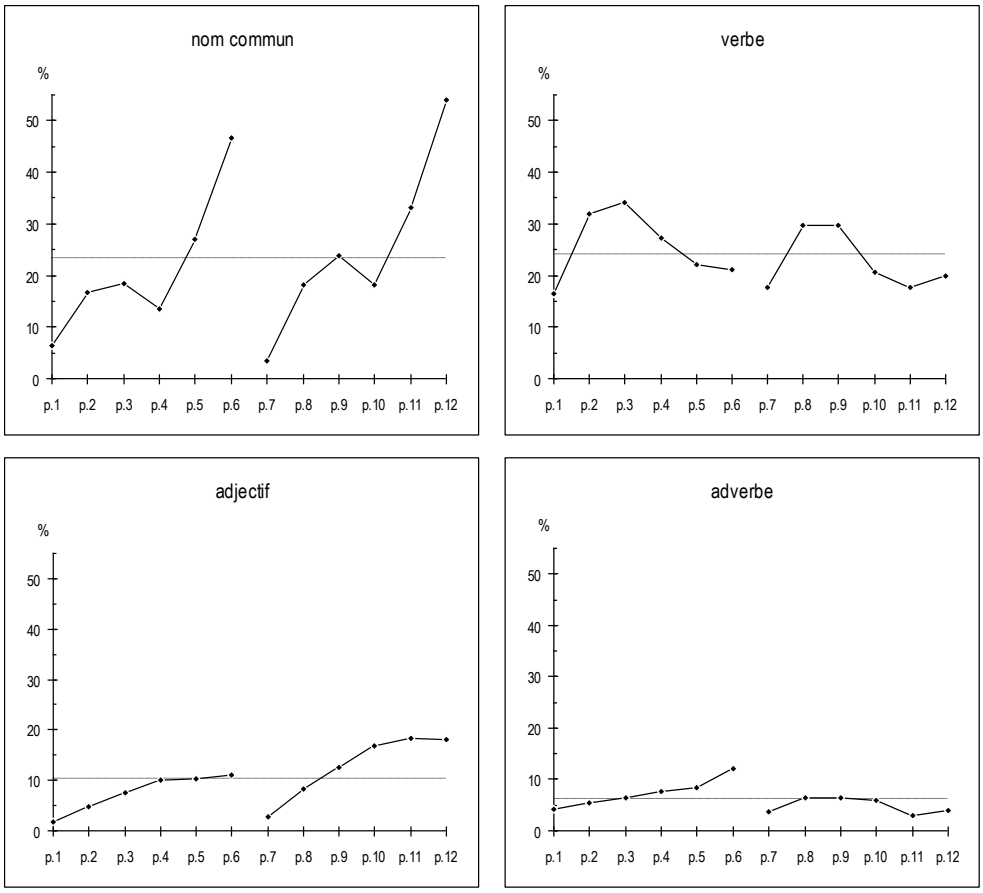
Clef de lecture : 46,8 % des vers ont une sixième position occupée par une syllabe de nom, cette proportion est grisée car elle est très significativement supérieure à la répartition moyenne des noms : 23,3 %. L'annexe V présente les résultats des tests statistiques.

Tableau 25. Répartition des catégories syntaxiques sur les douze positions

Outre le fait que les sixième et douzième positions sont quasi-exclusivement occupées par des mots pleins et que sur les positions 1 et 7 les mots-outils dominent largement, les détails par catégorie permettent de voir que le nom l'emporte très largement en positions 6 et 12 et que parmi les mots pleins, seuls les verbes n'occupent pas spécifiquement les positions finales du segment métrique : ils apparaissent surtout sur les positions centrales.

Le segment métrique coïncide en général avec une unité syntaxique. Nous avons le sentiment, mais cela mériterait d'être approfondi, que dans chaque segment les structures de type Sujet-Verbe-Objet sont relativement les mieux représentées.

Le Tableau 25 présente, pour chaque catégorie syntaxique, son taux d'occurrence sur les douze positions et la Figure 31 les représentations graphiques correspondant à chaque catégorie syntaxique.



Clef de lecture : Près d'une fois sur deux, la sixième position métrique est occupée par une syllabe appartenant à un nom.

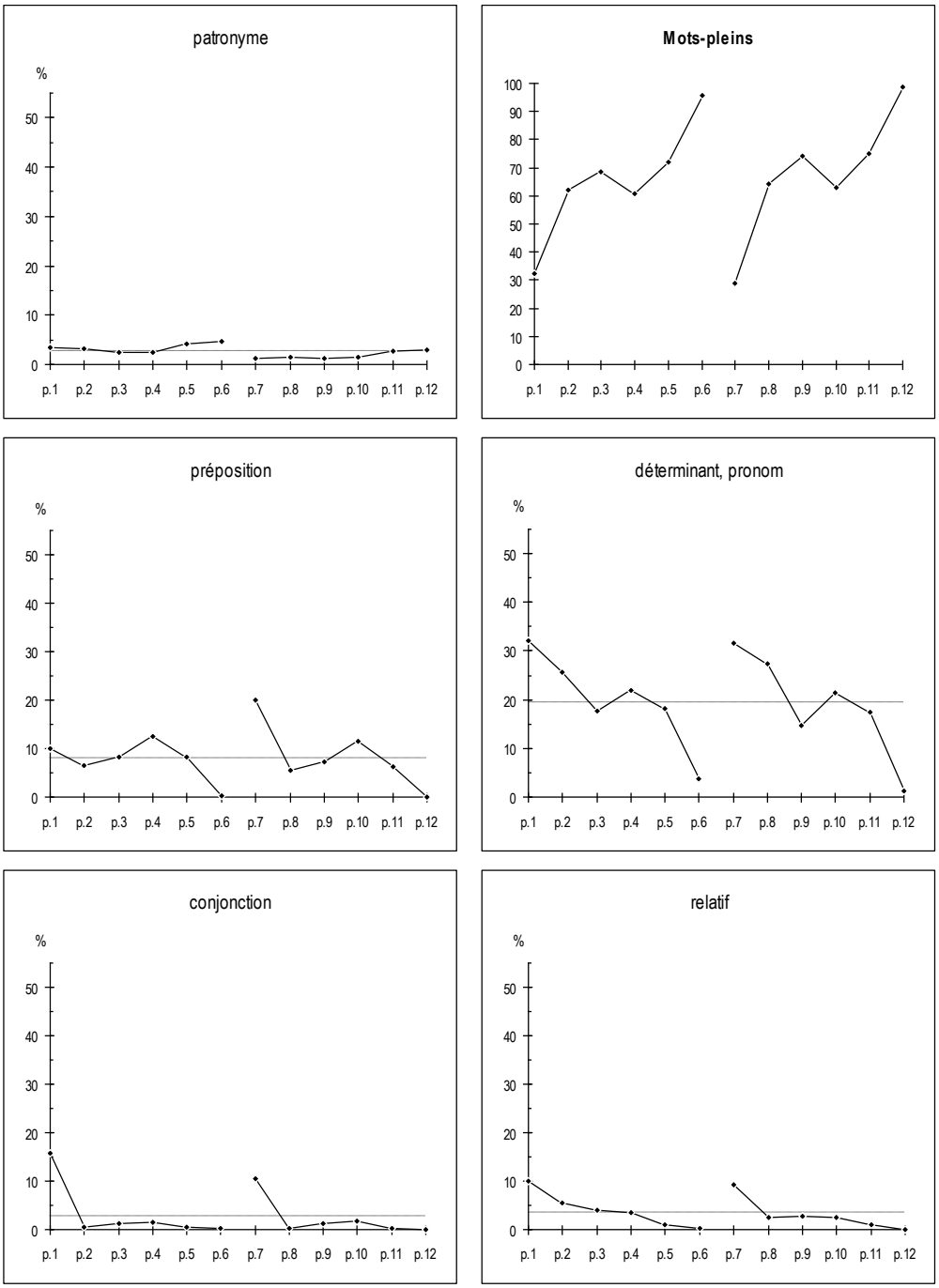


Figure 31. Répartition des catégories syntaxiques sur les douze positions

A l'exception des noms propres, chaque catégorie syntaxique a un profil différencié selon les positions, les mots pleins (à l'exception des verbes) étant mieux représentés sur les positions finales, les mots-outils l'étant davantage sur les positions initiales.

Tout comme pour les phonèmes vocaliques, nous observons pour chaque catégorie syntaxique une distribution similaire sur les positions métriques du premier et du second hémistiche. Ce parallélisme vient renforcer le fait que le vers est la concaténation de deux segments et que la période est à six positions. Il y a donc bien répétition d'une forme, et cela est d'autant plus remarquable que cette répétition se retrouve, indépendamment, sur toutes les catégories syntaxiques. Si l'on considère que chaque catégorie syntaxique est un instrument, chaque instrument, séparément, répète son motif d'un hémistiche sur l'autre.

Il existe cependant quelques différences significatives dans la répartition des catégories syntaxiques entre le premier et le second hémistiche (*cf.* Annexe V. B). Globalement, les noms et les adjectifs sont relativement plus fréquents sur le second hémistiche que sur le premier. Inversement, verbes, adverbes et noms propres sont plus fréquents sur le premier ; les verbes sur les positions 2 à 5, les adverbes sur les positions 4 à 6 et les noms propres sur toutes les positions.

Pour les mots-outils, la différence la plus marquante tient à la répartition des prépositions, très présentes en position 7 (relativement à la position 1) et des conjonctions de coordination, très spécifiques de la position 1.

La répartition des voyelles recoupe en partie celle des catégories syntaxiques, comme le prouve le tableau croisant les deux variables (*cf.* Annexe V. A. 4). Ces effets viennent principalement des mots-outils qui, peu nombreux mais très fréquents, impriment des sonorités particulières à leur catégorie. Ainsi la prédominance des conjonctions de coordination en première et septième positions tient essentiellement à la conjonction *et* (13 % de /e/[134] en position 1, 10 % en position 7, contre 3 % en moyenne) et dans une moindre

[134] Les trois-quarts des syllabes /e/ correspondent à la préposition *et.*

mesure à *mais* pour la première position (4 % de /m e/, au lieu de 0,7 %). Pour les prépositions, leur prédominance en position 7 tient à *à* (8 % de /a/ au lieu de 2 %) et à *de* (5 % de /də/ au lieu de 2 %). En position 1, les pronoms *je* et *il* sont particulièrement bien représentés, en position 7, le pronom *il* et le déterminant *un*. Se dessine discrètement une structure syntaxique dominante du type sujet (pronom) - verbe - complément (nom) qui s'inscrirait dans la structure de l'hémistiche.

VII.3.2. Mots-outils en fin d'hémistiches : quelques difficultés d'analyse syntaxique

Les mots-outils en positions 6 et 12 sont très rares. Nous examinerons dans cette sous-section les cas de mots grammaticaux situés en fin d'hémistiche.

Dans le vers classique, on s'attend en général à ne pas trouver de mots grammaticaux en fin de vers, ni en fin d'hémistiche. En effet, à l'époque classique, l'enjambement de vers à vers et d'hémistiche à hémistiche, est proscrit par Malherbe, et la règle est figée dans l'*Art poétique* de Boileau.

> Enfin Malherbe vint, et, le premier en France,
> Fit sentir dans les vers une juste cadence,
> D'un mot mis en sa place enseigna le pouvoir,
> Et réduisit la muse aux lois du devoir. (...)
> Les stances avec grâce apprirent à tomber,
> Et le vers sur le vers n'osa plus enjamber.
>
> Boileau, *Art Poétique*, Chant premier.

> Que toujours dans vos vers le sens coupant les mots
> Suspende l'hémistiche, en marque le repos.
>
> Boileau, *Art Poétique*, Chant premier.

Un mot phonologique ne peut donc se trouver à la fois sur deux segments métriques. Autrement dit, il faut que dans le vers il y ait coïncidence entre segment métrique et segment syntaxique, plus précisément que les fins de segments métriques coïncident avec des fins de syntagmes. En particulier, les mots-outils ou mots clitiques[135] qui ont pour fonction d'accompagner un noyau

[135] Cornulier [1982b, p. 138-139] emploie le terme de *clitique* pour désigner les mots-outils qui « n'apparaissent qu'au voisinage d'autres mots auxquels ils semblent

de syntagme ne peuvent être séparés de leur appui par la structure métrique. Les mots clitiques n'ont pas un fonctionnement autonome. Ainsi, un pronom, une préposition, une conjonction ou un relatif, ne peuvent être placés à la fin d'un vers ou d'un hémistiche, la suite étant reportée sur le vers ou l'hémistiche suivant. Si la règle était strictement respectée, on ne devrait trouver en douzième ni en sixième position aucun mot grammatical dont la base serait placée sur le segment métrique suivant.

Pour la sixième position, les règles devraient s'imposer aussi bien que sur la douzième, puisque, comme nous l'avons vu précédemment, il y a un très grand parallélisme entre les deux segments métriques : ce qui s'applique au deuxième doit sans doute s'appliquer aussi au premier.

Mais, comme l'ont montré de nombreuses études qui se sont intéressées aux perturbations de la zone frontière entre les hémistiches, la sixième position ne fonctionne pas tout à fait comme la dernière. En particulier, le non respect des règles de concordance entre métrique et syntaxe apparaîtra bien avant et bien plus fréquemment à l'hémistiche qu'en fin de vers. On se reportera à *La Vieillesse d'Alexandre*, qui montre comment l'édifice de l'alexandrin a peu à peu été altéré (les principaux artisans de cette altération étant Hugo, Baudelaire et Nerval d'une part, Mallarmé, Lautréamont et Rimbaud de l'autre), et comment cette destruction a d'abord porté sur la fin du premier segment avant de toucher, bien plus tard, la fin du vers, dans les formes avancées du vers libre. Cornulier [1982b] et Gouvard [1995] ont refait le même type de travail sur des corpus légèrement différents et avec des systèmes de description de ces phénomènes différents.

Il ne devrait donc pas se trouver beaucoup de mots-outils en positions 6 et 12. Pourtant, le métromètre nous en indique un nombre non négligeable. Que signifie la présence d'un mot-outil en position 6 ou 12 dans le vers classique ?

---

s'appuyer ». Quand les clitiques sont antéposés à leur base, ils sont dits *proclitiques*, quand ils sont postposés, *enclitiques*.

Cela peut provenir :

- D'une mauvaise analyse syntaxique qui attribue une mauvaise catégorie syntaxique. Cela peut arriver pour des homonymes, ainsi *vers* de poésie est confondu avec la préposition *vers*, le participe passé *tu* avec le pronom personnel..., mais aussi pour des termes qui selon le contexte syntaxique peuvent avoir des étiquettes différentes : des mots qui peuvent être adverbes ou prépositions... ;
- De certains pronoms qui, dans des contextes déterminés, ont les qualités des substantifs, ce que n'a pas toujours vu l'analyseur syntaxique. Ainsi *lui* après préposition a-t-il une valeur de nom et non plus de pronom. On a bien des mots grammaticaux en fin de segment métrique, mais ils sont dans ce cas-là toniques et prédicatifs, donc acceptables dans ces positions ;
- De certains pronoms placés après le verbe dont ils dépendent (pronoms enclitiques). Ils forment avec le verbe une même unité phonologique et héritent de l'accent du verbe qui se déplace en fin de mot phonologique. Par exemple dans *dis-le-moi*, l'accent de groupe est sur *moi*. Si ce dernier cas est vrai, cela nous conduira à poser que ce sont des règles prosodiques qui l'emportent sur les règles syntaxiques dans le vers ; autrement dit, que les particularités de la syntaxe du vers résultent de contraintes prosodiques. Le principe constituant de la fin du segment métrique serait plutôt l'accent de groupe que la catégorie syntaxique du dernier mot.
- De certains vers qui admettent quelques mots-outils « purs », c'est-à-dire non accentués, en positions 6 et 12.

Nous examinerons en détail les mots-outils en douzième position. Ceci nous permettra d'évaluer à la fois la qualité de l'analyse syntaxique et le type de mots-outils que l'on trouve en cette dernière position.

Nous avons vu précédemment la répartition des mots vides et des mots pleins sur les différentes positions (*cf.* Figure 30), ce qui nous avait conduit à nous demander si les mots vides étaient interdits en sixième et en douzième positions. En réalité, la position douze est occupée à 98,7 % par un mot plein et la position six à 95,6 % et non pas à 100 %. Ces résultats sont sujets à deux

interprétations : ou bien l'on considère que la douzième position découpe de manière très nette dans le lexique deux grandes catégories, celle des mots pleins et celle des mots vides, et donc que les mots vides identifiés par le métromètre ont en réalité une fonction de mots pleins, ou bien l'on part de l'hypothèse que les mots-outils ne sont pas admis en position six et encore moins en position douze, ce qui nous amène à examiner les exceptions qui devraient toutes être des erreurs d'analyse syntaxique, ou bien des violations sévères de la règle.

La réalité se situe dans une position intermédiaire : certains mots-outils sont admis en sixième et douzième, mais portent un accent, d'autres sont présents en tant que mots-outils et représentent de véritables violations, d'autres enfin relèvent d'erreurs d'analyse syntaxique.

Mots-outils en position douze

Examinons donc tous ces vers dont la position douze est occupée par un mot analysé comme mot-outil (déterminant ou pronom, préposition, conjonction, relatif) par l'analyseur syntaxique.

Comme précédemment, nous n'avons retenu dans CORRAC que les vers de douze positions (moyen assez radical pour éliminer les vers mal analysés) et la première interprétation syntaxique. Dans les décomptes effectués ci-dessous, nous avons éliminé les erreurs d'analyse syntaxique liées à des homographes (*lui* de *luire*, *vers* comme nom et non pas préposition, *tu* de *taire*...).

- Déterminants et pronoms

Commençons par les déterminants et pronoms. On ne relève aucun cas de déterminant en position 12. Nous nous intéresserons donc surtout à la question des pronoms. Rappelons pour commencer l'organisation des pronoms personnels : il faut distinguer les formes disjointes des formes conjointes qui n'ont pas le même fonctionnement syntaxique.

Nombre	Personne	Formes conjointes			Formes disjointes
		cas sujet	cas objet direct	cas objet second	
Singulier	1re	je	me, moi		moi
	2e	tu	te, toi		toi
	3e masc	il	le	lui	lui
	3e fém	elle	la	lui	elle
Pluriel	1re	nous	nous		nous
	2e	vous	vous		vous
	3e masc	ils	les	leurs	eux
	3e fém	elles	les	leurs	elles
	3e réfléchie	-	se		soi
Neutralisation	indéfinie	on	se		soi

Figure 32. Tableau des pronoms [Arrivé et al., 1986, p. 499]

Pour la catégorie des pronoms, on trouve en douzième position :

- des pronoms personnels à valeur substantive, ou prédicative (en ancien français, il s'agissait de pronoms toniques, donc accentués, qui avaient le droit d'occuper la dernière position du vers). Il s'agit des formes disjointes des pronoms personnels ou, dans la terminologie de Tesnière, des substantifs personnels qui appartiennent à la catégorie générale des substantifs généraux, par opposition aux substantifs particuliers (*cheval*, *Alfred*) [Tesnière, 1959, 1982, p. 66]. Ce sont principalement des pronoms personnels après préposition. Ils auraient dû être systématiquement analysés comme des noms, ce qui n'a pas toujours été le cas. Dans le vers suivant, *lui* est, par exemple, analysé à tort comme pronom :

> Le poursuivre, le perdre, et mourir après *lui*
>
> Corneille, *Le Cid*, vers 848.

On trouve également en douzième position :

- des pronoms numéraux :

> Je me procure un mal pour en éviter mille.
>
> Corneille, *La Place royale*, vers 250.
>
> Dis-lui que pour demain il choisisse un second
>
> Corneille, *La Suivante*, vers 1204.

On remarquera l'absence notable de déterminants numéraux en douzième position.

- des pronoms personnels enclitiques rattachés au verbe : leur présence en fin de vers est assez ambiguë. Ils appartiennent en effet en tant qu'indices personnels à la catégorie des mots vides et ne portent donc pas *a priori* d'accents. Cependant Henri Bonnard remarque que :

> Les formes conjointes (pronoms personnels qui précèdent ou suivent obligatoirement le verbe avec éventuellement *ne*, *en* et *y* entre les deux) sont souvent appelées atones, et elles le sont quand elles précèdent le verbe, mais elles cessent de l'être quand elles le suivent, à moins d'être suivies d'une autre forme conjointe.
>
> [Bonnard, 1981, p179].

On trouve des formes conjointes qui ont des fonctions de sujet, objet et objet second.

> Quel était ton dessein et que prétendais-tu ?
>
> Corneille, *Cinna*, vers 1499.

> Si vous me préférez, seigneur, punissez-la :
>
> Corneille, *Attila*, vers 1176.

> Et vous-même, seigneur, après tout, dites-moi,
>
> Corneille, *Œdipe*, vers 315.

En revanche, aucun pronom conjoint placé avant le verbe n'apparaît en douzième position, sauf ce vers exceptionnel, repéré par Roubaud [*1978*, 1988, p. 100][136] :

> Il monte à son retour, il frappe à la porte : elle
> Transit, pâlit, rougit, me cache en sa ruelle,
>
> Corneille, *Le Menteur*, vers 621-622.

C'est un violent outrage que la syntaxe inflige au vers, en séparant le pronom personnel du verbe. Il semble convenir à la situation. L'effet stylistique d'une telle construction est évident.

- Prépositions

[136] Cornulier [1995, p. 108] considère que *elle* peut s'interpréter comme la forme autonome, disjointe et tonique du pronom sujet et donc que « l'enjambement quoique notable est moins discordant... et plus vraisemblable pour l'époque ». Mais cette forme autonome ne fonctionne jamais seule. Il aurait fallu avoir : Elle, elle transit...

Dans le vers classique, on ne trouve aucun cas de préposition ou locution prépositionnelle en douzième position. Cela signifierait que le complément introduit par la préposition serait détaché de la préposition (enjambement ou rejet), ce qui créerait une distorsion très violente entre métrique et syntaxe. En effet, dans ce cas, le mot phonologique ne coïnciderait pas avec le vers.

Le métromètre trouve cependant des mots en douzième position, qu'il analyse à tort comme des prépositions. Ce sont des mots qui peuvent avoir tantôt la fonction d'adverbe, tantôt celle de préposition : *près*, *dedans*, *devant*, *derrière*, *au-dessous*... mais ils sont toujours employés dans le vers comme adverbes :

> Pour ne rien hasarder, cachez-vous là derrière.
>
> Corneille, *Clitandre*, vers 618.

Ont été analysés comme prépositions — à tort — les présentatifs *voici* et *voilà* (dans certains cas c'est effectivement leur fonction, quand ils introduisent un complément circonstanciel de temps : *Il s'est marié voilà dix ans* [Bonnard, 1982]). Les présentatifs constituent à eux seuls des groupes verbaux et ont pour fonction d'introduire des noms, des infinitifs ou des propositions introduites par *que* et ses variantes [Arrivé et al, 1986].

Dans le vers classique, les présentatifs *voici* ou *voilà* peuvent être présents en douzième position, à condition d'être précédés par ce qu'ils ont fonction d'introduire. *A priori*, le groupe introduit ne peut figurer en aucun cas sur le vers suivant.

> Votre Émilie en est, seigneur, et la *voici*.
>
> Corneille, *Cinna*, vers 1563.

> N'avez-vous pas reçu de l'huissier que *voilà*
>
> Racine, *Les Plaideurs*, vers 473.

On trouve cependant une seule exception à cette règle dans le vers classique, exception qui contribue à renforcer l'effet de surprise :

> Je vois déjà Mélite. Ah ! Belle ombre voici
> L'ennemi de votre heur qui vous cherchait ici :
>
> Corneille, *Mélite*, vers 1497-1498.

- Conjonctions de coordination

Les conjonctions de coordination sont très rares en douzième position. On trouve un *car* chez Racine, dans *Les Plaideurs*, comédie où se réfugient toutes les excentricités métriques de Racine (au point que certains ont pu se demander si Racine en était bien l'auteur).

> Je veux dire la brigue et l'éloquence. Car
> D'un côté le crédit du défunt m'épouvante ;
>
> Racine, *Les Plaideurs*, vers 729-730.

Le mot *donc* apparaît parfois en fin de vers, mais il n'a pas la fonction de conjonction de coordination. En effet, *donc* peut aussi avoir une fonction d'interjection, comme dans l'exemple suivant, également tiré des *Plaideurs :*

> CHICANNEAU.
> Oui-da : je verrai bien s'il est sergent.
> L'INTIMÉ, EN POSTURE D'ÉCRIRE.
>
> Tôt donc,
> Frappez : j'ai quatre enfants à nourrir.
> CHICANNEAU.
>
> Ah ! Pardon !
>
> Racine, *Les Plaideurs*, vers 429-430.

- Relatifs

Cette catégorie est censée regrouper des mots qui ont pour fonction d'introduire une relative. Dans le lexique, les mots qui ont souvent cette fonction sont : les pronoms relatifs, les conjonctions de subordination, mais aussi les adverbes comparatifs, interrogatifs et d'intensité. Elle est très hétéroclite. Les pronoms relatifs et les adverbes introduisant une proposition, quand ils se trouvent en position 12, introduisent une forte distorsion entre mètre et syntaxe, puisque la proposition est différée sur le vers suivant, comme le montrent les exemples suivants :

Simple, apprends que ta sœur n'aura jamais de quoi
Asservir sous ses lois des gens faits comme moi ;

Corneille, *La Veuve*, vers 99-100.

Disposez-en, madame, et me dites en quoi
Vous avez résolu de vous servir de moi.

Corneille, *Le Menteur*, vers 959-960.

Ce titre en eux me choque, et je ne sais pourquoi
Un roi que je commande ose se nommer roi.

Corneille, *Attila*, vers 23-24.

Enfin, au bout d'un an, sentence par laquelle
Nous sommes renvoyés hors de cour. J'en appelle.

Racine, *Les Plaideurs*, vers 208-207.

Cependant, *quoi* et *pourquoi* sont des formes toniques, et peuvent être grâce à leur accent placés en douzième position. Le dernier exemple introduit une distorsion métrique plus grande.

Mots-outils en sixième position

Nous avons également identifié les mots-outils en sixième position. Le métromètre indique 3243 vers avec un mot-outil. Comme précédemment, nous les avons examinés, éliminé les cas d'erreur d'analyse syntaxique pour ne retenir que les cas « réels » de mots-outils en sixième position.

- On trouve de rarissimes cas de pronoms conjoints, deux dans *Les Plaideurs* :

Il se tourmente ; il vous... fera voir aujourd'hui

Racine, *Les Plaideurs*, vers 360.

Et concluez.
L'INTIMÉ, *d'un ton pesant*.
Puis donc, qu'on nous, permet, de prendre,
Haleine, et que l'on nous(,) défend, de nous, étendre,

Racine, *Les Plaideurs*, vers 791-792.

Les points de suspension dans le premier cas, l'indication de ton et la virgule (qui n'apparaît pas dans l'édition d'origine de P. Mélèse) dans le second, introduisent des éléments de diction qui marquent la frontière des hémistiches, là où la syntaxe passe par-dessus la césure.

- On trouve une conjonction :

Je les fais taire, mais, Seigneur, à le bien prendre,

Corneille, *Suréna* vers 841.

- Les cas de relatifs sont beaucoup plus nombreux en sixième qu'en douzième position, mais sont de même nature.
- La différence majeure dans le traitement de la fin d'hémistiche par rapport à la fin de vers tient à la présence de conjonctions de subordination, de locutions prépositionnelles et de prépositions en sixième position.

Quoi qu'il en soit, depuis que je vous vois chez elle,

Corneille, *L'Illusion comique*, vers 541.

Si nos prêtres, avant que de sacrifier,

Corneille, *Horace*, vers 1773.

Je vous aime, et depuis que j'ai vu la lumière,

Corneille, Psyché, vers 1467.

Je me jette au-devant du coup qui t'assassine,

Corneille, *Pompée*, vers 1382.

Met la revanche hors de mon peu de pouvoir,

Corneille, *La Veuve*, vers 1602.

Si toutefois, après ce coup mortel du sort,

Corneille, *Polyeucte*, vers 559.

Quoi, Madame, parmi tant de sujets de crainte,

Racine, *Iphigénie*, vers 1021.

Je quitterais, et pour, / Hé bien, / Pour quelle mère,

Racine, *Athalie*, vers 100.

Il s'agit à chaque fois de mots-outils qui occupent au moins deux positions métriques, et qui, en raison de leur longueur, portent un accent sur leur dernière position. On ne trouve que deux exceptions : avec *hors* (dans *hors de*) et avec *pour* (dans *Les Plaideurs*, pièce qui est, rappelons-le, un concentré de transgressions).

Le tableau suivant résume les cas réels de mots-outils en sixième et douzième positions, une fois contrôlée chacune des propositions du métromètre. Nous ne recensons pas les mots-outils à valeur substantive qui sont tantôt considérés comme des noms (*moi*, *toi*, *nous*, *vous*) tantôt comme des mots-outils (*lui*) par le métromètre, ni ceux qui sont en réalité des adverbes... Pour la catégorie des « relatifs », nous n'avons retenu que ceux qui introduisaient une proposition reportée sur le segment métrique suivant. L'ensemble des

exemples, sauf les cas de pronoms conjoints post-posés, figurent dans l'annexe VII.

	Position 6		**Position 12**	
Type de mot-outil	**Corneille**	**Racine**	**Corneille**	**Racine**
Pronom conjoint post-posé	467	317	75	56
Indice personnel (pronom conjoint, non prédicatif)	0	2	1	0
Déterminant, adjectif possessif, démonstratif ou numéral	0	0	0	0
Préposition (y. c. locutions prépositionnelles)	28	15	0	0
Locution prépositionnelle	21	2	0	0
Préposition	7	13	0	0
Conjonction de coordination	1	0	0	1
Conjonction de subordination	3	3	0	0
Relatifs, interrogatifs, comparatifs	28	20	3	1
Nombre « réel » de mots-outils en fin de segment	**527**	**357**	**79**	**58**
% de mots-outils en fin de segment	0,93%	1,94%	0,14%	0,32%
Nombre de mots-outils trouvés par le métromètre en fin de segment	2409	833	682	268
Taux d'erreur du métromètre	3,33%	2,58%	1,07%	1,13%
Nombre de vers	56478	18438	56478	18438

Tableau 26. Répartition des mots-outils en positions 6 et 12

Les mots-outils ont une fonction de ciment syntaxique, ils ont principalement un rôle de jonction ou de translation [Tesnière, 1959, 1982]. En tant que jonctifs[137], ils ont tendance à être (en français) entre deux mots pleins, et en tant que translatifs[138], à précéder le mot-plein. En tout état de cause, le mot-outil ou mot vide a une très faible probabilité de se situer en fin de groupe syntaxique ou de mot phonologique. Il en résulte que dans le vers, s'il est placé en dernière position, ce qu'il introduit (joint ou translate) sera rejeté sur le vers suivant (enjambement). Le mot-outil en dernière position entraîne une

[137] « "mots vides" ayant pour fonction d'unir entre eux les "mots pleins" » ou les nœuds qu'ils forment [Tesnière, 1959, 1982, p. 668].

[138] « "mots vides" ayant pour fonction de transformer la catégorie des "mots pleins" » [Tesnière, 1959, 1982, p. 670].

distorsion entre la structure métrique et la structure syntaxique, inacceptable dans le cadre de l'esthétique classique.

Cependant, les mots-outils qui, par une opération de translation, entrent dans la catégorie des mots pleins, ont un fonctionnement différent : ils peuvent occuper la position finale du segment métrique. C'est en particulier le cas des pronoms disjoints toniques à valeur substantive. Les pronoms enclitiques post-posés peuvent également figurer en fin de segment métrique car ils constituent avec le verbe dont ils dépendent une seule unité phonologique. Si nous trouvons certains pronoms relatifs, ceux-ci sont en général toniques : *en quoi*, *pourquoi*, et donc plus acceptables pour l'oreille. Enfin, pour les prépositions, locutions prépositionnelles et conjonctions de subordination trouvées en sixième position, la plupart ont pour caractéristique d'occuper au moins deux positions métriques : *depuis*, *malgré*, *après*, *au-dessus de*, *depuis que*, *avant que*... En tant que polysyllabes, ils portent une forme d'accent.

VII.3.3. Conclusion

Le métromètre met en évidence des répartitions contrastées de chaque catégorie syntaxique sur les douze positions. Les catégories syntaxiques, comme tous les marquages vus jusqu'à maintenant, délimitent dans le vers deux segments sur lesquels les courbes de répartition pour chaque catégorie sont très semblables. La part des mots pleins augmente quand on avance vers la fin du segment métrique et devient quasiment exclusive. Inversement, les mots-outils occupent principalement les positions initiales du vers. Tout se passe comme si l'avancée vers la plénitude du sens, l'accomplissement du mot phonologique, s'accompagnait d'un enrichissement de la matière graphique et sonore, que nous avons décrite dans la première section (plus on avance dans le segment métrique, plus la consistance des syllabes métriques augmente). Les mots-outils souvent plus grêles occupent plutôt les premières positions. Par cette organisation syntaxique et sonore, le contraste entre la fin d'un segment et le début du suivant (fin d'hémistiche par rapport au début de l'hémistiche suivant, fin de vers par rapport au début du vers suivant) devient très marqué. On peut même se demander si la métaposition à l'intérieur d'un hémistiche ne contribue pas en partie à placer la préposition en début de segment et à accroître le

contraste entre la fin d'un segment et le début du segment suivant. Ainsi, dans le vers suivant :

> Soumis à tous leurs vœux, à mes désirs contraire,
>
> Racine, *Britannicus*, vers 1335.

l'inversion, qui produit un chiasme, n'est pas imposée par le mètre, elle l'est sans doute par la rime. Cependant, l'inversion permet aussi de placer la préposition *à* en première position du segment, accentuant le contraste avec la fin du segment précédent constituée d'un mot plein.

Les verbes occupent plutôt les positions centrales, tandis que les noms sont très majoritaires en positions 6 et 12. Les allures syntaxiques des deux hémistiches sont proches, même s'il existe quelques différences significatives : les noms et les adjectifs sont plus fréquents sur le second hémistiche, tandis que verbes, adverbes et noms propres le sont davantage sur le premier. Les prépositions sont très représentées en position 7, les conjonctions en position 1.

Sur un plan plus technique, les observations sur la présence des mots-outils en fin de segment métrique, en positions 6 et 12 montrent que :

- Le métromètre, en raison des difficultés liées à l'analyse syntaxique, est un outil imparfait pour la syntaxe, puisque la recherche des mots vides en douzième position sur près de 75000 alexandrins de CORRAC conduit à 950 vers, sur lesquels seuls 139 sont dotés de vrais mots-outils en douzième position. La grosse majorité des erreurs provient d'une mauvaise analyse des pronoms personnels disjoints ou substantifs personnels ayant une fonction de substantif.
- Globalement, dans notre corpus, les infractions à l'unité syntaxique des hémistiches sont extrêmement rares. Le nombre très faible de mots-outils de début de syntagme dans les positions finales de segment métrique témoigne de l'effort entrepris pour faire coïncider syntagme et segment métrique. Dans le vers classique, la métrique domine la syntaxe en lui imposant des distorsions comme la métaposition. Cette domination s'assouplira au XIXe siècle, notamment à la suite des transgressions de Hugo.

- Les mots-outils présents en positions 6 ou 12 sont le plus souvent des formes toniques :

1. pronoms personnels disjoints à valeur tonique,
2. pronoms conjoints placés après le verbe, qui héritent de l'accent du verbe,
3. mots-outils polysyllabiques qui acquièrent un accent en dernière position.

La contrainte de l'accent semble donc plus forte que celle de la catégorie syntaxique : en posant que la dernière position du vers doit toujours être tonique, on en déduit immédiatement que les mots-outils sont très rares en fin de vers, car très peu d'entre eux sont toniques.

Cependant, les mots-outils, par leur présence rarissime en positions 6 et 12, acquièrent une valeur *métrique*. Les distributions contrastées des différentes catégories syntaxiques, l'opposition dans la répartition des mots pleins et des mots-outils sur les différentes positions et le contraste entre la fin d'un segment et le segment suivant contribuent à l'élaboration du rythme des vers.

VII.4. LA MARQUE ACCENTUELLE

Il nous reste maintenant à explorer la distribution sur notre corpus de la marque accentuelle, seul marquage prosodique que nous ayons finalement retenu.

Nous avons vu dans ce chapitre consacré aux marquages que la prosodie pouvait aussi être approchée par la ponctuation. Dans la seconde partie de Dynastie, Roubaud [1988] a utilisé la marque de la ponctuation pour l'analyse du vers classique. Le poids 1 est attribué aux virgules, le poids 3 aux points et le poids 2 à toutes les ponctuations intermédiaires. Le poids est affecté à la voyelle précédant le signe de ponctuation. Mais si celle-ci est un *e* muet, le poids est alors affecté à la position précédente. Sur les mille premiers vers de *Bérénice*, les résultats sont les suivants :

	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7	P. 8	P. 9	P. 10	P. 11	P. 12
Marque fonda-mentale	99	142	112	129	8	447	36	47	38	5	0	1772

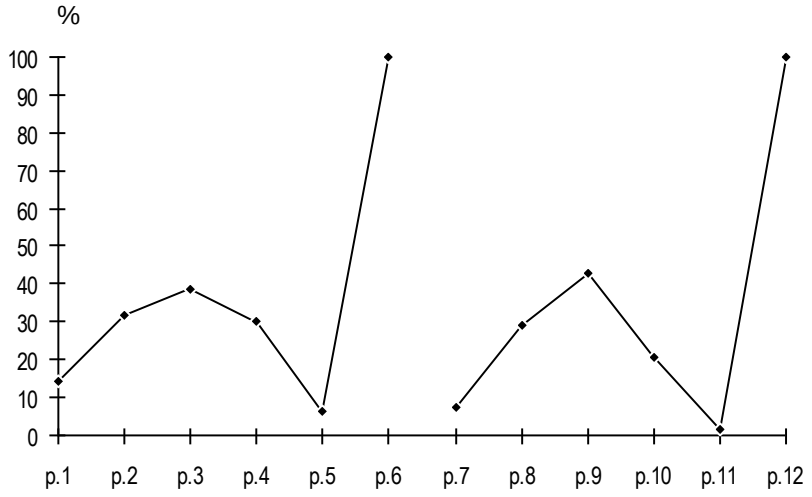
Clef de lecture : sur les mille premier vers de *Bérénice*, le marquage de ponctuation retenu donne à la position 1 le poids 99 et à la position 12 le poids 1772.

Tableau 27. La marque de ponctuation sur les mille premiers vers de* Bérénice *[Roubaud, 1988, p. 52]

A sa manière, le marquage de la ponctuation est un indicateur de la position des pauses prosodiques et permet d'identifier la position de la césure. Comme nous l'avons vu dans le chapitre consacré aux limites du métromètre, nous n'avons pas introduit dans la version actuelle de cet outil de marquage de ponctuation. Mais les résultats de Roubaud sont suffisamment éloquents.

Toujours dans le même article, Roubaud avait défini la « marque fondamentale » comme suit : marquage de la dernière position ou avant-dernière (si la dernière est occupée par un *e* muet) des mots pleins (noms, verbes, adjectifs, adverbes). Ce marquage, appliqué aux mille vers de *Bérénice*, conduit à la courbe présentée sur la Figure 33.

marque accentuelle	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7	P. 8	P. 9	P. 10	P. 11	P. 12
1000 vers de *Bérénice*	14,5	31,7	38,6	29,9	6,3	99,8	7,4	29,1	43,0	20,4	1,7	100



Clef de lecture : sur les mille premiers vers de *Bérénice*, la position 1 est "marquée" dans 14,5 % des vers.

Figure 33. La marque fondamentale sur les mille premiers vers de* Bérénice *[Roubaud, 1988, p. 49]

On retrouve une forme de distribution symétrique que le lecteur commence à bien connaître. C'est ce dernier marquage, la « marque fondamentale », que nous avons repris dans le métromètre.

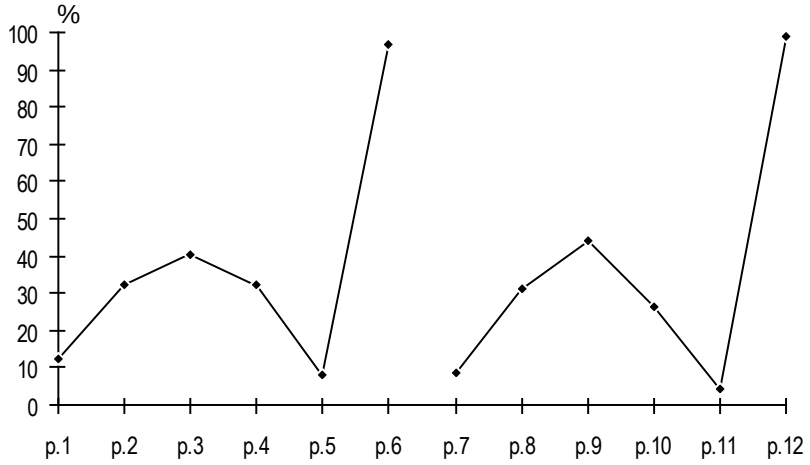
VII.4.1. Le marquage accentuel des positions

Nous avons donc appliqué d'une manière systématique le marquage présenté précédemment : marquage de la dernière position métrique des mots pleins, ou avant-dernière si le mot a une terminaison féminine[139]. La Figure 34 présente le résultat du marquage sur le corpus CORRAC. L'allure de la courbe

[139] Dans le cas d'un pronom post-posé, le marquage est appliqué au mot-outil final : dans *demande-lui*, la marque fondamentale est portée par *lui* (*cf.* chapitre V).

est sensiblement identique à celle construite à partir des mille premiers vers de *Bérénice*.

marque accentuelle	**P. 1**	**P. 2**	**P. 3**	**P. 4**	**P. 5**	**P. 6**	**P. 7**	**P. 8**	**P. 9**	**P. 10**	**P. 11**	**P. 12**
CORRAC	12,6	32,1	40,1	32,0	8,3	96,6	8,8	31,0	44,0	26,2	4,6	98,9



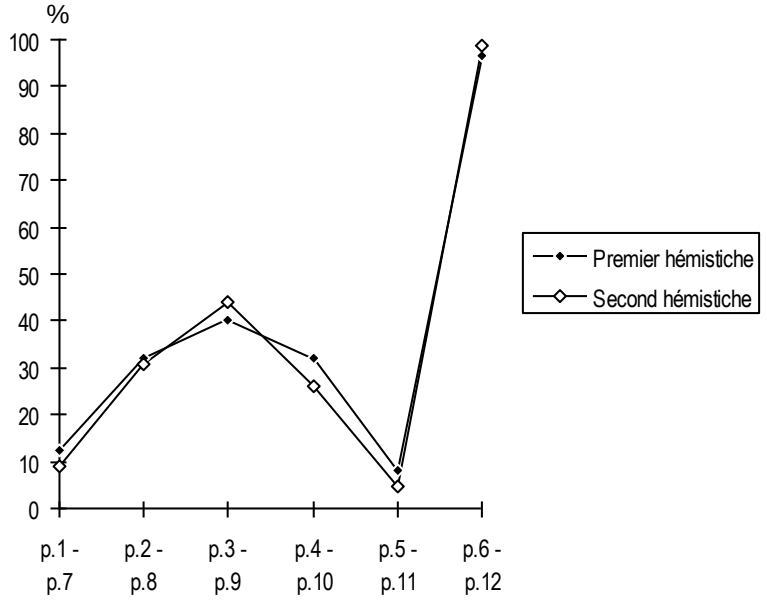


Figure 34. Répartition des « marques accentuelles » sur CORRAC

Ce marquage accentuel (*cf.* Figure 34) délimite à l'intérieur du vers deux segments métriques identifiés par deux éléments :

- le marquage systématique des positions 6 et 12 (le taux de marquage n'est pas égal à 100%, comme il le serait si le traitement avait été manuel..., en raison d'erreurs d'analyse syntaxique : les pronoms disjoints, comme *lui*, qui sont porteurs d'accents, ont été analysés comme des pronoms conjoints...) ;
- le parallélisme des figures accentuelles des deux segments métriques.

On retrouve la similitude des figures du premier et du second hémistiche, déjà observée pour tous les autres marquages. Il faut cependant signaler quelques différences. Sur le second hémistiche, la position 9 est sensiblement plus marquée que la position 3. Inversement, la position 10 connaît une déficience de marquage par rapport à la position 4. Enfin, sur le second hémistiche, les écarts de marquage entre positions voisines sont plus accentués : entre les positions 12 et 11, on trouve un écart de 94 points, et de 88 points entre les positions 6 et 5. La position 12 est mise en valeur par la diminution des marquages sur les positions 10 et 11.

Le marquage que nous venons de voir repose sur une hypothèse de langue, l'existence d'un « accent » final en français (*cf.* chapitre V.4). Nous cherchons par ce marquage à placer les accents du discours dans l'espace métrique du vers. Cependant, cette notion même d'accent, est problématique en français surtout quand il s'agit d'accents de mots. Nous avons donc imaginé une autre approche en partant d'une hypothèse métrique. Rappelons cette approche déjà décrite au chapitre V.

Comme la position 12 est la position la plus éminente du vers puisqu'elle marque à la fois la fin de l'hémistiche et la fin du vers et qu'elle porte la rime, on utilise ses caractéristiques pour définir un marquage. Autrement dit, la position 12, qui est indiscutablement accentuée, va nous servir d'étalon de marquage. Elle correspond à la dernière position d'un mot ou à l'avant-dernière si le mot se finit par un *e* muet : c'est donc sur la dernière ou avant-dernière position de chaque mot que s'appliquera le marquage.

La répartition des catégories syntaxiques en fin de vers nous montre que 54 % des fins de vers correspondent à des noms, 20 % à des verbes...

Répartition des catégories en P. 12	**%**
Nom commun	54,04
Verbe	19,88
Adjectif	17,98
Adverbe	3,93
Patronyme	2,89
Déterminant, pronoms	1,16
Préposition	0,08
Relatif	0,02
Conjonction	0,01

Tableau 28. Répartition des catégories syntaxiques en position 12

Nous affecterons à la dernière (ou avant-dernière) position d'un mot le poids de sa catégorie syntaxique en position 12 (54,04 si c'est un nom, 19,88 si c'est un verbe...) [140] :

De cette nuit, Phénice, as-tu vu la splendeur ?

Racine, *Bérénice*, vers 302

syllabes métriques	d ə	s ɛ	t əf	n ɥ i	f e	n i s	a	t y	v y	l a	sp l ɑ̃	d œ r
marquage accentuel	0,08	1,16	0	54,04	0	2,89	0	1,161	17,98	1,16	0	54,04

Nous avons réalisé cette opération sur tous les vers de CORRAC. Le poids moyen de chaque position qui en résulte est présenté dans la Figure 35.

[140] On affecte le poids de la catégorie syntaxique (défini par la répartition des catégories en position 12) à une position métrique, uniquement si celle-ci correspond à la dernière position d'un mot ou à l'avant-dernière si le mot finit par un *e* muet. Dans le vers suivant, la troisième position (première syllabe de *crime*) aura pour valeur 0,54 et la quatrième aura pour valeur 0 :

Par quel **cri**me, seigneur, ou par quelle imprudence (Corneille, *Suréna*, vers 900)

Position	Marquage de la dernière ou avant-dernière position des mots
P. 1	2,7
P. 2	8,6
P. 3	10,3
P. 4	7,1
P. 5	1,4
P. 6	32,1
P. 7	1,8
P. 8	8,2
P. 9	12,3
P. 10	6,2
P. 11	0,7
P. 12	36,6

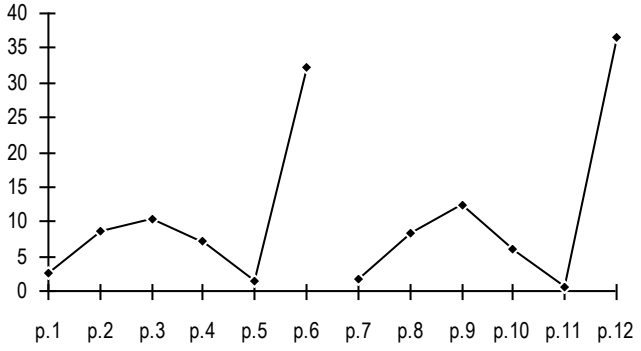


Figure 35. Poids accentuels obtenus avec la position 12 comme étalon de marquage

La figure ressemble étrangement à celle que l'on obtient avec le simple marquage des fins de mots pleins : même allure des hémistiches et mêmes différences entre les deux segments, avec un poids plus fort sur le second hémistiche de la position 9 par rapport à la position 3.

Il existe donc une concordance très forte entre un marquage construit à partir de données de langue et un marquage construit à partir d'une hypothèse métrique qui considère que la douzième position est la plus saillante du vers. Cette concordance montre que le vers mobilise et informe à sa manière les

caractéristiques de la langue de manière si puissante qu'il arrive à imposer sa forme à travers les indicateurs les plus incertains. Cette résistance à la déconstruction force notre admiration.

VII.4.2. Les profils prosodiques des hémistiches

La figure globale du vers construite statistiquement à partir du marquage « accentuel » masque la variété des séquences métrico-rythmiques des hémistiches. Ces séquences se définissent par la succession de six événements élémentaires qui peuvent prendre les valeurs « 1 » ou « 0 », selon qu'ils portent ou non une marque accentuelle. Voici, à titre d'exemple, quelques types d'hémistiches :

Approu**vez** ma fai**bless**e, et souf**frez** ma dou**leur ;**
H1_001001 H2_001001
Corneille, *Horace*, vers 1.

Qui du **moin**dre pé**ril** se **fait** une infor**tu**ne ;
H1_001001 H2_010001
Corneille, *Horace*, vers 16.

Je **sais** qu'il **doit** s'ac**croî**tre, et que tes **grands** des**tins**
H1_010101 H2_000101
Corneille, *Horace*, vers 41.

Albe, mon **cher** pa**ys**, et mon pre**mier** a**mour ;**
H1_100101 H2_000101
Corneille, *Horace*, vers 30.

Vous mou**rez** ? / **Jus**te **ciel** ! **qu'ai**-je **fait** aujourd'**hui** ?
H1_001101 H2_101001
Racine, *Phèdre*, vers 839.

Dans le marquage simple que nous avons adopté, aucune règle de réduction de séquence [Roubaud et Lusson, 1974] ou d'effacement d'accents en contexte [Dinu, 1993] n'a été introduite. Ces règles ont en général pour effet de modifier les marquages quand des accents apparaissent sur des positions adjacentes (soit en supprimant les deux accents, soit en supprimant le plus faible au bénéfice de l'autre). Elles font tendre la séquence concrète vers une séquence plus régulière. Ainsi, grâce à ces règles d'effacement qui s'appliquent principalement à des positions marquées voisines, toutes les séquences peuvent finalement être réduites à des séquences de type 010101 ou 001001 et à des

séquences 000001 au statut indécidable. Par commodité, mais sans que jamais il n'y ait eu la moindre assimilation avec des métriques quantitatives ou accentuelles[141], les deux premières séquences ont été désignées par Roubaud et Lusson [1974] comme *ïambe* et *anapeste*. Pour éviter les critiques que ces appellations ont suscité, nous ne les désignerons que par leur forme : 001001 et 010101.

Nous n'avons pas appliqué de règles de réduction pour différentes raisons. Tout d'abord, le fondement phonétique des filtres de réduction n'est pas toujours très clair. Ensuite et surtout, notre marquage accentuel souffre d'un grave défaut : il marque les « accents » de mots et non les accents de groupes qui sont les seuls admis en français. Il ne hiérarchise évidemment pas non plus les accents au sein du groupe. Comment procéder à des réductions de séquences quand les marquages se limitent à des « 0 » et des « 1 » ?

Si nous avions pu mettre en place un marquage hiérarchisé d'accent, la réduction des séquences aurait été assez simple : nous aurions conservé comme positions marquées celles qui correspondent aux événements les plus éminents localement, c'est-à-dire davantage « marqués » que les positions adjacentes.

Si l'on refuse de réduire les séquences métriques construites à partir de la marque accentuelle, la variété des séquences devient très grande. Le Tableau 29 présente les différents types d'hémistiches que l'on trouve chez Corneille et Racine. Il y a 61 séquences différentes (sur les 64 possibles) dont trois n'apparaissent que sur le premier hémistiche. Les séquences les plus fréquentes correspondent à des segments métriques au profil mélodique régulier : les segments marqués sur au moins deux positions métriques paires (010001, 000101 et 010101)[142] et ceux qui sont marqués sur les positions multiples de trois (001001). Il faut ajouter, parmi ces séquences les plus fréquentes, la séquence 000001 qui est à la fois compatible avec 001001 et avec 010101.

---

[141] En effet, ni la quantité des syllabes, comme dans la poésie gréco-latine, ni l'accent, comme dans la poésie anglaise où il existe réellement un accent fixe de mot, n'interviennent dans le vers français ; la notion de pied non plus n'appartient pas à la versification française.

[142] On peut regrouper les séquences 010001 et 000101 avec la séquence 010101 car elles sont compatibles avec le mètre 010101.

Premier hémistiche		
	fréq.	**%**
001001	17963	24. 0
010001	11970	16. 0
000101	10967	14. 6
000001	6380	8. 5
010101	4833	6. 5
011001	2839	3. 8
001101	2692	3. 6
101001	2629	3. 5
100101	1897	2. 5
100001	1855	2. 5
000011	1703	2. 3
001011	1242	1. 7
010011	1116	1. 5
110001	1009	1. 3
000100	610	0. 8
001000	605	0. 8
000111	588	0. 8
110101	389	0. 5
011101	377	0. 5
101101	373	0. 5
100011	301	0. 4
010100	279	0. 4
010111	252	0. 3
011011	218	0. 3
101011	202	0. 3
001100	196	0. 3
111001	191	0. 3
010000	153	0. 2
001111	119	0. 2
011000	103	0. 1
110011	98	0. 1
100111	97	0. 1
100100	94	0. 1
101000	90	0. 1
001010	66	0. 1
010010	53	0. 1
000010	39	0. 1
011100	39	0. 1
000000	38	0. 1
110100	31	0. 0
110111	31	0. 0
111101	28	0. 0
101100	20	0. 0
000110	18	0. 0
111011	15	0. 0
101010	14	0. 0
110000	12	0. 0
010110	10	0. 0
100010	10	0. 0
011111	9	0. 0

Second hémistiche		
	fréq.	**%**
001001	23837	31. 8
010001	14213	19. 0
000101	11016	14. 7
000001	6678	8. 9
010101	3648	4. 9
011001	2740	3. 7
001101	2119	2. 8
101001	2091	2. 8
100001	1625	2. 2
100101	1145	1. 5
000011	926	1. 2
001011	888	1. 2
110001	773	1. 0
010011	577	0. 8
000111	354	0. 5
011101	241	0. 3
110101	232	0. 3
101101	218	0. 3
000100	194	0. 3
001000	194	0. 3
011011	163	0. 2
111001	158	0. 2
010111	125	0. 2
010100	112	0. 1
100011	89	0. 1
101011	80	0. 1
001100	61	0. 1
001111	50	0. 1
010000	47	0. 1
101000	46	0. 1
110011	43	0. 1
100111	39	0. 1
011000	38	0. 1
100100	30	0. 0
010010	16	0. 0
001010	14	0. 0
000010	11	0. 0
000110	10	0. 0
110111	9	0. 0
011100	8	0. 0
110100	8	0. 0
111011	8	0. 0
111101	7	0. 0
000000	5	0. 0
010110	5	0. 0
111000	5	0. 0
011111	4	0. 0
101100	3	0. 0
101111	3	0. 0
110010	3	0. 0

Premier hémistiche		
	fréq.	**%**
100000	9	0. 0
101111	9	0. 0
100110	8	0. 0
111000	8	0. 0
011010	7	0. 0
110010	5	0. 0
001110	4	0. 0
011110	2	0. 0
101110	1	0. 0
111010	1	0. 0

Second hémistiche		
	fréq.	**%**
011010	2	0. 0
100000	1	0. 0
100110	1	0. 0
101010	1	0. 0
110000	1	0. 0
110110	1	0. 0
111111	1	0. 0

Clef de lecture : la séquence 001001 apparaît 17963 fois sur le premier hémistiche (24 %) et 23837 fois sur le second (31,8 %).

Tableau 29. Figures accentuelles des premier et second hémistiches

Les formes de type 001001 et 010101 (et assimilées) apparaissent dans 65% des hémistiches. Les séquences marquées en positions paires (38 %) sont plus fréquentes que les séquences 001001 (28 %). Cette comparaison n'a cependant pas grand sens dans la mesure où 35% des hémistiches n'ont pas été assimilés à l'un de ces deux modèles.

Premier hémistiche	%	**Second hémistiche**	%
001001	24,0	**001001**	31,8
010101	37,0	**010101**	38,6
010001	16,0	**010001**	19,0
000101	14,5	**000101**	14,7
010101	6,5	**010101**	4,9
Autres figures	39,0	**Autres figures**	29,6
Effectifs	74917	**Effectifs**	74917

Tableau 30. Les profils prosodiques les plus fréquents

Globalement, sur l'ensemble des hémistiches, les formes régulières apparaissent dans près de deux tiers des cas. Cependant, il y a des différences sensibles entre les deux hémistiches. Sur le premier segment métrique, les formes régulières couvrent 61 % des cas, tandis qu'elles couvrent 70 % des cas sur le second. Le second hémistiche est donc plus régulier que le premier[143].

[143] On voit ici l'intérêt de n'avoir pas réduit les séquences. Si nous l'avions fait, nous n'aurions pu mettre en évidence cette différence entre les deux segments métriques.

Dans la progression du vers, la tendance vers des formes régulières s'accroît. Alors que jusqu'à présent nous avons surtout souligné le parallélisme des deux segments métriques, nous voyons ici apparaître de manière très nette un facteur de différenciation. La répartition des accents est plus régulière sur le second hémistiche que sur le premier et privilégie très nettement la figure 001001.

Cette progression vers la régularité observée dans le passage du premier segment métrique au second est principalement due à l'augmentation de la part absolue et relative de la séquence 001001. Nous avions déjà vu (Figure 34) que sur le second hémistiche, la position 3 était plus marquée que sur le premier.

Ces résultats confirment et complètent ceux qu'avait obtenus Gasparov [1987] sur de petits échantillons. Il a en effet montré pour l'alexandrin français qu'il y avait moins d'accents sur le second hémistiche que sur le premier (pour 100 vers on a de 113 à 117 accents sur les positions 1 à 5 et 106-107 sur les positions 7 à 11)[144]. Il a également montré que la proportion d'anapestes (001001), c'est le terme qu'il emploie, était plus élevée sur le second hémistiche que sur le premier.

VII.4.3. Existe-t-il des régularités dans l'enchaînement des hémistiches ?

Nous venons de voir que le second hémistiche est plus régulier, plus rythmique que le premier, et que le type 001001 y est mieux représenté. On peut alors s'interroger sur la nature de l'enchaînement entre les deux hémistiches : lorsque le poète produit un premier hémistiche d'un certain type, aura-t-il tendance à privilégier tel autre type pour le second segment ? Autrement dit, peut-on déceler un souci d'organisation dans l'enchaînement des segments métriques dans le vers ? Ou les combinaisons relèvent-elles du hasard ?

[144] À propos du tétramètre ïambique anglais, Gasparov a observé que le nombre d'accents tendait à diminuer au fil du vers et que plus le vers tend vers sa fin, plus le nombre d'accents sur les positions impaires diminue.

Nous avons fait un premier test en ne conservant que les vers dont chacun des segments métriques pouvait être rattaché soit à la figure 010101 (y compris 000101 et 010001), soit à la figure 010001. Un test classique du chi-2 effectué sur le croisement de ces deux variables, montre qu'il n'y a pas de dépendance significative entre le premier et le second segment métrique. La figure du second segment métrique ne semble pas dépendre de celle du premier.

Nous avons fait un second test en conservant cette fois-ci tous les vers et en classant les hémistiches en trois groupes : 010101, 001001 et « autres ». Là encore, nous avons constaté qu'il n'y avait pas de dépendance significative entre la structure du premier et celle du second hémistiche.

Il semble exister une autonomie entre la structure du premier et celle du second hémistiche dans le cadre du vers. Qu'en est-il si l'on prend en compte les couples de vers unis par la rime commune, ou distiques ?

Pour répondre à cette question, le travail a été cette fois-ci plus ardu, puisqu'il a fallu en premier lieu identifier les couples de vers rimant ensemble. Ce travail de préparation complexe, produit principalement dans le but d'étudier la rime, a été présenté dans le chapitre III.

Pour commencer, la structure du premier vers ne diffère pas grandement de celle du second, comme le montre le Tableau 31 : le second segment du vers est toujours plus régulier que le premier. La seule différence remarquable vient du troisième segment métrique, plus régulier que le premier : 63 % d'hémistiches de type 010101 ou 001001, contre 60 % pour le premier segment.

CORRAC	Premier segment métrique	Deuxième segment métrique	Troisième segment métrique	Quatrième segment métrique
001001	22,9	32,5	25,1	31,4
010101	36,8	38,2	37,6	37,6
010001	15,8	18,6	16,3	19,3
000101	14,4	14,7	15,0	14,6
010101	6,6	4,9	6,3	4,7
Autres figures	40,3	29,2	37,4	30,1
Effectifs	35135	35135	35135	35135

Tableau 31. Figures des segments métriques sur les couples de vers

Nous avons donc fait un premier test statistique sur les couples de vers composés de segments métriques « réguliers », apparentés à l'un des deux modèles 010101 (y compris 010001 et 000101) et 001001. Il n'y a finalement que 6882 couples de vers, sur plus de 35000, dont tous les hémistiches sont « réguliers ». Une fois encore, aucune liaison significative n'a pu être mise en évidence entre la structure du premier vers et celle du deuxième, ni entre celle des segments métriques pris par couples successifs : pas de lien entre le premier et le deuxième segment métrique, ni entre le deuxième et le troisième, ni entre le troisième et le quatrième.

Un second test conservant tous les couples de vers conduit à des résultats proches. Une dépendance très faible entre la structure du premier vers et celle du second a pu être identifiée, les vers avec un hémistiche « autre » étant *légèrement* plus souvent appariés avec un autre vers ayant un hémistiche « autre ». Quand on considère, non plus le vers, mais les hémistiches deux à deux, la seule dépendance significative observée concerne le second et le quatrième hémistiches : quand le second hémistiche est de type « autre », il y a plus de chances (par rapport au hasard) que le quatrième soit aussi de type « autre ».

La structure accentuelle du vers renforce l'effet de la rime, dans la mesure où les figures d'hémistiches les plus régulières ont une place encore plus importante dans la deuxième partie du vers. Mais il n'y a pas, semble-t-il, dans l'organisation du couple de vers, de structure qui renforce l'organisation du couple par la rime.

VII.4.4. Comparaison avec d'autres résultats

Nous venons, grâce à l'analyse des séquences métrico-rythmiques, des segments métriques de montrer que :

- sur chaque hémistiche, les formes les plus régulières (001001, 010001, 000101 et 010101), proches de formes métriques, sont largement majoritaires (deux tiers des hémistiches) ;
- le second segment du vers est plus structuré que le premier : les formes les plus régulières y sont encore mieux représentées que sur le premier hémistiche ; plus on s'approche de la rime, plus la figure du vers devient régulière ;
- le second hémistiche est plus fréquemment marqué sur sa troisième position que le premier ;
- il n'y a pas de lien entre la structure du premier et celle du second hémistiche ;
- dans les couples de vers rimant ensemble, il n'y a pas de différence significative entre la figure du premier et celle du second vers et aucun type d'enchaînement privilégié d'un hémistiche et d'un vers à l'autre.

Dans un récent article, Mihai Dinu [1993], à partir de l'examen empirique des deux cents premiers vers *d'Andromaque*, présente cinq résultats en partie contradictoires avec les nôtres.

Pour examiner la structure métrique des vers et des couples de vers rimant ensemble, il propose un système de marquage accentuel en quatre niveaux :

3[145] note la syllabe proéminente d'un mot "plein" polysyllabique
2 note un mot "plein" monosyllabique
1 note la syllabe proéminente d'un mot grammatical polysyllabique
0 note toute autre syllabe qui n'appartient pas aux classes précédentes.

[Dinu, 1993, p. 64].

Sur les 200 vers d'*Andromaque*, l'allure de la courbe (*cf.* Figure 36) ressemble à celle que l'on obtient avec le métromètre. Il y a légèrement plus d'accents chez Dinu, parce qu'il marque les mots *grammaticaux* polysyllabiques. On voit déjà sur cette Figure 36 que, relativement à CORRAC, les 200 premiers vers d'*Andromaque* ont une allure assez particulière, principalement sur le premier hémistiche.

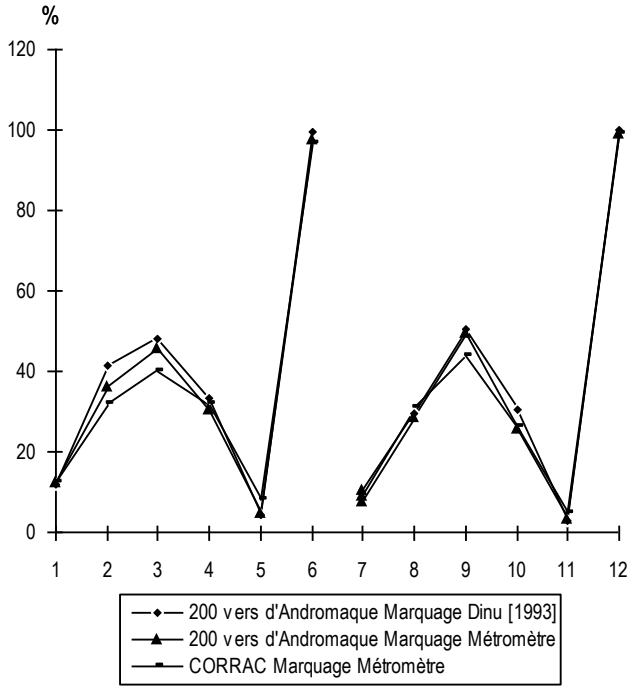


Figure 36. Marquage de Dinu [1993] et marquage du métromètre

L'auteur associe à son système de marquage une « Règle d'effacement en contexte » qui tient compte de la difficulté à prononcer deux accents de suite :

[145] Ne parvenant pas à reproduire les symboles utilisés par Dinu, nous les avons remplacés par des chiffres.

- 2 3 => 0 1
- 3 2 => 1 0
- 3 1 => 1 0
- 2 1 => 1 0
- 1 2 => 0 1
- 2 2 => 0 1.

Ce type de marquage lui permet d'aboutir à cinq schémas accentuels :

1) 010101
2) 010001
3) 000101
4) 001001
5) 000001

En s'appuyant sur ce marquage, il montre dans un premier temps que les hémistiches se répartissent en parts égales entre ïambes (schémas 1, 2 et 3) et anapestes (schéma 4)[146]. Nos propres marquages où, rappelons-le, plus de 30 % des segments métriques ne peuvent être assimilés ni au modèle ïambique, ni au modèle anapestique, montrent cependant une prédominance des segments 010101 par rapport aux segments 001001. Cette différence tient en partie au type de marquage choisi, mais surtout au choix de l'échantillon. Celui-ci explique en partie les distorsions entre les résultats obtenus selon les deux types de marquages. Il apparaît en effet que dans les 200 premiers vers d'*Andromaque*, le type 001001 est très nettement supérieur à ce qu'il est sur l'ensemble de CORRAC.

	premier hémistiche		**second hémistiche**	
	200 vers Andromaque	**CORRAC**	**200 vers Andromaque**	**CORRAC**
001001	31,0	24,0	41,5	31,8
010101	38,0	37,0	40,5	37,6
010001	17,5	16,0	19,0	19,0
000101	12,0	14,6	16,0	14,7
010101	8,5	6,4	5,5	4,9
autres	31,0	39,0	18,0	29,6

Tableau 32. Un échantillon plus « anapestique »

[146] Il considère le schéma 5 comme négligeable.

Ensuite, Dinu signale une très légère préférence pour le type ïambique dans le premier hémistiche, au détriment du type anapestique. Dinu n'indique pas cependant l'existence d'une nette préférence pour le type anapestique sur le second hémistiche. En effet, nous avons vu que la part relative des 001001 était relativement beaucoup plus forte sur le second que sur le premier segment métrique : si l'on ne retient que les formes régulières, la forme 001001 représente 39 % des premiers segments métriques et 45 % des seconds (*cf.* Tableau 30).

Le troisième résultat introduit par Dinu, et que nous confirmons, est que « l'assemblage des unités métriques qui composent le vers n'est soumis qu'aux lois du hasard, la volonté organisatrice du poète n'étant pas décelable à ce niveau » [Dinu, 1993, p. 70].

Les tests effectués sur l'ensemble des vers de Corneille et Racine confirment l'absence de liaison significative entre la structure du premier et celle du second segment métrique.

Les deux résultats suivants apportés par Dinu concernent les couples de vers rimant ensemble. Tout d'abord, Dinu montre que sur les cent premiers couples de vers d'*Andromaque*, la structure du second vers est différente de celle du premier :

> Le premier hémistiche du vers initial de chaque distique a 65 % de chances d'appartenir au type ïambique et seulement 35 % au type anapestique. L'écart n'est pas écrasant, mais il reste sensible et se détache surtout si l'on compare cette situation à celle de l'autre vers du distique qui commence dans une proportion de 56 % par un hémistiche de type anapestique Ces résultats semblent soutenir l'idée d'un parallélisme entre la structure rythmique et le groupage des rimes.
>
> [Dinu, 1993, p. 71].

Avec notre marquage, nous mettons en évidence une légère préférence pour le type anapestique sur le troisième segment métrique, par rapport au premier, et ce au détriment de la catégorie « autres ». L'écart tient essentiellement aux caractéristiques et à la taille réduite de l'échantillon. En effet, nous avons vu précédemment que sur CORRAC, il n'y avait pas de différence significative entre le premier et le second vers (*cf.* Tableau 31). La part des hémistiches de type 001001 y est un peu plus élevée sur le troisième segment métrique (25,1 %) que sur le premier (22,9 %), tandis que sur les 200 premiers vers d'*Andromaque*, on

observe avec le métromètre un écart de 14 points (38 % contre 24 %), comme l'indique le Tableau 33.

200 vers Andromaque	Premier segment métrique	Deuxième segment métrique	Troisième segment métrique	Quatrième segment métrique
001001	24,0	35,0	38,0	48,0
010001	22,0	15,0	13,0	23,0
000101	13,0	17,0	11,0	15,0
010101	7,0	8,0	10,0	3,0
Total	66,0	75,0	72,0	89,0
Autres figures	34,0	25,0	28,0	11,0

Tableau 33. Figures des segments métriques sur les paires de vers

Ensuite, Dinu - sans l'aide de tests statistiques - affirme qu'il y a un certain ordre dans l'organisation des hémistiches dans les couples de vers. Les tests que nous avons effectués sur les liens entre les quatre segments métriques ne permettent pas de mettre en évidence des liens significatifs.

Il était intéressant d'essayer de mettre en évidence un éventuel lien entre la structure rythmique des vers et des hémistiches et la rime. Si, sur un petit échantillon, peu représentatif, avec un certain type de marquage, Dinu a pu en partie montrer une certaine dépendance, la confrontation à de vastes corpus invalide ses hypothèses. Dinu avait pleinement conscience des limites de son échantillon :

> Il est évident que, vu les dimensions modestes de l'échantillon utilisé, les résultats obtenus ne peuvent présenter qu'un intérêt méthodologique.
>
> [Dinu, 1993, p. 73].

Sur CORRAC, il n'y a pas de dépendance significative entre la structure du premier hémistiche et celle du second, ni entre la structure du premier vers du distique et celle du second. Il nous faut donc admettre qu'en tout cas sur les œuvres complètes de Corneille et Racine, la structure de la rime n'est pas amplifiée par une organisation rythmique particulière des deux vers, ni des quatre hémistiches.

En revanche, l'effet de la rime est effectivement renforcé par la préférence du second hémistiche pour les figures les plus régulières et, en particulier, pour

la figure 001001. Le fait d'introduire des règles d'effacement et de réécriture des accents, comme l'a fait Dinu, a pour effet de gommer la préférence du second hémistiche pour les figures les plus régulières.

VII.4.5. Conclusion

Le marquage accentuel met en évidence que le vers est la concaténation de deux segments, aux allures générales similaires : marquage systématique de la dernière position, faible marquage des positions 1 et 5 et marquage intermédiaire des positions 2 à 4. La position finale est d'autant plus éminente que les positions voisines, la position 5 et la position 1 du segment suivant, sont faiblement marquées.

En approchant de plus près les figures des deux segments, il apparaît que le contraste entre l'avant-dernière et la dernière position est encore plus grand sur le second hémistiche que sur le premier, et surtout que la position 3 du second hémistiche est beaucoup plus souvent marquée que la position 3 du premier hémistiche.

L'analyse des séquences « accentuelles » de chaque hémistiche confirme ce résultat. La répartition de ces figures montre tout d'abord qu'un nombre réduit de figures (001001, 010001, 000101 et 010101) couvre une très grande partie des hémistiches, à savoir les deux tiers, mais elle montre aussi que, sur le second hémistiche, la forme 001001 (marquage en position 3) est nettement plus fréquente que sur le premier. Apparaît ici un facteur net de différenciation entre les deux hémistiches.

VII.5. CONCLUSION SUR LA FIGURE DU VERS CLASSIQUE

Nous venons de voir dans ce chapitre comment certaines propriétés phonétiques, lexicales, syntaxiques et prosodiques, du type « être occupé par un *e* muet », « appartenir à un nom », « porter un accent », se réalisaient ou non sur chacune des positions du vers. Nous avons décrit les régularités dans la disposition des différents constituants linguistiques, vérifiant ce qu'annonçait Halle :

> Les énoncés composant un morceau de poésie métrique révèlent des régularités dans la disposition séquentielle de leurs constituants phonétiques, morphologiques et syntaxiques, que l'on ne trouve logiquement pas dans la langue ordinaire.
>
> [Halle, 1972, p. 103].

Le premier résultat est la validation statistique de certaines propriétés connues ou déjà signalées dans la littérature :

- les sixième et douzième positions coïncident toujours avec une fin de mot ;
- les sixième et douzième positions correspondent quasiment toujours à un mot plein (nom, verbe, adjectif, adverbe) ;
- les sixième et douzième positions portent toujours un accent ;
- il n'y a jamais de *e* muet en sixième, ni en douzième position ;
- il n'y a jamais de *e* muet faible en première ni en septième position.

Ces propriétés qui, de vers en vers, sont systématiquement vérifiées donnent un statut très particulier aux positions finales 6 et 12, et dans une moindre mesure aux positions initiales 1 et 7.

Si nous considérons que le domaine de la métrique englobe tous les phénomènes qui se reproduisent à l'identique de vers en vers, la liste précédente délimite, et cela n'a rien d'original, ce qui au niveau du vers et des hémistiches relève du mètre.

Les différents marquages adoptés nous ont permis de mettre à jour d'autres phénomènes qui ne sont pas métriques au sens défini précédemment, parce qu'ils ne sont pas systématiquement attestés, mais qui dessinent des tendances rythmiques dans le vers.

Tout d'abord, l'hémistiche comme segment métrique n'est pas uniquement défini par l'éminence de la dernière position. Nous avons en effet observé pour tous les marquages (de voyelle, de catégorie syntaxique, de fin de mot, d'accent, de nombre de phonèmes) une remarquable similitude de l'allure des courbes du premier et du second hémistiche (à quelques exceptions près). Au regard de *chaque* marquage, l'hémistiche constitue indiscutablement la période et délimite la forme qui se répète.

Ensuite, sur la plupart des marquages, le différentiel est très fort entre la position de fin de segment et ses positions adjacentes. Si l'on se limite aux marquages de mots pleins, de fin de mot, « d'accent » de mot plein, et même de nombre de phonèmes par position, l'écart se creuse très sensiblement entre la sixième et la septième positions et entre la douzième et la première du vers suivant. De même, relativement à la sixième et à la douzième, les cinquième et onzième positions sont très faiblement marquées. L'éminence des positions de fin de segment est donc renforcée par le faible marquage des positions immédiatement adjacentes. Bref, le segment métrique du vers alexandrin a une forme propre et ses contours sont renforcés par les contrastes qui s'établissent avec les segments contigus.

La plupart des marquages étudiés présentent des courbes de répartition avec ces caractéristiques : éminence de la dernière position, symétrie des courbes du premier et du second hémistiches. C'est cette figure commune que l'on peut appeler le rythme de l'alexandrin. Chaque marqueur, qu'il soit vocalique, syntaxique ou prosodique, s'inscrit dans le rythme commun, en présentant sa propre courbe, comme les instruments d'un orchestre symphonique jouent chacun leur partition et pourtant participent de la même symphonie. En cela, le rythme est une structure feuilletée, mais dont chaque « couche » suit les mêmes contrastes. La synchronisation des variations de tous les marquages observés sur la séquence métrique est particulièrement remarquable. On peut penser que cela produit un phénomène de résonance, en réception, entre toutes ces formes ayant la même période. On comprend alors mieux que Lusson et Roubaud aient parfois proposé un unique marquage métrico-rythmique qui combine tous ceux qui « vont ensemble » pour mieux faire ressortir le phénomène du rythme [Lusson et Roubaud, 1981 ; Lusson, 1982].

Enfin, et là nous nous appuyons principalement sur le marquage accentuel, à l'intérieur des hémistiches, un faible nombre de figures métrico-rythmiques (001001, 010001, 000101 et 010101) domine largement. Dans le cadre de la théorie du rythme, les formes 001001 et 010101 sont considérées comme métriques car les séquences peuvent s'analyser comme la concaténation de groupements identiques à un groupement générateur (001 et 01). Il y a de plus une forme d'indépendance entre les structures du premier et du second hémistiche.

L'unité hémistiche est assurée par le marquage systématique de la dernière position de chaque hémistiche, par le contraste entre la fin de segment et le début de segment suivant, et par la similitude des allures des courbes sur les deux hémistiches pour chaque marquage.

Les segments métriques donnent le sentiment d'être autonomes par le parallélisme de leurs figures pour chaque marquage et par l'absence de débordement d'un segment sur l'autre. Peut-on soutenir pour autant, si l'on se place au niveau des positions métriques, que l'unité du vers n'est assurée que par la rime ?

Comme nous l'avons vu en étudiant la question des positions métriques, des interactions entre les deux segments métriques ont lieu au point de jonction : la césure. Tous les contextes ne sont pas autorisés : V#C[147], VC#C, V*e*#V, VC*e*# sont admis, tandis que C*e*#C, V*e*#C, *ent*#V, *ent*#C, *es*#V, *es*#C, V#V ne le sont pas. L'interdiction des césures épiques, lyriques et italiennes (même si celles-ci ont peu fait partie du répertoire métrique du vers français), impose des contraintes fortes dans le lieu de rencontre entre les deux hémistiches : la fin du premier segment et le début du suivant dépendent l'un de l'autre. Ces contraintes d'agencement n'existent pas pour la fin du vers, qui, elle, est soumise à la seule question de la rime.

L'unité du vers est assurée par la rime, par le mode d'agencement des deux hémistiches, autrement dit par les phénomènes de frontières. Elle l'est aussi dans une certaine mesure par la prédominance des figures rythmiques régulières sur le second hémistiche par rapport au premier. Le second hémistiche privilégie nettement les figures les plus métriques du type 001001 et 010101.

Bref, le vers alexandrin classique a bien cette particularité observée par Roubaud d'être à la fois un seul (le vers) et deux segments (les hémistiches) métriques. Cette particularité se voit très nettement dans la forme rythmique telle que nous avons pu la représenter aux niveaux phonétique, syntaxique et accentuel : la répétition d'une ondulation très marquée à la fin, la seconde occurrence étant légèrement plus contrastée que la première.

---

[147] V : voyelle ; # : frontière de mot ; C : consonne.

VII.6. APPENDICE SUR L'EVOLUTION DU VERS

Nous disposons des caractéristiques de l'alexandrin de Corneille et Racine, qui a de fortes chances d'être un bon représentant de l'alexandrin classique. Maintenant que cette référence est établie, nous pouvons voir comment le vers a évolué dans le temps. Nous disposons en effet d'autres corpus, dont les seuls exploitables sont ceux du XVIIIe et du XIXe siècles : *Tancrède* de Voltaire, le premier livre des *Contemplations* et la *Légende des siècles* de Hugo, les *Fleurs du Mal* de Baudelaire, les poésies de Mallarmé et de Rimbaud. Tous ces corpus proviennent de l'InaLF, sauf celui de Rimbaud aimablement prêté par Max Reinert. Ces corpus ont été analysés plus sommairement[148]. Nous ne retiendrons que quatre résultats portant sur le *e* muet, sur le marquage de fin de mot, sur les clitiques en sixième et douzième positions et sur la marque accentuelle[149].

Auteur	Œuvre	Nbre de vers	%	Date de publication
Corneille	*Œuvres théâtrales*	60982	49,47	1629-1674
Racine	*Œuvres théâtrales*	18939	15,36	1664-1691
Voltaire	*Tancrède*	1648	1,34	1761
Hugo	*Les Contemplations (Livre I)*	1631	1,32	1856
Hugo	*La Légende des siècles*	24752	20,08	1859-1883
Baudelaire	*Les Fleurs du mal*	3281	2,66	1861
Rimbaud	*Poésies*	1946	1,58	1870-1871
Mallarmé	*Poésies*	1403	1,14	1898
Total		123274	100,00	

Tableau 34. Répartition des vers selon les auteurs dans « TOT »

Commençons par le *e*. Dans le métromètre, nous avons distingué deux types de *e* muets : le *e* muet fort toujours prononcé en langue et le *e* muet faible d'un polysyllabe qui a une forte propension à l'élision même devant consonne.

[148] Nous renvoyons à l'annexe VI pour la description de ces autres corpus, leur mise en forme, le taux d'erreur avec le métromètre, les nouvelles règles sur la diérèse...

[149] On trouvera dans les annexes VI et VII des résultats complémentaires, en particulier les vers ayant un *e* muet en sixième ou septième positions, les vers avec des mots qui « traversent » l'hémistiche et enfin la liste des vers qui présentent un mot-outil en sixième ou douzième positions.

La répartition des deux types de *e* muets chez les différents auteurs montre que les *e* muets faibles, qui tombent systématiquement en langue, ont tendance à être de plus en plus fréquents au fil du XIXe siècle (sur l'ensemble des positions possibles : P. 2 à P. 5 et P. 7 à P. 11), tandis que les *e* muets forts, toujours prononcés dans la prose quand ils sont graphiquement présents, se font plus rares (sur les positions possibles : P. 1 à P. 5 et P. 7 à P. 11, mais la baisse est très sensible sur les positions 1 et 7) (*cf.* Figure 37). En ce qui concerne le *e* muet, le vers évolue en s'écartant de la diction courante.

Par ailleurs, à partir de Rimbaud, certains *e* muets, rares, apparaissent en des positions interdites. On trouvera dans *La Vieillesse d'Alexandre* [Roubaud, *1978*, 1988, p. 42-43], une liste assez riche de ces différentes violations[150]. D'un point de vue statistique, ces quelques violations passent inaperçues mais elles ébranlent de manière violente l'édifice alexandrin. Il y a dans le passage du zéro au un, de l'absence à la présence, de l'impossibilité à l'existence, un saut qualitativement beaucoup plus saisissant qu'une évolution de dix points entre deux pourcentages.

[150] Le métromètre nous permet évidemment d'identifier ces vers avec *e* muet en sixième position. Mais le travail ayant été déjà fait, nous ne le présentons qu'en annexe (VI).

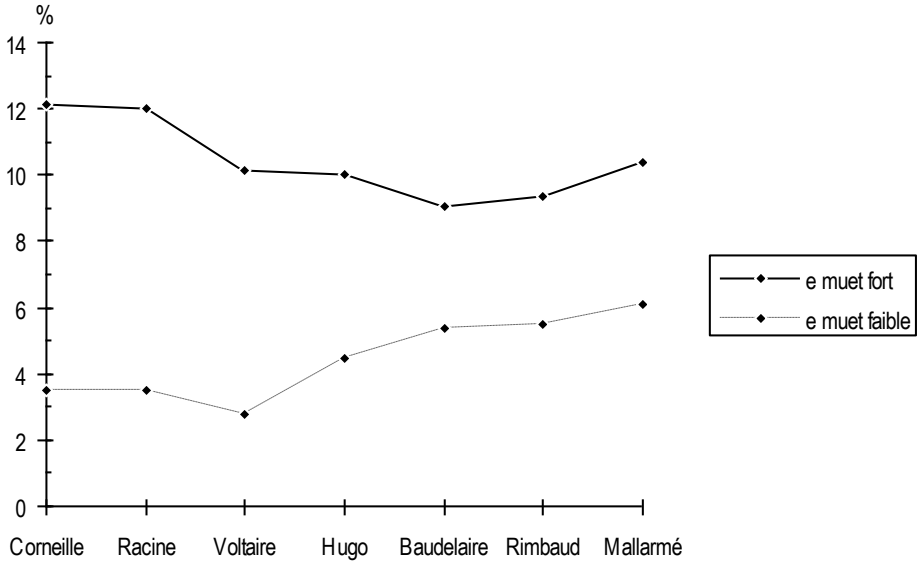


Figure 37. Évolution de la proportion de* e *muets dans le vers

En résumé, pour le *e* muet, deux tendances sont ici mises en évidence. Le vers s'éloigne au cours du XIXe siècle de la diction courante, en introduisant de plus en plus de *e* muets comptés qui ne sont pas prononcés en langue. Le vers s'éloigne aussi de l'alexandrin classique en violant la règle d'interdiction du *e* muet en sixième position. On a donc la conjonction de deux principes de dérive du vers : par rapport à la langue et par rapport au mètre. Cela constitue un premier résultat fort sur l'évolution du vers.

En ce qui concerne la répartition des fins de mots selon les auteurs, le premier résultat est que la dernière position de chaque hémistiche coïncide toujours statistiquement avec une fin de mot[151]. La proportion de frontières de mots varie peu sur les positions 1, 4 et 10. En revanche, les variations les plus fortes concernent les positions 5 et 11. On ne peut cependant tirer des conclusions sur l'évolution du vers : en effet, chez Baudelaire les fins de mots

[151] Sauf quelques rares exceptions, qui transgressent fortement les règles : 2 chez Rimbaud, 5 chez Mallarmé. Nous renvoyons à *La Vieillesse d'Alexandre* [Roubaud, 1978, 1988, p. 42-43] qui répertorie ces cas de mots qui « traversent » l'hémistiche. Le métromètre sait produire cette liste (*cf.* Annexe VI), mais n'apporte rien de neuf dans ce cas.

en positions 5 et 11 sont plus rares que chez Corneille et Racine, tandis que chez Rimbaud elles sont plus fréquentes.

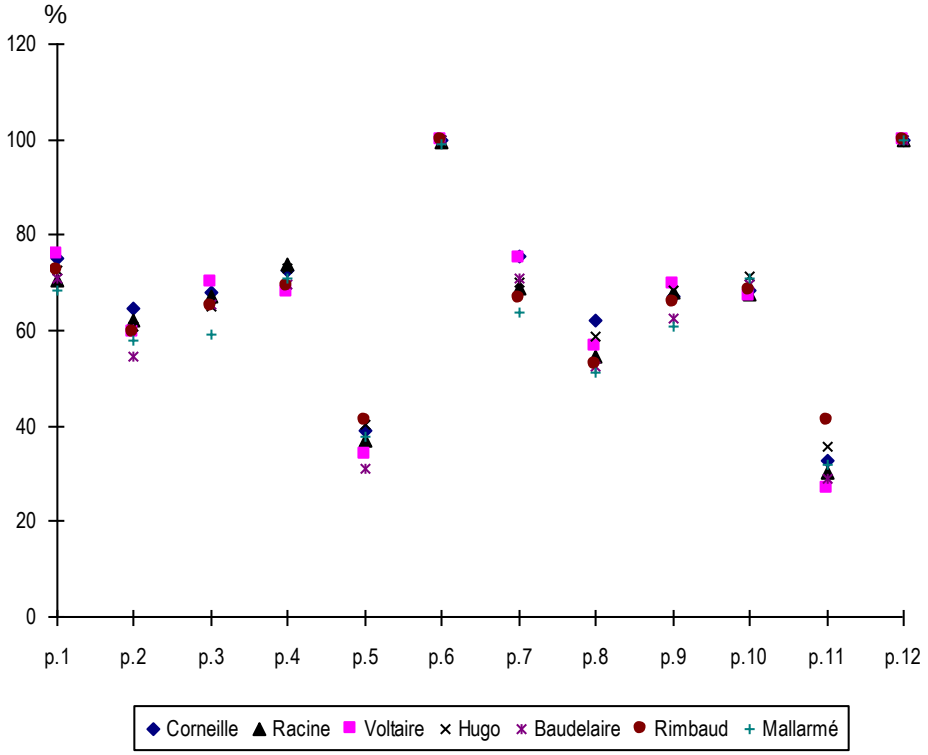


Figure 38. Répartition des fins de mots selon les positions

Passons maintenant aux aspects syntaxiques et examinons comment évolue la proportion de mots-outils en sixième et douzième positions[152]. Chez Corneille et Racine, les différences entre le premier et le second hémistiche sont minimes : en gros il n'y a quasiment pas d'emploi de mots-outils en sixième ni en douzième positions. En revanche, dès Hugo, l'écart se creuse entre les deux hémistiches et s'accentue encore chez Rimbaud, puis chez Mallarmé. La proportion de mots-outils en sixième position reste minoritaire

[152] À partir des listes (produites par le métromètre) des vers ayant un mot-outil en position 6 ou 12, nous avons retenu les vers contenant des mots-outils clitiques, c'est-à-dire ceux qui sont rattachés à un mot plein situé dans le segment métrique suivant (créant par là des distorsions entre mètre et syntaxe). La liste complète des vers présentant des mots-outils en sixième ou douzième position figure dans l'annexe VII.

(moins de 10 %), mais augmente vite d'un auteur à l'autre. La douzième position connaît une évolution plus lente (*cf.* Figure 39).

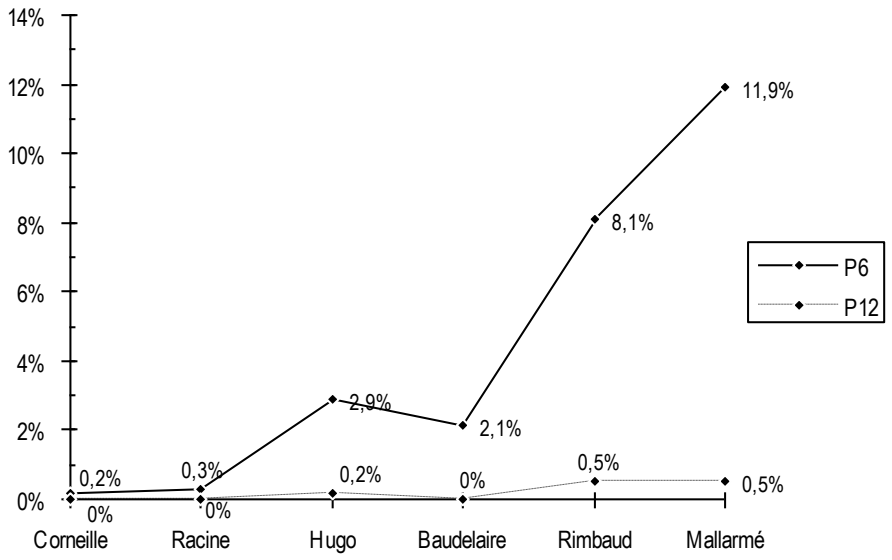


Figure 39. Proportion de mots-outils en sixième et douzième positions

Les mots-outils en fin de segment métrique, tels que nous les avons comptés, sont un indice de l'évolution des rapports entre syntaxe et mètre : ils signalent toujours des cas de rejet ou d'enjambement[153]. Tout se passe comme si dans le vers classique la métrique dominait la syntaxe en lui imposant des distorsions rares dans la langue courante, comme la métaposition, alors qu'à partir du XIXe siècle la syntaxe regagne du terrain par rapport au mètre en s'affranchissant partiellement de la contrainte de la césure, mais en continuant cependant à respecter celle du vers.

Nous voyons avec clarté que le « vieillissement » de l'alexandrin se manifeste principalement à la fin du premier segment métrique et que l'unité du vers est préservée. Elle le sera encore longtemps : Roubaud [*1978*, 1988,

[153] Les mots-outils en positions 6 et 12 ne permettent pas de repérer tous les cas de distorsion entre mètre et syntaxe. Il faudrait aussi étudier les cas où des syntagmes sont traversés par des frontières métriques (entre un adjectif et le verbe qui le suit, entre un verbe et un adverbe...).

p. 126] montre comment dans le vers libre classique la ligne-vers continue de coïncider avec la phrase, autrement dit, le vers libre « respecte la règle de concordance des frontières métriques et syntaxiques » [p. 127].

Pour finir sur la question de l'évolution du vers, et là encore il ne s'agit que de pistes à approfondir, voyons comment se répartissent selon les auteurs les marques accentuelles sur les douze positions. Premier résultat remarquable, les variations entre auteurs sont nettement plus sensibles sur le second hémistiche que sur le premier, comme si les différences de style cherchaient à s'affirmer sur le segment métrique dominant du vers, le second hémistiche qui porte la rime et qui est globalement plus régulier...

La Figure 40 montre que Racine et Baudelaire sont proches dans leur manière de traiter le vers. Tous deux privilégient les positions 3 et surtout 9, ce que confirme la proportion élevée d'hémistiches de type 001001, surtout sur le second hémistiche. A l'inverse, Hugo préfère aux positions multiples de trois les positions paires et en particulier la position 8. Les figures d'hémistiches avec un marquage en positions paires sont sureprésentées. Mais nous ne voyons pas apparaître de manière significative les deux formes d'hémistiches qui correspondraient à l'alexandrin ternaire (H1_000101 et H2_010001) : si le type H2_010001 semble caractéristique de Hugo, ce n'est guère le cas de H1_000101.

En ce qui concerne la marque accentuelle, la place de Rimbaud est très spécifique : il marque plus fréquemment que les autres les positions qui ont les niveaux de marquage les plus faibles : P. 11, P. 5, P. 7 et P. 1. Mallarmé, pour finir, ne se distingue des autres que par le marquage de la position 10.

Marque accentuelle	P. 1	P. 2	P. 3	P. 4	P. 5	P. 6	P. 7	P. 8	P. 9	P. 10	P. 11	P. 12
Racine	13,7	34,9	40,8	30,9	7,1	97,0	9,0	29,1	48,8	21,1	3,4	98,8
Baudelaire	12,8	32,3	43,7	30,3	5,1	97,3	10,4	31,5	46,0	24,2	4,8	99,5
Mallarmé	15,0	34,1	39,1	33,8	6,6	91,1	9,6	34,2	42,8	29,6	5,6	98,8
Corneille	12,2	31,2	39,9	32,4	8,7	96,5	8,8	31,5	42,4	27,6	5,0	98,9
Rimbaud	15,7	35,3	38,7	35,4	12,3	95,6	12,7	37,8	42,1	29,2	9,4	99,0
Hugo	16,7	36,9	40,9	36,1	9,2	96,0	11,4	41,3	40,9	29,5	6,8	99,0
Ensemble	13,6	33,3	40,7	32,8	8,5	96,5	9,6	33,2	43,9	26,7	5,1	99,0

Clef de lecture : la case grisée indique que le marquage est surepresenté pour telle position chez tel auteur par rapport à la moyenne, la case hachurée indique qu'il est sous-représenté.

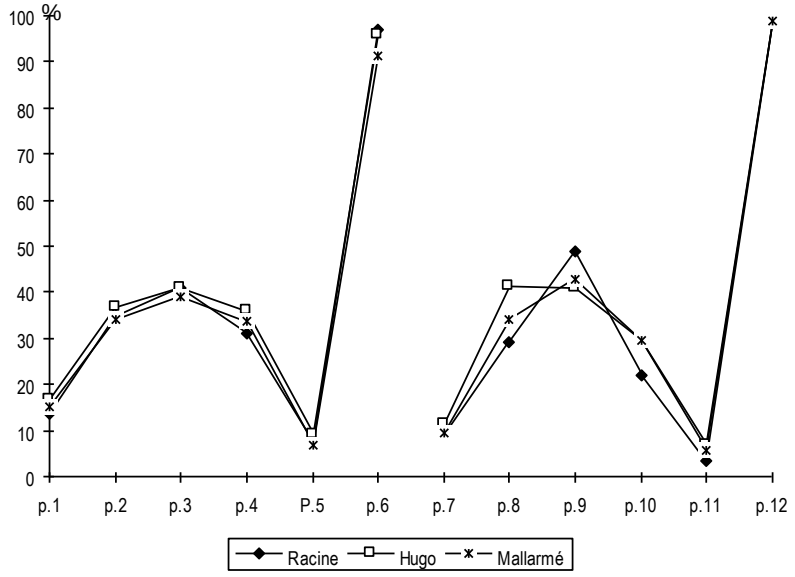


Figure 40. Marquages accentuels les plus différenciés

Le taux d'hémistiches réguliers est un indicateur stylistique supplémentaire (*cf.* Tableau 35). Il est particulièrement élevé chez Baudelaire et Mallarmé, tout comme chez Racine, et particulièrement bas chez Corneille, Hugo et Rimbaud. On peut donc difficilement imputer à des effets de genre ces particularités, mais bien à des styles d'écriture.

Premier SM	Corneille	Racine	Hugo	Baudelaire	Rimbaud	Mallarmé
001001	23,4	25,8	21,7	29,2	22,2	26,1
000101	15,0	13,5	11,7	13,1	13,9	14,3
010001	15,3	18,2	14,4	17,7	13,7	15,4
010101	6,3	7,0	10,4	6,9	7,7	7,9
Autres	40,1	35,5	41,8	33,2	42,6	36,3
Second SM	**Corneille**	**Racine**	**Hugo**	**Baudelaire**	**Rimbaud**	**Mallarmé**
001001	29,8	37,9	27,2	36,7	28,8	33,0
000101	15,3	12,8	11,5	13,8	11,5	16,4
010001	18,7	19,8	22,7	21,0	21,0	20,2
010101	5,3	3,7	9,6	4,6	8,3	6,7
Autres	30,9	25,8	29,1	23,9	30,4	23,8
Total	56479	18438	22759	2334	1370	942

Tableau 35. Répartition des types d'hémistiches selon les auteurs

Nous voyons dans ce genre d'analyse diachronique se mêler des effets liés à l'évolution du vers, comme la montée du taux de mots-outils en position 6, ou l'augmentation du nombre de positions occupées par un *e* muet faible et des effets stylistiques, comme la préférence de Hugo pour la position 8, celle de Mallarmé pour la 10 et celle de Baudelaire et Racine pour les hémistiches 001001. Ces résultats sont très partiels, et mériteraient d'être explorés de manière plus approfondie.

VIII. LA CONTRIBUTION DU RYTHME

A LA FORME POETIQUE

Dans la suite logique de ce que nous avons présenté jusqu'ici, nous voudrions maintenant chercher à montrer dans quelle mesure les unités sémantiques contribuent au rythme des vers, au même titre que les phonèmes, la syntaxe, le schéma accentuel. Il nous faudrait pour cela mettre en place des marquages sémantiques du vocabulaire qui permettent de voir comment certains mots se répartissent de manière non aléatoire sur les différentes positions métriques du vers. Bien entendu, appréhender le sens à travers le seul vocabulaire est assez insatisfaisant, mais représente cependant une première étape.

Quelles pistes pourrions-nous retenir pour ce type d'analyse, sachant que la prise en compte de l'ensemble du vocabulaire d'un texte est difficile ?

Deux possibilités simples peuvent être mises en œuvre. La première revient à établir le vocabulaire spécifique d'un texte par rapport à un cadre de référence, par exemple d'une tragédie face à l'ensemble des tragédies d'un auteur et à examiner ensuite comment ces termes spécifiques occupent les positions du vers. Nous verrions sans doute chez Corneille et Racine que les mots les plus significatifs tendent à occuper de manière notable les positions finales des segments métriques, surtout la position de la rime. La seconde possibilité consiste à définir une liste de mots *a priori*, en fonction des hypothèses que l'on cherche à tester et à examiner alors leur répartition métrique. Nous pourrions, par exemple, examiner le comportement métrique des mots qui dans la tragédie donnent souvent lieu à des syllepses de sens comme *cœur, flamme, couronne...* : le sens concret et le sens figuré de ces mots

se trouvent simultanément actualisés dans le contexte. Plus finement, la présomption d'une isotopie pourrait nous conduire à identifier les léxèmes comportant un même sème isotopant (le sème étant une unité de contenu inférieure au léxème) et à examiner la répartition de ces léxèmes[154]. Mais la notion de mot en tant que forme graphique est loin d'être satisfaisante pour une approche sémantique. Il faudrait tout d'abord effectuer une décomposition morphématique des mots, puis une décomposition sémique, si l'on considère, comme la sémantique componentielle le fait, que le sens peut être décomposé en unités sémantiques minimales, les sèmes.

La décomposition sémique est loin d'être évidente. En effet, quelle liste de sèmes retenir ? Comment les associer à des mots du discours ? Comment le faire d'une manière systématique ?

Ce travail d'analyse de la contribution du sens au rythme des vers n'a pas été effectué. Cela devrait être envisagé dans les versions ultérieures de l'outil.

Nous aurions dû montrer comment le sens contribue, au niveau microstructural, à la structure rythmique du vers, mais nous avons plutôt cherché à montrer qu'il pouvait exister des corrélations entre rythme et sens, et que le rythme pouvait jouer le rôle de marqueur stylistique. En inversant le point de vue, nous chercherons ici à mettre en évidence que le rythme contribue, au niveau macro-structural, et non plus seulement au niveau du vers, à l'élaboration d'une forme littéraire spécifique. Nous allons examiner les liens éventuels existant entre la figure du vers et d'autres aspects formels, ceci en vue de comprendre la contribution des aspects formels à la cohésion d'un texte.

Rythme et sens

Les liens entre structure rythmique ou métrique et sens ont déjà été étudiés. Donnons-en trois illustrations différentes et complémentaires.

Dans *La Poétique du vers*, ouvrage consacré à *Aire Nuestro* de Jorge Guillén, P. Bonhomme vise à « établir les liens entre le champ métrique et le

---

[154] Nous renvoyons à *Sémantique interprétative* de François Rastier [1987] pour un examen théorique du concept d'isotopie et à *Sens et textualité* du même auteur [1989] pour des exemples d'applications de la théorie à des textes littéraires.

champ sémantique », en partant de l'hypothèse qu'un « poème construit de tel mètre et ordonné de telle façon doit nécessairement signifier de par cette construction » [1995, p. 9]. Cette approche lui permet de proposer une lecture des signifiés des poèmes qui s'appuie sur les signifiants métriques.

> Il s'agira donc de montrer que le discours guillénien met sans cesse à nu, par-delà l'évidente et paraphrastique déclaration d'une complétude désirée que les signifiés des poèmes ne cessent de ressasser, l'histoire des signifiants qui lui ont permis d'accéder à l'architecture métrique accomplie.
>
> En d'autres termes, le grand Poème guillénien raconte une histoire qui est aussi celle des mètres qui l'ont informé.
>
> [Bonhomme, 1995].

Dans cet ouvrage, ce sont les liens entre les signifiants *métriques*, principalement les types de mètres, et ce que disent les poèmes qui sont analysés.

La recherche de relations peut également porter sur des phénomènes rythmiques internes au vers. Par exemple, Marina Tarlinskaja [1987], dans un article intitulé « Rythm and Meaning : "Rythmical Figures" in English Iambic Pentameter, Their Grammar and Their Links with Semantics », s'intéresse à la question des relations entre sens et rythme. En se limitant à un cas précis, les transgressions de l'ordre métrique, et en s'appuyant sur un corpus de taille considérable (18 000 pentamètres ïambiques tirés principalement de Shakespeare - poésie et théâtre -, mais aussi de Pope, Shelley et Arnold), elle cherche à mettre en évidence des relations entre des figures rythmiques qui transgressent l'ordre ïambique (principalement des inversions de pieds) et, soit des connotations émotionnelles, soit des catégories sémantiques. Elle se limite aux cas de transgression métrique car ils constituent pour elle un lieu idéal pour la mise en valeur d'un mot quand la position accentuée d'un mot en langue (respectivement non accentuée) se trouve sur une position non marquée (respectivement marquée) du vers. Elle montre par exemple que les adjectifs à connotation émotionnelle sont plus fréquents là où il y a transgression de l'ordre métrique qu'ailleurs (inversion des positions accentuées / inaccentuées).

Roubaud et Lusson proposent certaines « lectures rythmiques » de poèmes (y compris de poèmes mis en musique) dans lesquelles ils ont introduit un certain nombre de marquages sémantiques. La lecture rythmique, telle que l'entendent les auteurs, résulte de la combinaison d'un certain nombre de

« marquages, qui, combinés, dessinent la figure des vers et celle des strophes. Certains marquages textuels (syntaxe, ponctuation, marque accentuelle) et musicaux sont systématiquement utilisés. D'autres sont définis en fonction des particularités des textes. Ils ajoutent en général des marquages propres aux caractéristiques des textes étudiés. Prenons deux exemples de lecture rythmique : un sonnet de Jodelle et une canso de Rudel. A propos du sonnet de Jodelle « Des astres, des forests et d'Achéron l'honneur... », composé pour Claude-Catherine, maréchale de Retz, la lecture rythmique [Lusson et Roubaud, 1981] permet entre autre de montrer que dans le vers, dont le modèle s'apparente à un trimètre, avec les positions impaires peu marquées, la position 9 est au contraire fortement marquée. Or, en cette même position 9, du neuvième vers, se trouve le mot-clef *rets* (*Retz*), référence onomastique, lourdement chargée d'un point de vue sémantique. La construction d'une figure rythmique spécifique du vers est ici mise au service du sens.

Ta beauté par ses rais, par son rets, par la craincte

Jodelle, « Des astres, des forests, et d'Achéron l'honneur… »

Prenons comme autre exemple la lecture rythmique que Le Vot, Lusson et Roubaud [1979] ont proposé de la chanson de « l'amour de loin » de Jauffré Rudel. L'analyse rythmique propose à la fois une lecture musicale et textuelle de la canso et montre l'articulation de ces deux composantes. Concernant la lecture textuelle, un marquage sémantique est introduit, qui marque les positions du vers dans lesquelles apparaît un mot appartenant au registre poétique du grand chant courtois. Les auteurs montrent, en considérant le fait que les vers contiennent ou non un mot du grand chant courtois, que ce registre poétique renforce métriquement la structure des strophes composées de 7 vers organisés par la rime en deux groupements *ababccd*.

A notre tour, nous allons rechercher des liens entre les formes du discours et les caractéristiques métrico-rythmiques et rimiques du vers. Autrement dit, nous allons nous demander quelles relations entretiennent les « signifiants métriques », pour reprendre l'expression proposée par P. Bonhomme [1995], avec les univers de discours. Nous avons tenté sur notre corpus de préciser quelles pouvaient être les relations entre le rythme des vers et les univers de discours.

Signification de la variation métrique

Étudier la variation métrique des vers de Corneille et Racine, essentiellement des alexandrins, ne présente ici qu'un intérêt limité, contrairement au travail fait par P. Bonhomme [1995] sur un recueil qui mobilise tous les mètres de l'histoire du vers espagnol. Les passages qui ne sont pas en alexandrins à rime plates, dans les pièces largement dominées par ce vers, correspondent à des stances ou à du discours rapporté.

Tout un ensemble d'éléments contribuent à constituer la singularité des stances dans les pièces. Elles définissent, puisque les frontières des scènes sont définies par les entrées et sorties, une scène isolée où l'unique personnage se livre à un monologue insolite. Comme l'écrit Le Tellier :

> Je pense que personne ne s'étonne de la présence de monologues au théâtre, alors qu'il n'y a tout de même que les dingues qui parlent tout seuls.
>
> [Le Tellier, 1997, p. 12].

Ces stances sont très homogènes par leur contenu, puisqu'elles expriment le déchirement intérieur du personnage. A ce monologue est associé une triple rupture métrique : les stances introduisent des strophes, une variation dans la disposition des rimes, et enfin, le plus souvent, des mètres autres que l'alexandrin. Le bouleversement de la régularité et de la platitude fait écho à celui de l'âme, d'une manière certes très convenue, puisque la régularité est parfaite dans l'organisation des strophes. On est loin des silvas de la poésie espagnole. Ainsi les stances de Don Rodrigue sont constituées de 6 strophes de 10 vers organisées comme suit :

	mètre	rime
Percé jusques au fond du cœur	8	a
D'une atteinte imprévue aussi bien que mortelle,	12	b
Misérable vengeur d'une juste querelle,	12	b
Et malheureux objet d'une injuste rigueur,	12	a
Je demeure immobile, et mon âme abattue	12	c
Cède au coup qui me tue.	6	c
Si près de voir mon feu récompensé,	10	d
Ô Dieu, l'étrange peine !	6	e
En cet affront mon père est l'offensé,	10	d
Et l'offenseur le père de Chimène !	10	e

Corneille, *Le Cid*, vers 291-300.

On notera que dans chacune des 6 strophes les mots-rimes *peine* et *Chimène* sont repris à l'identique dans les vers 8 et 10.

Comme l'écrivait Regnault [1996], « les stances se prêtent fort bien à l'exposition du nœud entre l'amour et le politique, puisqu'elles en donnent une version inquiète et irrésolue ». Mais Corneille y renonce petit à petit et le principe n'est repris qu'une fois par Racine dans sa première tragédie, *La Thébaïde*. Ceci conduit Regnault à poser que les vers des stances paraissent moins naturels et vraisemblables que l'alexandrin, qui devient ainsi « *le vers naturel du discours théâtral* » [Regnault, 1996, p. 188].

La variation dans la disposition des rimes et dans la longueur des vers apparaît également quand le personnage lit un billet, une lettre ou rapporte des paroles prophétiques, ou dramatiques.

« Une main qui nous fut bien chère	8	a
Venge ainsi le refus d'un coup trop inhumain.	12	b
Régnez ; et surtout, mon cher frère,	8	a
Gardez-vous de la même main.	8	b
C'est... » La Parque à ce mot lui coupe la parole ;	12	c
Sa lumière s'éteint, et son âme s'envole ;	12	c

Corneille, *Rodogune*, vers 1643-1648.

« Une main qui nous fut bien chère ! »	8	a
Madame, est-ce la vôtre, ou celle de ma mère ?	12	a

Corneille, *Rodogune*, vers 1663-1664.

« Thébains, pour n'avoir plus de guerres,	8	a
Il faut, par un ordre fatal,	8	b
Que le dernier du sang royal	8	b
Par son trépas ensanglante vos terres. »	10	a

Racine, *La Thébaïde*, vers 393-396.

Quelle fut sa réponse ! et quel devins-je, Arcas,	12	a
Quand j'entendis ces mots prononcés par Calchas !	12	a
« Vous armez contre Troie une puissance vaine,	12	b
Si dans un sacrifice auguste et solennel	12	c
Une fille du sang d'Hélène	8	b
De Diane en ces lieux n'ensanglante l'autel.	12	c
Pour obtenir les vents que le ciel vous dénie,	12	d
Sacrifiez Iphigénie. »	8	d

Racine, *Iphigénie*, vers 55-62.

La dimension tragique des paroles prononcées transparaît sans doute mieux sur scène grâce à la rupture métrique. La rupture permet d'isoler, de mettre en relief, l'horreur de l'oracle.

Voici un exemple de lettre :

« Après tant d'injustes détours,	8	a
Faut-il qu'à feindre encor votre amour me convie ?	12	b
Mais je veux bien prendre soin d'une vie	10	b
Dont vous jurez que dépendent vos jours.	10	a
Je verrai la sultane ; et par ma complaisance,	12	c
Par de nouveaux serments de ma reconnaissance,	12	c
J'apaiserai, si je puis, son courroux.	10	d
N'exigez rien de plus. Ni la mort, ni vous-même	12	e
Ne me ferez jamais prononcer que je l'aime,	12	e
Puisque jamais je n'aimerai que vous. »	10	d

Racine, *Bajazet*, vers 1135-1144.

Cette lettre rapportée s'insère mal dans le corps du texte, au point où, quand elle est reprise, un vers ne trouve pas d'écho à sa rime.

Il peut l'avoir écrit sans m'avoir offensée.	12	a
Il peut même... lisons, et voyons sa pensée :	12	a
... « Ni la mort, ni vous-même	6	b
Ne me ferez jamais prononcer que je l'aime,	12	b
Puisque jamais je n'aimerai que vous. »	10	c
Ah ! de la trahison me voilà donc instruite !	12	d
Je reconnais l'appas dont ils m'avaient séduite.	12	d

Racine, *Bajazet*, vers 1265-1270.

Quand il s'agit simplement de paroles rapportées, elles sont en alexandrins à rimes plates. Pour qu'il y ait variation métrique et modification de la disposition des rimes, il faut que les paroles relèvent du caractère tragique de l'oracle, ou du caractère plus solennel de l'écrit.

La variation métrique signale un changement de registre : passage du dialogue au monologue ou introduction de paroles hors-scène, mais ces changements ne s'accompagnent pas toujours d'une variation métrique : bien des monologues et des texte ou discours rapportés sont en alexandrins[155].

[155] Il aurait été également intéressant comme l'ont fait Evrard et Purnelle [1998] d'étudier la coïncidence entre fin de paires de vers rimant ensemble et fin de tour de parole. Les auteurs montrent en particulier que Corneille comme Racine tendent à faire coïncider les deux, et qu'inversement, Hugo évite fréquemment ce type de situations.

Structure rythmique et univers lexico-sémantiques

Les variations métriques étant dans notre contexte marginales et, par conséquent, de signification limitée, nous pouvons en revanche essayer d'identifier les univers de discours dans les pièces de notre corpus et mettre éventuellement à jour l'existence de liens entre ces univers et le rythme des alexandrins. Autrement dit, la structure rythmique des vers est-elle ou non indépendante de ce qui est dit et de la configuration scénique (personnages présents sur scène, situation de la scène dans la pièce...) ?

Nous cherchons donc à voir s'il existe une mobilisation des effets rythmiques du vers pour renforcer les effets de sens. Pour cela, nous allons mobiliser d'une part l'analyse rythmique, d'autre part l'analyse lexicale. Mais qu'entend-on par analyse rythmique et analyse lexicale ?

Le métromètre (*cf.* chapitres IV à VI) produit pour chaque vers un ensemble de descripteurs rythmiques, en particulier le marquage accentuel[156], ce qui permet de construire aisément la figure accentuelle de chaque hémistiche.

Par exemple, le vers suivant présente deux figures d'hémistiches : H1_000101 et H2_001001 :

[156] Nous avons choisi d'indiquer la présence d'un accent sur la dernière position des mots pleins (noms, verbes, adjectifs et adverbes), ou sur l'avant-dernière si la dernière est occupée par un *e* muet. Nous avons détaillé dans les chapitres V et VI les limites de ce marquage.

À cet effort pour vous qui pourrait le contraindre,

Corneille, *Tite et Bérénice*, vers 39

sm	a	s ɛ	tɩ e	f ɔ r	p u r	v u	k i	p u	r ɛ	l ə	k ɔ̃	t r ɛ̃ d r
ma	0	0	0	1	0	1	0	0	1	0	0	1
hém	H1_000101						H2_001001					

Clef de lecture : sm : syllabe métrique ; ma : marquage accentuel ; hém : hémistiche. La première position est occupée par la syllabe a, qui est non accentuée. La figure accentuelle du premier hémistiche est H1_000101.

Il existe plus de 60 figures d'hémistiches[157] qui auraient pu être regroupées de diverses manières. Nous avons adopté, après quelques tests comparatifs, un regroupement en trois catégories : H_001001, H_010101 (qui regroupe 000101, 010001 et 010101) et H_autres (qui regroupe toutes les autres figures). Chaque hémistiche peut donc être caractérisé par l'une de ces trois modalités. C'est ce que nous appellerons ici la forme rythmique abstraite de l'hémistiche.

L'analyse lexicale, également désignée sous les termes d'analyse statistique des données textuelles ou de statistique textuelle [Lebart et Salem, 1988 et 1994], est un ensemble de méthodes qui permettent d'appliquer aux textes les méthodes statistiques de calculs de spécificités, d'analyse factorielle et de classification. L'annexe VIII est consacrée à une rapide présentation de l'origine de ces méthodes, en particulier aux rôles de Charles Muller et de Jean-Paul Benzécri, de leurs principes et des logiciels que nous avons utilisés (Alceste et Hyperbase, conçus et développés par Max Reinert et Étienne Brunet)[158]. Nous utiliserons ces méthodes principalement pour identifier des univers lexico-sémantiques dans nos corpus.

L'objectif est simple : attribuer à chaque vers un marquage sémantique. Pour cela, on procède en deux temps : on détermine les univers lexicaux pertinents dans le corpus, et l'on attribue à chaque vers l'univers qui lui correspond le mieux en fonction de son contenu lexical.

---

[157] Il existe théoriquement 64 combinaisons possibles de 0 et de 1 sur six positions : avec le métromètre 61 des combinaisons sont attestées sur notre corpus.

[158] Pour de plus amples développements, nous renvoyons à Muller [1967 et 1977], à Benzécri [1981], à Lebart et Salem [1994], à Reinert [1993] et à Brunet [1989].

Décrivons succinctement le principe du logiciel Alceste [Reinert, 1983, 1987 et 1993] que nous avons utilisé : les pièces sont découpées en unités de contexte, constituées de séquences de vers (par exemple des suites d'une dizaine de vers). Chaque unité de contexte est alors considérée comme un « sac de mots » [Lahlou, 1998]. L'objectif est, en partant de l'examen de tous ces sacs de mots, de construire une typologie des unités de contexte : seront classées ensemble les unités de contexte qui emploient le même type de vocabulaire. Il ne s'agit pas d'avoir une typologie *a priori* par rapport à laquelle on classerait chacune des unités de contexte, mais de construire les classes et de classer les unités dans un même mouvement. Concrètement, l'ensemble des unités de contexte est séparé en deux, de telle sorte que les unités de contexte de chaque groupe soient le plus homogènes possible en termes de vocabulaire utilisé et que les différences soient maximales entre les deux groupes. Le processus est itératif et en fin de calcul, on se trouve face à une typologie des unités de contexte construite sur la base des cooccurrences lexicales. Le nombre final de classes de la typologie résulte d'un calcul de stabilité.

Chaque classe est décrite par son vocabulaire spécifique, par les unités de contexte les plus typiques (séquences de quelques vers) et par les catégories syntaxiques et pragmatiques les plus spécifiques, éléments qui permettent de construire une interprétation des contenus sémantiques. Tous ces éléments nous permettent en particulier de nommer les classes.

Les classes une fois nommées constituent les univers lexico-sémantiques de référence, avec lesquels on va marquer les vers. La procédure est triviale : chaque vers est marqué par la classe à laquelle il appartient (puisque les classes ne sont rien d'autre que des ensembles de vers, qui forment une partition du corpus). On peut ensuite procéder à toutes les analyses classiques entre les variables « classe lexico-sémantique » et le rythme, soit au niveau des vers, soit au niveau des ensemble de vers que sont les classes, les pièces, les auteurs, les genres…

La première section de ce chapitre sera consacrée au corpus complet. Nous montrerons l'existence de corrélations significatives entre univers de discours et rythme, qui seront interprétées comme le résultat d'effets de genre et

d'auteur. Cela nous amènera à montrer que le rythme contribue au même titre que d'autres indices stylistiques à caractériser genres et auteurs. La question des genres au théâtre ayant largement été traitée par la critique, aucun résultat ne paraîtra surprenant. Cette étude permettra de valider, sur un champ bien balisé, la pertinence des outils et de leur utilisation.

Dans la seconde section du chapitre, nous nous concentrerons sur une partie du corpus particulièrement homogène : les tragédies centrales de Racine, d'*Andromaque* à *Phèdre*. Là, nous montrerons l'existence de liens, ténus mais significatifs, entre des fragments de pièces et des rythmes spécifiques.

VIII.1. UNIVERS DE DISCOURS ET RYTHME : RETOUR SUR LES NOTIONS D'AUTEUR ET DE GENRE

À partir de l'ensemble des pièces de Corneille et Racine, cherchons à identifier des univers lexicaux spécifiques et leurs liens éventuels avec des structures rythmiques. Cela nous conduira à montrer la prégnance des effets de genre et d'auteur. Nous montrerons alors comment un faisceau de faits linguistiques et rythmiques contribue à définir le style d'un genre et d'un auteur.

VIII.1.1. Univers lexico-sémantiques et rythme

Nous cherchons à identifier des liens éventuels entre ce qui est dit et le rythme des vers. Pour cela, nous allons considérer l'ensemble de notre corpus, identifier les univers lexico-sémantiques qui le composent et examiner si ces univers sont caractérisés par des types de rythmes spécifiques. Pour l'analyse, nous définissons des unités de contexte élémentaire (UCE) d'environ 10 vers. Ce sont ces unités qui vont être classées en fonction de leur contenu lexical. Chaque classe sera considérée comme un univers lexico-sémantique qui regroupera des unités de contexte partageant une partie de leur vocabulaire. Chaque classe sera très distincte des autres en termes de vocabulaire.

Chaque classe de contextes est définie par son vocabulaire spécifique. À partir de ce vocabulaire, nous pouvons identifier des isotopies en classant ensemble les léxèmes contenant un sème défini. Par exemple, dans le vocabulaire spécifique de la classe suivante, on peut identifier au moins deux isotopies, construites autour des sèmes /pouvoir/ et /mariage/ :

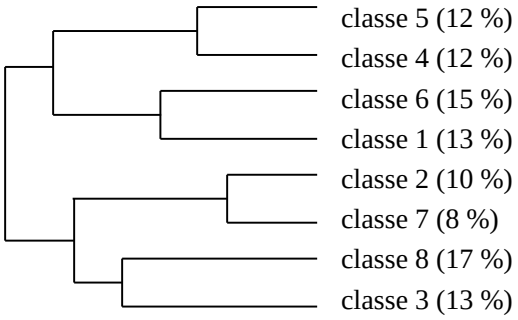
grand, haut, maître, souverain, choix, couronne, empereur, empire, emploi, époux, état, grandeur, ordre, pouvoir, puissance, rang, refus, reine, roi, suffrage, trône, choisir, donner, épouser, estimer, mettre, sceptre, sénat, autorité, politique, droit, illustre, suprême, comte, général, loi, respect, gouverner, proposer, déférent, âge, ambassade, cour, don, éclat, hymen, impératrice, obéissance, offre, règle, plaire, diadème, ambition, civile, différent, prudent, sujet, gendre, art, avis, exil, incertitude, mari, marque, monarchie, preuve, union, affermir, borner, commander, disposer, hasarder.

Les mots sont classés en fonction de leur parenté sémantique et l'on essaie ensuite de comprendre l'articulation de ces différentes isotopies au sein d'une classe pour l'interpréter [Lahlou, 1995], puis l'articulation de ces classes dans les pièces. Nous renvoyons pour une démarche similaire à une analyse de *La Route des Flandres* par Pascal Mougin [1995], dans laquelle l'auteur conclut que la démarche adoptée par le logiciel Alceste peut être considérée comme une mise en œuvre de la procédure d'extraction de Greimas proposée dans *Sémantique Structurale* [1986].

L'analyse ne porte que sur les mots pleins, tels qu'ils sont définis dans le logiciel (noms, verbes, adjectifs et certains adverbes). Nous avons considéré comme mots illustratifs les noms de personnages, les noms de lieux et de peuples et obtenu huit classes : ces mots n'interviennent pas dans la construction de la typologie des unités de contexte, mais interviennent dans la description du contenu des classes obtenues [159].

[159] Nous avons retenu une analyse standard qui croise deux classifications fondées sur deux découpages différents du corpus. Ce croisement permet d'aboutir à un taux de classement correct de 70 % des unités, ce qui est assez satisfaisant.

Le processus de classification descendante utilisé permet de représenter sous forme d'arbre la classification, comme dans la figure ci-dessous :



Guide de lecture : le logiciel Alceste utilisé fait une classification descendante hiérarchique : il part de l'ensemble des unités de contexte (à gauche du graphique) qu'il sépare en deux, il considère ensuite la plus importante des sous-classes restante et la sépare en deux, ainsi de suite. Le nombre de classes final résulte d'un arbitrage entre le choix de l'utilisateur (qui fixe un nombre de classes maximum) et les calculs de stabilité que fait le logiciel. La numérotation des classes est arbitraire.

Figure 41. Arbre de la classification

La classification ici présentée est stable : différents tests menés avec des paramètres différents ont conduit à des résultats proches. Néanmoins, la méthode est loin d'être parfaite et l'on ne saurait prendre une typologie comme un classement définitif.

La classification distingue en premier lieu deux grands groupes. Dans le premier (classes 5, 4, 6 et 1), les catégories de mots les plus spécifiques sont : les verbes modaux (*devoir*, *pouvoir*, *croire*,...) et, comme mots-outils, les marqueurs de modalisation et d'intensité. Inversement, le second groupe (2, 7, 8 et 3) est caractérisé par une prédominance des noms, noms propres et adjectifs, par des déictiques spatiaux et par des interjections. Si l'on examine les mots les plus spécifiques, le sème /amour/ caractérise une grande partie du vocabulaire spécifique du premier groupe, tandis que le sème /mort/ est largement représenté dans le second.

Examinons à présent la répartition des types d'hémistiches dans ces deux groupes. Le tableau suivant montre que dans les séquences de vers classées dans le premier groupe (/amour/) les hémistiches H_autres sont plus fréquents et que, d'autre part, les formes H_001001 sont relativement moins présentes que dans le second groupe (25 % contre 30 %). Inversement, la classe /mort/ se caractérise par une plus forte présence d'hémistiches H_010101 et H_001001. Nous interprétons ces deux dernières figures comme étant plus régulières que H_autres dans la mesure où il y a itération d'une même cellule métrique (01 ou 001). Le couple Eros et Thanatos se voit ici caractérisé par des modalités rythmiques différentes :

	/amour/ (cl4, cl5, cl6, cl1)	**/mort/ (cl2, cl7, cl8, cl3)**	**Ensemble**
H_001001	25%	30%	28%
H_010101	38%	39%	38%
H_autres	37%	31%	34%
Total	100 %	100 %	100 %

La probabilité que cette répartition soit due au hasard est infime (probabilité inférieure à 2.10^{-96} avec le test du Chi-deux).

Si les séquences de vers sont dans un premier temps organisées en deux grandes catégories dominées respectivement par les sèmes de l'amour et de la mort, un découpage plus fin permet d'identifier huit univers lexico-sémantiques plus spécifiques, dont quatre sont rattachées au pôle /amour/ (« jeu et mensonge », « marivaudage[160] », « passion amoureuse », « trône et hymen») et quatre au pôle /mort/ (« gloire et honneur », « victoire ou défaite », « mort et culpabilité», « l'ailleurs »). Nous avons rapidement interprété le contenu et baptisé ces univers à partir de leur vocabulaire spécifique et des séquences de vers les plus caractéristiques. Nous présentons pour chaque classe les isotopies que nous avons identifiées en rattachant au sème isotopant les mots qui le contiennent : le recensement des isotopies est loin d'être exhaustif et seules les principales, les plus évidentes, sont signalées. La liste complète du vocabulaire spécifique de chaque classe est présentée dans l'annexe IX.

[160] On nous permettra cet anachronisme.

La première classe (classe 5 : jeu et mensonge) permet d'identifier quatre domaines sémantiques : ceux du jeu, du mensonge, des affaires et de la lettre. Les catégories syntaxiques qui dominent sont les marqueurs de la modalisation et de l'intensité, les verbes modaux, les auxiliaires *être* et *avoir* et les interjections (*eh*, *oh*). Cette classe allie la topique du rendez-vous galant à celle du mensonge et de la duperie. La situation prototypique est la visite nocturne, où le galant, déguisé, passe incognito. La lettre et le billet, messages décontextualisés, instruments du quiproquo, sont particulièrement adaptés à ces situations d'ambiguïté.

Voici quelques isotopies (accompagnées de la liste des mots présentant le sème isotopant) construites à partir de l'examen du vocabulaire spécifique de la classe :

/divertir/ : jeu, conter, divertir, jouer, déguiser, jeu, loisir.
/mensonge/ : mentir, fourbe, déguiser, histoire, ruse, double, croyance.
/lettre/ : billet, adresse, lettre, mot, écrire.
/affaires/ : affaire, argent, crédit, vendre.
/rencontre/ : venir, rencontrer, entretenir, entrer, parler, surprendre.
/soir/ : nuit, soir, tard.

Voici quelques fragments de pièces répertoriés dans cette classe (nous n'avons retenu dans ces fragments, qui font en général dix vers, que les vers les plus significatifs) :

H1_000101 H2_001001	Me divertir une heure à m'en faire justice :
H1_010001 H2_010001	Ces lettres fourniront assez d'occasion
H1_010001 H2_000001	D'un peu de défiance et de division.
H1_001101 H2_011001	Si je prends bien mon temps, j'aurai pleine matière
H1_000100 H2_001001	À les jouer tous deux d'une belle manière.

Corneille, *Mélite*, vers 1016-1020.

H1_011001 H2_110101	Il est vrai que je sors fort peu souvent de jour :
H1_010001 H2_010001	De nuit, incognito, je rends quelques visites ;

Corneille, *Le Menteur*, vers 256-257.

H1_010101 H2_000101	Je crains à tous moments qu'on me surprenne ici ;
H1_010001 H2_001001	Encor que déguisée, on pourrait me connaître.

Corneille, *Le Menteur*, vers 1076-1077.

H1_001001 H2_001001	Du grotesque récit de vos rares exploits
H1_001001 H2_010101	La sottise ne plaît qu'alors qu'elle est nouvelle :
H1_001101 H2_010001	En un mot, faites reine une autre qu'Isabelle.

Corneille, *L'Illusion comique*, vers 720-723.

La seconde classe (classe 4 : « marivaudage ») partage avec la première le champ du mensonge et de la tromperie, mais elle est centrée sur la relation amoureuse : triangle amoureux, personnages volages... Les pronoms personnels et les marqueurs d'intensité sont spécifiques de la classe.

Voici les isotopies caractéristiques de cette classe :

/tromperie/ : feint, volage, soupçon, déloyal, changer, faux, trahison, ruse, dissimuler, infidèle, perfidie, feindre, trahir, jalousie, présumer, tromper, confus.

/parole/ : discours, entretien, propos.

/amour/ : amant, feu, affection, appas, attrait, brûler, ardeur, inclination, maîtresse, amour, flamme.

/dédain/ : dédain, mépris, blâmer, endurer, froideur, glace.

/parties du corps/ : visage, bouche.

/chaud/ : feu, brûler.

/froid/ : froideur, glace.

Voici des extraits spécifiques de cette classe :

H1_001001 H2_010001	Mais Daphnis se contraint, de peur de vous déplaire,
H1_001101 H2_000101	Et sa bouche est toujours à ses désirs contraire,

Corneille, *La Suivante*, vers 845-846.

H1_101001 H2_000101	J'aime ailleurs ; et lassé d'un compliment frivole,
H1_001001 H2_010001	Et de feindre une ardeur qui blesse mes amis,
H1_011001 H2_000101	Ma flamme est véritable et son effet permis.
H1_010001 H2_010011	J'adore une beauté qui peut disposer d'elle,
H1_000101 H2_001001	Et seconder mes feux sans se rendre infidèle.

Corneille, *La Suivante*, vers 1170-1174.

H1_010001 H2_001001	Il tient pour Angélique, il la suit, le rebelle !
H1_000001 H2_011001	Parmi mes trahisons il veut être fidèle ;

Corneille, *La Place royale*, vers 1001-1002.

H1_011001 H2_001001	Ce n'est pas sans raison que ma feinte passée
H1_001001 H2_000001	A jeté cette erreur dedans votre pensée.

Corneille, *La Galerie du palais*, vers 1557-1558.

Avec la classe 6 (passion amoureuse), nous restons dans le domaine de l'amour, mais la tromperie cède la place à l'alternance amour/haine. L'intensité des sentiments atteint des degrés plus élevés, comme en témoigne l'importance des marqueurs d'intensité.

Voici quelques isotopies caractéristiques de cette classe :

/amour/ : amour, cœur, flamme, passion, aimer, désir, sentiment, charmer, allumer, adorer.
/rupture amoureuse/ : haïr, ingratitude, éteindre, soupir, soupirer, inconstant, ingrat, dédain, mépris.
/feu/ : flamme, éteindre, allumer.
/douceur/ : tendre, doux, charme, douceur, grâce, bonté.
/force/ : force, rude, effort, forcer, imposer, dur, rigoureux.

et quelques fragments spécifiques :

H1_000101 H2_100101	Et par pitié d'un cœur trop ardemment épris,
H1_000101 H2_010001	Éteignez-en la flamme à force de mépris.

Corneille, *Agésilas*, vers 1656-1657.

H1_010101 H2_010001	Je veux qu'en vain son âme, esclave de la mienne,
H1_001001 H2_001001	Me demande sa grâce, et jamais ne l'obtienne,
H1_001001 H2_000011	Qu'il soupire sans fruit ; et pour le punir mieux,
H1_011001 H2_001001	Je veux même à mon tour vous aimer à ses yeux.

Corneille, *Pertharite*, vers 463-466.

Avec la classe suivante (classe 1 : trône et hymen), nous changeons de domaine sémantique pour entrer pleinement dans celui du pouvoir. Mais il est encore question d'amour, par le biais de l'hymen. Le mariage peut renforcer la

puissance, l'amour s'opposer au pouvoir... L'alliance des deux ne va pas de soi et elle est sujette à délibérations : choix, refus, obéissance.

On notera que le sème /grand/ est récurrent dans ces fragments et qu'il n'est jamais remis en cause par la moindre occurrence, dans son environnement proche, de mots portant un sème d'évaluation négative. On comprend que ces pièces, écrites pour la cour aient pu être appréciées : elles sont un merveilleux faire-valoir pour le pouvoir royal.

On trouvera ci-dessous quelques-unes des isotopies construites à partir du vocabulaire de cette classe :

/le pouvoir/ : maître, souverain, empereur, empire, état, pouvoir, puissance, rang, reine, roi, autorité, sénat, politique, droit, loi, respect, gouvernement, impératrice, monarque, commander, cour, Seigneur.
Métaphore du pouvoir : couronne, trône, sceptre, diadème.
/grand/ : grand, haut, grandeur, illustre, suprême, éclat.
/mariage/ : époux, reine, épouser, hymen, mari, union, impératrice.
/choix/ : choix, refus, choisir, disposer, proposer, offrir.

et quelques fragments caractéristiques :

H1_001001 H2_000101	Vous n'osez m'accepter, et disposez de moi.
H1_001001 H2_001001	Accordez le respect que mon trône vous donne

Corneille, *Sertorius*, vers 556-557.

H1_010001 H2_000101	Remet à notre reine à nous choisir un roi.
H1_000011 H2_010001	Ne vous abaissez plus à suivre la couronne :
H1_001001 H2_000001	Donnez-la, sans souffrir qu'avec elle on vous donne ;
H1_010001 H2_010101	Réglez notre destin, qu'ont mal réglé les dieux :
H1_001101 H2_101001	Notre seul droit d'aînesse est de plaire à vos yeux ;

Corneille, *Rodogune*, vers 910-914.

Nous quittons ici les univers de discours où domine la thématique de l'amour pour aborder ceux où domine celle de la mort.

Dans la classe 2 (gloire et honneur), le vocabulaire spécifique est centré sur l'énonciation des qualités des personnages. Deux univers sémantiques s'opposent terme à terme : ceux de la gloire et de l'infamie. Cette articulation des deux figures antithétiques est mobilisée par le personnage dont l'honneur est mis en péril.

Les isotopies que nous avons pu identifier sont :

/honneur/ : gloire, honneur, mémoire, nom, vertu, digne.
/évaluation positive/ : généreux, gloire, honneur, mémoire, nom, vertu, digne, magnanime, brave.
/évaluation négative/ : ennemi, infâme, lâche, crime, infamie, parricide, souillé, ignominie, honte, indigne, lâcheté, déshonorer.

Voici quelques fragments caractéristiques :

H1_001001 H2_010001	Si ma main en devient honteuse et profanée,
H1_001011 H2_010001	Vous pouvez d'un seul mot trancher ma destinée :
H1_001001 H2_000001	Reprenez tout ce sang de qui ma lâcheté
H1_000001 H2_010001	A si brutalement souillé la pureté.

Corneille, *Horace*, vers 1423.

H1_001001 H2_010001	Souviens-toi de ton nom, soutiens sa dignité ;
H1_001001 H2_000001	Et prenant d'un Romain la générosité,
H1_100011 H2_001011	Sache qu'il n'en est point que le ciel n'ait fait naître
H1_000101 H2_001001	Pour commander aux rois, et pour vivre sans maître.

Corneille, *Cinna*, vers 999.

La classe suivante (classe 7 : victoire ou défaite) met en scène le combat, et ses issues : victoire ou défaite. Les déictiques spatiaux et temporels sont très spécifiques de cette classe. Il s'agit de récits de combats rapportés par un personnage, puisque l'unité de lieu interdit le passage de l'antichambre au champ de bataille.

Voici les isotopies que nous avons pu identifier à partir du vocabulaire spécifique :

/combat/ : arme, défait, vainqueur, champ, combat, conquête, défaite, guerre, exploit, guerrier, héros, laurier, paix, soldat, vaincu, victoire, combattre, conquérir, bataille, captif, bras, fer, attaquer, triompher, vaincre, prisonnier, danger, char.
/victoire/ : vainqueur, conquête, héros, laurier, valeur, vaillant, glorieux.
/défaite/ : vaincu, défaite, tomber, captif, succomber, chuter.

et quelques vers caractéristiques :

H1_011001 H2_000101	Ce grand roi fut défait, il en perdit la vie,
H1_001001 H2_001001	Et laissait sa couronne à jamais asservie,
H1_000101 H2_001001	Si pour briser les fers de son peuple captif,
H1_100001 H2_010001	Rome n'eût envoyé ce noble fugitif.

Corneille, *Sertorius*, vers 441.

H1_010001 H2_010011	En vain quelques guerriers, qu'anime son grand cœur,
H1_000101 H2_001001	Ont ramené l'effroi dans le camp du vainqueur :
H1_011001 H2_001001	Il faut bien qu'il succombe, et qu'enfin son courage
H1_100101 H2_010001	Tombe sur tant de morts qui ferment son passage.

Racine, *Alexandre le Grand*, vers 1359.

La classe 8 (mort et culpabilité) nous fait entrer dans un univers où dominent les noms, les noms propres et les interjections. Cette classe s'articule autour de la mort et de la douleur. Les sentiments y sont exacerbés.

Les isotopies de cette classe sont :

/mort/ : fatal, funeste, mort, sanglant, bourreau, vie, victime, expirer, mourir, percer, mortel, trépas, poison, tombeau, étouffer.
/colère/ : colère, rage, haine, furie, éperdu, furieux.
/coupable/ : tourment, vengeance, punir, innocent, châtiment, enfer.
/douleur/ : malheureux, désespoir, douleur, larme, malheur, pleur, triste, affliger, infortune.
/cruauté/ : horreur, cruel, supplice, criminel, inhumain, injuste.
interjections : Ô, ciel, ah, hélas.

et les vers caractéristiques :

H1_110001 H2_011001	Ô ciel, que tes rigueurs seraient peu redoutables,
H1_001001 H2_001001	Si la foudre d'abord accablait les coupables !
H1_000001 H2_010001	Et que tes châtiments paraissent infinis,
H1_001001 H2_010001	Quand tu laisses la vie à ceux que tu punis !

Racine, *La Thébaïde*, vers 595.

H1_101001 H2_001001	Oui, cruelles, en vain vos injustes secours
H1_010001 H2_001011	Me ferment du tombeau les chemins les plus courts,
H1_000101 H2_000011	Je trouverai la mort jusque dans vos bras même.

Racine, *Mithridate*, vers 1497.

La dernière classe (classe 3 : l'ailleurs) se situe dans l'univers du sacré et elle est marquée par l'évocation d'espaces extérieurs à la scène, constitutifs d'une forme de paysage stylisé : mer, port, palais, ciel, nuage... Une isotopie à connotation négative associe à cet ailleurs une menace.

Voici les isotopies caractéristiques de cette classe :

/sacré/ : sacré, saint, autel, cieux, déesse, dieu, prêtre, profane, temple, oracle.
/espace hors-scène/ : flot, mer, palais, port, temple, vaisseau, vent, nuage, rive, lieu, séjour, soleil, ciel.
/famille/ : fille, enfant, mère.
/mal/ : affreux, perfide, sombre, monstre, méchant.

et quelques vers spécifiques :

H1_001001 H2_001001	Au delà de ce lieu gardez-vous d'avancer.
H1_100101 H2_001001	C'est des ministres saints la demeure sacrée ;
H1_010101 H2_001001	Les lois à tout profane en défendent l'entrée.
H1_000101 H2_001001	Qui cherchez-vous ? mon père, en ce jour solennel,
H1_000101 H2_101001	De l'idolâtre impur fuit l'aspect criminel ;

Racine, *Athalie*, vers 850.

H1_010011 H2_010101	À peine son sang coule et fait rougir la terre,
H1_011001 H2_010001	Les dieux font sur l'autel entendre le tonnerre ;
H1_010101 H2_010001	Les vents agitent l'air d'heureux frémissements,
H1_001001 H2_000001	Et la mer leur répond par ses mugissements ;

Racine, *Iphigénie*, vers 1777.

.

Nous venons d'interpréter rapidement le contenu des classes de cette typologie : « jeu et mensonge », « marivaudage », « passion amoureuse », « trône et hymen», « gloire et honneur », « victoire ou défaite », « mort et culpabilité», « l'ailleurs ».

A ce niveau de détail, examinons les liens existant entre univers de discours et rythme. On voit émerger des associations nettes entre des thématiques et des rythmes. L'interaction est très significative ($P < 6.10^{-115}$). Prenons les deux extrémités de la Figure 42 : la classe 5 « jeu et mensonge » se caractérise par un taux élevé d'hémistiches irréguliers, tandis que la classe 3 « l'ailleurs » a un taux élevé d'hémistiches réguliers marqués en position 3 (H_001001), tout comme la classe 7. Ces deux classes ont comme point commun de correspondre aux récits d'événements qui se déroulent hors de la scène (dans les lieux sacrés, dans les espaces naturels et sur les champs de bataille). On remarquera que les variations selon les classes sont fortes pour les figures H_001001 et H_autres, mais que la proportion d'hémistiches de type H_010101 est assez constante quelles que soient les classes.

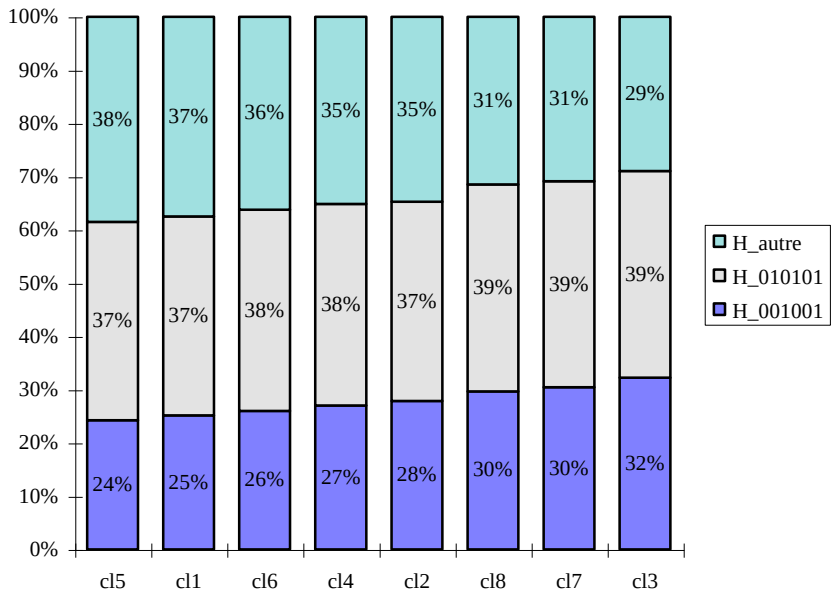


Figure 42. Rythme et univers lexicaux chez Corneille et Racine

On sent bien que chacun des univers lexico-sémantiques recouvre des thématiques propres à des genres et aux deux auteurs de notre corpus. Nous pouvons maintenant examiner comment ces univers se répartissent dans chaque pièce. Le profil varie considérablement selon les genres et les auteurs, comme le montre la Figure 43, p. 366.

Le bas du graphique est occupé par des pièces dans lesquelles dominent largement les classes 5 (jeu et mensonge) et 4 (marivaudage) : ce sont toutes des comédies. On notera que la pièce *Les Plaideurs* est presque exclusivement constituée d'un seul univers de discours (jeu et mensonge). La transition entre comédies et tragédies est assurée par deux pièces : *L'Illusion comique* et *Clitandre*.

Après les comédies, nous entrons dans l'univers des tragédies et pièces diverses. Alors que les classes 4 et 5 représentaient plus de 80 % de chacune des comédies, elles entrent pour moins de 10 % dans les autres pièces. La présence majoritaire de contextes appartenant aux classes 4 et 5 semble marquer le genre de la comédie. Si nous continuons la progression du bas vers

le haut, nous voyons apparaître, après les comédies, les pièces où dominent les classes centrées sur la passion amoureuse et sur les liens entre pouvoir et hymen. Viennent ensuite quelques tragédies où l'exaltation des qualités du personnage et le refus de l'infamie prennent une place importance (classe 2). La classe 2 apparaît surtout dans les tragédies de Corneille. Nous y retrouvons l'exaltation de soi comme une caractéristique du héros cornélien [Bénichou, 48] qui s'affirme dans un vouloir être :

> La louange de soi-même, si fréquente chez Corneille, constitue l'équivalent verbal de la répétition de l'exploit glorieux.
>
> [Starobinski, 1961, p. 58].

Cette exaltation de soi contribue à la renommée, et illustre la grandeur du nom : « la gloire prend corps dans le Nom » [Starobinski, 1961, p. 61].

La classe 7 (victoire ou défaite) est moins spécifique d'un genre ou d'un auteur particulier, mais elle occupe une place écrasante dans *Alexandre le Grand*.

Les classes 2 et 7 qui ont en commun l'opposition entre la gloire et l'infamie d'un côté et la victoire et la défaite de l'autre caractérisent un certain nombre de tragédies de Corneille : *Horace*, *Héraclius*, *Le Cid*, *Cinna*, *Pompée*, *Nicomède*... Les deux premières pièces de Racine, *La Thébaïde* et *Alexandre*, ont des profils proches de certaines tragédies de Corneille, par l'importance des classes 2 et 7 (exaltation de soi, victoire).

De plus, toujours en remontant vers le haut du graphique, les classes liées à la mort et à la culpabilité d'une part, et à l'évocation d'un ailleurs (monde des dieux, espace hors-scène) d'autre part, prennent de l'importance. On notera que la progression de ces deux thématiques suit l'évolution chronologique des pièces de Racine. En particulier, le champ lexical du sacré progresse au fil des tragédies de Racine, pour atteindre une place prépondérante dans les deux dernières pièces *(Esther* et *Athalie*).

> Racine a commencé par des tragédies conformes au goût du temps, et on aurait peine à différencier nettement la *Thébaïde* et l'*Alexandre* des tragédies de Corneille et de ses successeurs.
>
> [Bénichou, 1948, p. 176].

Les univers de discours très représentés dans les tragédies et surtout les classes 8 et 3 sont quasiment absents des comédies. Le pôle /mort/ qui correspond au premier niveau de découpage que nous avons identifié y est totalement exclu. En reprenant les deux pôles initiaux qui découpent le corpus, on pourrait être amené à conclure que la thématique de l'amour est minoritaire dans une bonne partie des tragédies. Cette impression provient du calcul des mots spécifiques de chaque classe qui indique les mots qui sont relativement plus présents dans une classe, ce qui ne veut pas dire qu'ils ne soient pas présents dans l'autre. La thématique de l'amour est relativement moins prégnante dans les tragédies que dans les comédies et surtout la passion amoureuse est souvent attachée à la thématique de la mort.

A quelques exceptions près, les zones de recouvrement entre tragédie et comédie sont quasiment inexistantes. Nous avions fait un test [Beaudouin, 1993] en prenant quatre pièces, une tragédie et une comédie pour Racine et Corneille, que nous avions soumises au même type de découpage et de classification. Il apparaissait que le premier niveau de classification opposait clairement presque tous les fragments (plus de 90 %) des comédies à ceux des tragédies. Ce n'était qu'au deuxième niveau de la classification qu'apparaissait l'opposition entre auteurs à l'intérieur de chacun des genres. Pour cette expérience, nous avions retenu des pièces très représentatives des genres étudiés (*Le Menteur*, *Sertorius*, *Les Plaideurs*, *Andromaque*).

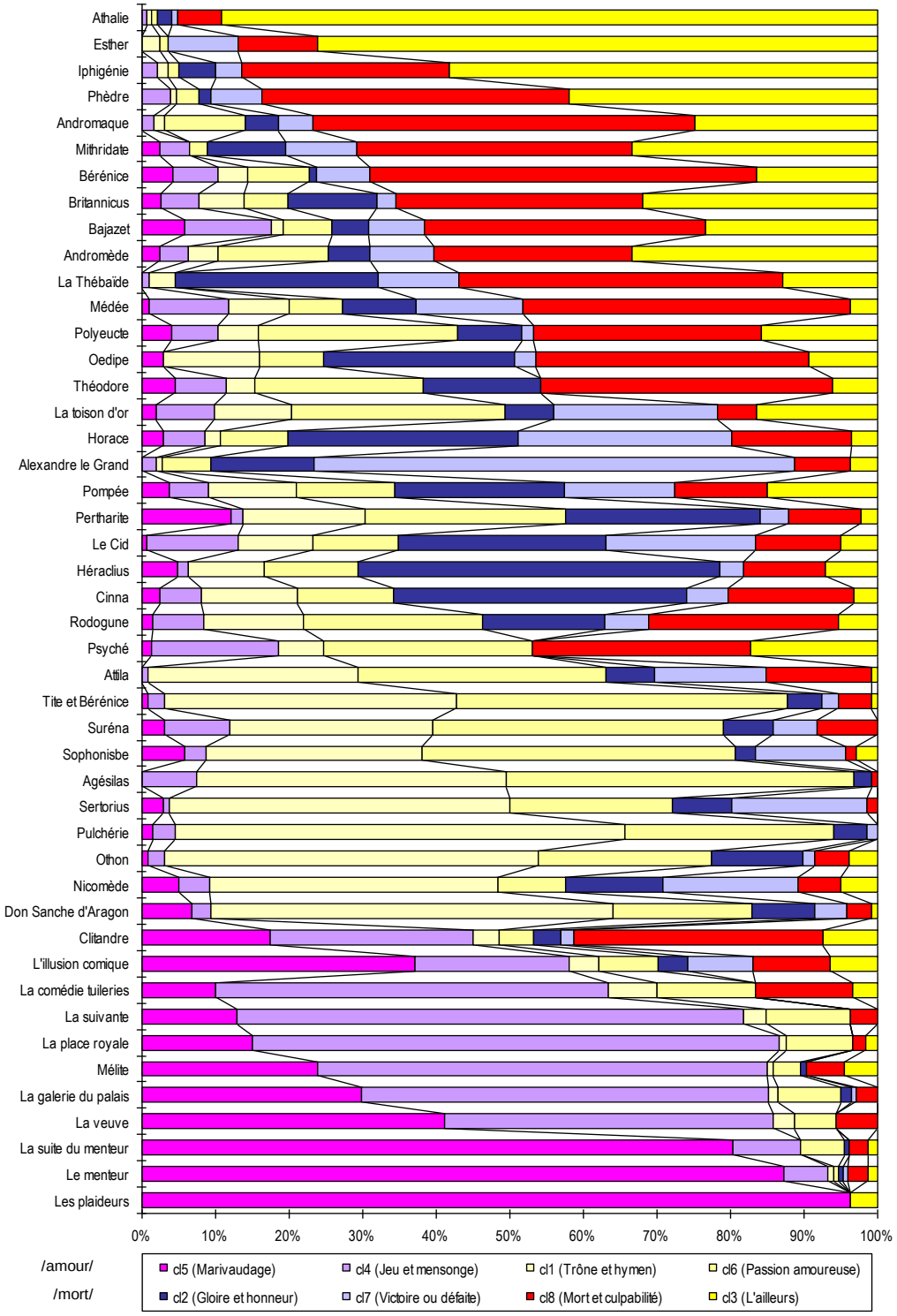
Revenons maintenant aux pièces situées à la frontière des deux genres. *L'Illusion comique*, classée par Corneille comme comédie, a un profil équilibré entre comédie et tragédie. La tragédie qui se joue dans la pièce, et qui s'avère être une représentation, introduit les univers lexicaux de la tragédie dans la pièce. On comprend que Ch. Muller se soit appuyé sur cette œuvre pour démontrer les différences lexicales entre les genres. Le cas de *Clitandre* est tout aussi intéressant. Dans cette pièce, les univers de discours se partagent en effet de manière équilibrée entre tragédie et comédie. Cette œuvre influencée par le théâtre du siècle d'or espagnol, tout en relevant de la tragédie, est cependant assez proche du profil des comédies, ce dont a bien rendu compte la mise en

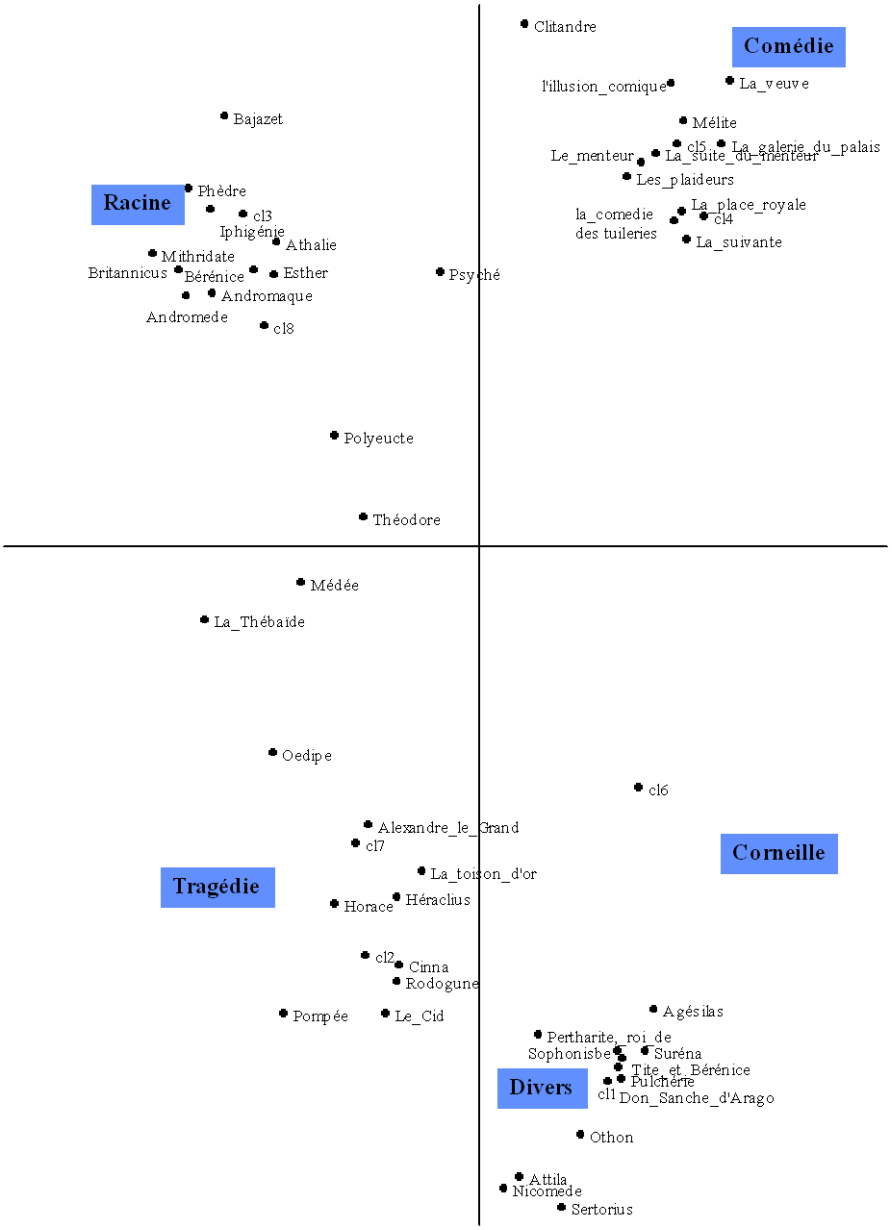
scène de Muriel Mayette à la Comédie-Française en 1997. Les nombreuses péripéties qui traversent la pièce peuvent expliquer ce profil hybride.

Figure 43. Profils lexico-sémantiques des pièces

(Page suivante)

Clef de lecture : *L'Illusion comique* est composée de près de 40 % de fragments de la classe « marivaudage », de près de 20 % de la classe « jeu et mensonge », de 5 % de « trône et hymen », de 10 % de « passion amoureuse », etc. et enfin d'environ 5 % venant de « l'ailleurs ».


Athalie
Esther
Iphigénie
Phèdre
Andromaque
Mithridate
Bérénice
Britannicus
Bajazet
Andromède
La Thébaïde
Médée
Polyeucte
Oedipe
Théodore
La toison d'or
Horace
Alexandre le Grand
Pompée
Pertharite
Le Cid
Héraclius
Cinna
Rodogune
Psyché
Attila
Tite et Bérénice
Suréna
Sophonisbe
Agésilas
Sertorius
Pulchérie
Othon
Nicomède
Don Sanche d'Aragon
Clitandre
L'illusion comique
La comédie tuileries
La suivante
La place royale
Mélite
La galerie du palais
La veuve
La suite du menteur
Le menteur
Les plaideurs
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
/amour/
/mort/
cl5 (Marivaudage)
cl4 (Jeu et mensonge)
cl1 (Trône et hymen)
cl6 (Passion amoureuse)
cl2 (Gloire et honneur)
cl7 (Victoire ou défaite)
cl8 (Mort et culpabilité)
cl3 (L'ailleurs)



Clef de lecture : l'axe horizontal oppose, à droite, le pôle /amour/ (comédies, pièces diverses) au pôle /mort/ (tragédies), à gauche. L'axe vertical oppose les tragédies de Racine (en haut à gauche) à celles de Corneille (en bas à gauche) et les comédies (en haut à droite) aux pièces diverses.

Figure 44. Premier plan de l'analyse factorielle

La Figure 44 représente sur le premier plan de l'analyse factorielle les distances entre pièces, au regard du vocabulaire. On y constate la grande proximité des tragédies de Racine à partir d'*Andromaque*, sensiblement différentes de celles de Corneille, la confusion entre *Les Plaideurs* et les comédies de Corneille, et la particularité des comédies héroïques et pièces diverses qui occupent le cadran en bas à droite[161].

Ayant observé qu'il existait des liens, d'une part, entre des univers de discours et des rythmes particuliers et, d'autre part, entre ces mêmes univers de discours et des genres et auteurs spécifiques, il nous reste à voir si les corrélations observées ne sont pas des effets de liens directs entre auteur et genre et types de rythmes.

VIII.1.2. Les genres : des indices convergents de spécificité

Tout a déjà été dit ou presque sur les différences de genres au théâtre. Ainsi, l'examen de la question des genres chez Corneille et Racine sera-t-il pour nous l'occasion de tester sur un domaine balisé la pertinence des outils utilisés.

Comme nous l'avons vu lors de la description du corpus (chapitre II), nous disposons de l'ensemble de quarante-six pièces de théâtre de Corneille et Racine. Si Corneille s'est illustré dans la comédie comme dans la tragédie, Racine, quant à lui, n'a écrit qu'une comédie. Corneille et Racine ont classé chacune de leurs pièces dans un genre théâtral : comédie, tragédie, tragi-comédie... Ce classement a pu changer dans le temps, sans que le contenu des pièces n'ait évolué. Ainsi *Clitandre* et *Le Cid* furent des tragi-comédies avant de devenir des tragédies [Muller, *1967*, 1992, p. 24].

Rappelons quelques caractéristiques de base des genres.

---

[161] Nous aurions également pu nous intéresser à l'évolution du vocabulaire chez chacun de ces auteurs ou à des indices de proximité du vocabulaire entre les pièces. Nous renvoyons pour cela à Muller [*1967*, 1992], à Bernet [1983] et aux travaux de D. Labbé sur l'accroissement du vocabulaire [1993] et sur la connexion des vocabulaires [1998], tous ces travaux s'appuyant sur les textes de Corneille et/ou de Racine.

La tragédie se distingue de la comédie par la condition des personnages, par la nature des situations, actions et sentiments et par le dénouement [Couprie, p. 3].

> La plainte des Seigneurs fut dite Tragédie.
> L'action du commun fut dite Comédie.
>
> Ronsard, *Pièces hors recueils*, «à O. de Magny», Élégie.

La tragédie se doit, comme le rappelle Corneille dans le *Discours sur le poème dramatique*, de présenter des actions et sentiments exemplaires.

> Lorsqu'on met sur la scène une simple intrigue d'amour entre des rois, et qu'ils ne courent aucun péril, ni de leur vie, ni de leur État, je ne crois pas que, bien que les personnes soient illustres, l'action le soit assez pour s'élever jusqu'à la tragédie. Sa dignité demande quelque grand intérêt d'État ou quelque passion plus noble et plus mâle que l'amour, telles que sont l'ambition ou la vengeance (...). La tragédie (...) veut pour son sujet une action illustre, extraordinaire, sérieuse (...).
>
> Corneille, Discours sur le poème dramatique.

Si la tragédie se finit mal, la comédie a en général un heureux dénouement.

Les modalités de la réception distinguent également comédie et tragédie. Nous retrouvons le concept de catharsis d'Aristote, pour qui la tragédie, contrairement à la comédie, vise à purger les passions par le biais de la crainte et de la pitié. Si les deux genres visent le plaisir du spectateur, par l'entremise de la dénégation-illusion [Ubersfeld, *1977*, 1982, p. 42], la comédie a pour fonction de divertir le public, tandis que la tragédie doit conduire à cette purgation des passions. A. Ubersfeld apporte une interprétation cohérente de la catharsis aristotélicienne, qui permet de comprendre comment crainte et plaisir peuvent être mêlés dans la réception :

> (...) La construction d'un réel concret qui est en même temps l'objet d'un jugement qui en nie l'insertion dans la réalité, libère le spectateur, qui voit s'accomplir ou s'exorciser ses craintes et ses désirs, sans qu'il en soit victime, mais non sans participation.
>
> [Ubersfeld, *1977*, 1982, p. 43].

Le drame bourgeois au XVIIIe siècle, et surtout le drame hugolien, refuseront la distinction des genres au nom de la vraisemblance. Mêler le grotesque au sublime, mettre de côté les règles d'unité de temps et de lieu, vouloir un vers « passant d'une naturelle allure de la comédie à la tragédie, du sublime au grotesque » pour que l'œuvre soit le miroir de la nature, tel est le projet qu'Hugo expose dans la préface de *Cromwell*. A l'époque classique, nous

sommes dans une période idéale pour identifier les caractéristiques des genres, car les frontières sont clairement marquées, et les genres nettement définis. Néanmoins, ce propos doit être nuancé : certaines pièces sont déjà chez Corneille à la frontière des genres. C'est le cas de *L'Illusion comique* qui en son sein représente une tragédie, mais aussi de *Clitandre*, comme a pu nous le montrer la précédente analyse.

Puisque comédie et tragédie ont des effets opposés sur leur public, divertissement d'un côté, catharsis de l'autre, on peut penser que les ressorts stylistiques qui provoquent ces effets diffèrent d'un genre à l'autre. En dehors de l'étiquette du genre qui caractérise la pièce, quelles sont les traces stylistiques de ces différents genres que nous pouvons identifier ? Quels indices, étant donné les outils dont nous disposons, pouvons-nous construire pour identifier d'éventuelles spécificités de genre ?

Le lexique et la prosodie du vers portent-ils les traces de la différence de genre ? Y a-t-il cohérence entre les marquages lexicaux et rythmiques ?

Nous utiliserons différents types de descripteurs (longueur des mots, type de vocabulaire, figures d'hémistiches) et nous verrons à l'aide de techniques statistiques s'ils sont ou non répartis de manière différenciée entre les genres et les auteurs.

VIII.1.2.1. Nombre de mots par vers

Le nombre de mots par vers est un indice stylistique simple pour caractériser un auteur, un genre, une pièce par rapport à d'autres textes. Charles Muller [1967, 1993], et plus tard Charles Bernet [1983], dans leurs études respectives sur le vocabulaire de Corneille et Racine, ont étudié le nombre moyen de mots par alexandrins, le rapport N/A.

Muller a montré que, chez Corneille, N/A était toujours plus élevé dans les comédies que dans les tragédies, à condition de comparer des pièces de la même époque [Muller, 1967, 1993, p. 50] ; Bernet [1983, p. 162] a montré que le nombre de mots par vers est de 9,201 chez Corneille et de 8,871 pour Racine. Muller [1967, 1993] a aussi montré qu'au fil du temps, le nombre moyen de mots par vers avait eu tendance à augmenter chez Corneille. Le rapport N/A

offre donc un premier indice stylistique de différenciation des genres, des auteurs et des périodes.

Nos décomptes ne conduisent pas forcément à des résultats identiques à ceux de Muller et Bernet, puisque les normes de dépouillement ne sont pas les mêmes. Comme le préconise Muller [1977], nous avons conservé la même norme de dépouillement pour tous les corpus, ce qui nous permet d'établir des comparaisons. Ici, on a considéré que tous les signes de ponctuation étaient des séparateurs, sauf le tiret et l'apostrophe. Ainsi *a-t-il* est compté comme un mot tout comme *j'aime*. Cette norme, imparfaite par certains aspects, a l'avantage d'être assez cohérente avec des unités prosodiques. Ainsi quand le pronom est placé après le verbe, le marquage accentuel le plus fort a-t-il tendance à se reporter sur lui.

Va-t'**en** chercher Philandre, et dis-**lui** que Mélite

Corneille, *Mélite*, vers 571.

Les résultats produits avec notre norme de dépouillement sur l'ensemble des pièces apparaissent dans le Tableau 36.

	Tragédie	Comédie	Divers	Total
Nombre de vers	50181	15359	11646	77186
Moyenne (N /A)	8,17	8,21	8,28	8,20
Ecart-type	1,30	1,31	1,30	1,30

Tableau 36. Nombre de mots par vers selon les genres

Cette différence entre les moyennes est-elle significative ? Le test d'égalité des moyennes conduit à rejeter l'hypothèse d'égalité. Le nombre de mots par vers est significativement plus élevé dans la comédie que dans la tragédie : les mots y sont en moyenne un peu plus courts. Comme, en général, les mots grammaticaux sont plus courts, on peut se demander si cette différence ne tient pas à un taux de positions métriques occupées par des mots grammaticaux plus élevé.

VIII.1.2.2. Caractéristiques lexicales des deux genres

Existe-t-il des mots spécifiques de chaque genre ? Si oui, quels sont-ils ?

Dans son ouvrage sur les pièces de Corneille, Muller [*1967*, 1993, p. 139-142] a montré que le vocabulaire ne se répartissait pas de manière stable selon les époques et les genres. Ainsi, les pronoms de dialogue (première et deuxième personne), les noms propres et les substantifs appellatifs (*Seigneur, Madame, Sire, Prince, Monsieur...*) sont-ils répartis de manière très différente selon les genres. Le vocabulaire spécifique de la tragédie est constitué par des noms de personnages de haut rang (*roi, reine, prince, tyran, peuple, empire, état, trône, gloire, sang, mort...*), par un certain vocabulaire psychologique (*chagrin, vertu, haine, choix*) et par quelques mots comme *héros, cœur, fierté, hymen, devoir, pouvoir...*

A ce vocabulaire, qui apparaît de manière contrastée selon les époques, il faut ajouter le vocabulaire spécifique de la tragédie qui reste stable au fil du temps : *bras, tête, vaincre, contre, justice, horreur, trépas, forcer, soutenir, poison, soldat, mortel*. A l'inverse, le vocabulaire caractéristique de la comédie renvoie pour Muller à un ensemble de termes relatifs à la vie mondaine (*esprit, humeur, feu, affection, beauté, ami, maîtresse, obliger, discours...*) et à des mots que l'on retrouve d'une manière constante sur les trois périodes : *je, mot, parole, prendre, discret, chez, autant, quelque, deviner...*

Le logiciel Hyperbase nous a permis d'identifier chez Corneille et Racine les termes les plus caractéristiques de chacun de ces genres théâtraux. Il s'agit maintenant de formes graphiques telles qu'elles apparaissent dans les corpus et non d'unités lexicales lemmatisées[162] comme chez Muller. On a choisi une norme interne constituée par l'ensemble de la production théâtrale de chacun de ces auteurs. On compare donc les tragédies de Corneille, puis ses comédies, à l'ensemble de son œuvre ; et enfin l'unique comédie de Racine à l'ensemble de sa production.

Le tableau suivant présente, pour chacun des cas, la liste des termes les plus caractéristiques, par spécificité décroissante. Dans la première colonne, on lit la fréquence du mot dans le sous-corpus (pour le premier cas, il s'agit des tragédies de Corneille), dans la seconde, sa fréquence dans l'ensemble du

[162] La lemmatisation consiste à ramener les mots à leur racine, c'est-à-dire à ce qui serait leur entrée dans un dictionnaire : toutes les formes fléchies des verbes sont ramenées à l'infinitif, les noms au singulier et les adjectifs au masculin singulier.

corpus et dans la troisième un indice de spécificité ou d'écart. Volontairement, nous avons supprimé les noms propres qui apparaissent dans les listes de spécificité.

Tragédies de Corneille / pièces de Corneille				**Comédies de Corneille / pièces de Corneille**				**Comédie *Les Plaideurs* / pièces de Racine**			
Fréq. trag.	Fréq. com. + trag..	Ecart	Mot	Fréq. com.	Fréq. com. + trag.	Ecart	Mot	Fréq. Plai-deurs	Fréq. com. + trag..	Ecart	Mot
651	726	15	roi	137	154	13	monsieur	115	115	44	monsieur
847	1027	13	seigneur	966	2011	11	tu	32	38	21	bon
582	674	12	sang	115	151	10	humeur	18	18	17	affaire
547	670	10	mort	145	207	10	maîtresse	18	18	17	messieurs
484	603	9	gloire	543	1058	10	te	18	18	17	procès
298	348	9	prince	63	73	9	amours	13	13	15	sergent
260	298	9	rois	203	345	9	fort	13	13	15	souffleur
166	172	9	tyran	1981	4704	8	?	12	13	13	exploit
339	408	8	haine	127	196	8	ami	23	45	13	là
117	120	8	romains	89	121	8	beauté	207	1623	12	!
274	319	8	vertu	162	264	8	belle	92	580	11	à
269	328	7	bras	136	224	8	discours	73	403	11	bien
129	143	7	couronne	92	134	8	nuit	13	21	11	fort
337	416	7	crime	50	64	7	affaire	25	79	10	j‘
400	508	7	dieux	859	1947	7	bien	17	40	10	juge
314	395	7	fils	30	30	7	blanche	11	19	10	juger
136	151	7	héros	81	119	7	cela	12	25	9	cela
875	1184	7	leur	237	441	7	esprit	9	14	9	chose
455	590	7	main	28	28	7	marquis	11	22	9	homme
164	183	7	peuple	1791	4262	7	me	98	742	9	on
230	273	7	trône	27	27	7	paris	15	35	9	voilà
157	179	7	victoire	38	43	7	portrait	37	186	8	hé
738	1019	6	aux	28	29	7	rire	8	14	8	partie
209	257	6	craindre	1325	3170	6	!	7	13	7	maison
								48	321	7	père

Figure 45. Vocabulaire spécifique des différents genres

L'examen de ces listes montre que tragédies et comédies se distinguent par les modalités de l'interlocution et par les univers de référence.

Les noms des personnages, qui n'apparaissent pas ci-dessus, sont les éléments qui distinguent le plus une pièce d’une autre. Ils sont de bons marqueurs lexicaux, puisqu’ils apparaissent souvent. Sans doute pour faciliter la compréhension des spectateurs, l’auteur se soucie-t-il de rappeler régulièrement le nom de ses héros. Nous aurons l’occasion de revenir sur ce point.

A ce niveau apparaît crûment l’une des différences majeures entre comédie et tragédie : les écarts de registre. Pour Racine, le simple examen des noms de

personnages permet d'identifier le genre de la pièce. Dans la tragédie racinienne, les noms de personnages, venus de la mythologie ou de l'histoire ancienne, sont savants et rares (*Néron*, *Porus*, *Achille*, *Titus*, *Pyrrhus*, *Roxane*...). Même les confidents ont des noms aux sonorités gréco-latines (*Phénice*, *Phœnix*, *Œnone*...).

Dans l'unique comédie, les noms de personnages ont des connotations toutes différentes, bien plus communes (*Chicanneau*, *L'intimé*, *Léandre*, *Dandin*, *Petit-Jean*, *Isabelle*).

Pour Corneille, en revanche, les noms des personnages ne semblent pas être autant marqueurs de genre :

- Tragédies : Jason, Camille, Marcelle, Pauline, Placide, Attila, Cinna, Cléopâtre, Eurydice, Médée, Œdipe, Othon, Phocas, Polyeucte, Pompée, Rodrigue, Sertorius, Spitridate.
- Comédies : Dorante, Cliton, Clarice, Lyse, Philiste, Isabelle, Daphnis, Florame, Hippolyte, Lysandre, Tircis, Alcidon, Cloris, Dorimant...

Par ailleurs, dans l'interlocution, dans tous les éléments destinés à produire un certain effet sur le récepteur (fonction conative du langage, [Jakobson, 1963]), apparaissent d'autres marqueurs de l'orientation vers le destinataire qui distinguent les deux genres : les titres utilisés pour s'interpeller, et le vouvoiement/tutoiement. Chez Racine, les termes *Seigneur* et *Madame* sont presque exclusivement présents dans la tragédie. Dans la comédie, les personnages s'interpellent plutôt par leurs prénoms ou par *monsieur*. Chez Corneille, la ligne de partage entre comédie et tragédie se fait autour de *Seigneur*, *Prince* / *Monsieur* ou patronyme. Enfin, dans la comédie, on se tutoie beaucoup plus facilement que dans la tragédie.

Les types d'appellation propres aux deux genres traduisent la différence entre les univers sociaux : d'un côté les sommets du pouvoir royal ou princier (registre élevé), de l'autre le monde des bourgeois (registre commun). Et la situation sociale globale, caractéristique du genre, a un effet sur l'ensemble du personnel de la pièce. Ainsi, le registre de langage de la comtesse dans la comédie ne sera pas sensiblement différent de celui de *Chicanneau* ou de *Petit-Jean* et inversement, dans la tragédie, les héros comme les suivants emploieront le même registre. Tout se passe comme si le genre, parce qu'il impose une uniformité de registre de langue, par-delà les différences de position sociale, uniformisait les manières de parler des personnages.

Ensuite, le lieu et l'époque où se déroule l'action sont caractéristiques des genres. L'action dans les tragédies est lointaine aussi bien dans le temps que dans l'espace, la distance spatio-temporelle étant garante de la noblesse du genre. Les comédies se situent le plus souvent à l'époque et dans la ville où elles sont représentées. Ainsi, le nom de ville le plus spécifique des tragédies de Corneille est-il *Rome* (351 des 398 occurrences de *Rome* apparaissent dans les tragédies), tandis que celui des comédies est *Paris* (les 27 occurrences de *Paris* appartiennent aux comédies).

Il est à noter que ces spécificités ne sont pas *a priori* nécessaires à l'intrigue. Ce sont des traits qui activent un ensemble de connotations qui renforcent la tonalité de genre : grandeur pour la tragédie, familiarité pour la comédie.

L'examen avec Hyperbase des mots caractéristiques montre chez Corneille comme chez Racine l'opposition nette (et somme toute prévisible) des thématiques référentielles entre les univers tragique et comique.

Alors que dans la tragédie, il est question de pouvoir (*roi*, *seigneur*, *prince*, etc.), de *gloire*, de *haine*, de *vengeance* et de *mort*, dans la comédie il est question d'*humeur*, d'*amours* (pluriel qui rend trivial le sentiment), de *discours* et d'*affaire*. D'un côté le sérieux, la noblesse des sentiments et la constance, de l'autre le jeu, le rire et la légèreté. Tels sont les éléments qui sont *exclusivement* spécifiques de chacun des genres. Les marques exclamatives, ponctuations (!, ?) et interjections (*hé*), accentuent les différences de ton. Ce point est d'autant plus notable que les marques de l'interrogation sont aussi très fréquentes dans la tragédie (*cf. infra*).

Comme le montrent les marqueurs lexicaux, la tragédie est plutôt du côté du sentiment, de la délibération intérieure, elle est là où s'affrontent des sentiments contradictoires (amour et honneur...). La comédie, ancrée dans les préoccupations matérielles, tourne autour d'affaires, d'argent... L'ancrage dans la réalité sociale y est beaucoup plus marqué.

VIII.1.2.3. Rythmes selon les genres

L'analyse de la figure des alexandrins selon les genres et les auteurs met également à jour des différences stylistiques importantes. Sur l'ensemble des

pièces de Corneille[163], le vers de la tragédie n'est pas très différent de celui de la comédie (*cf.* Tableau 37).

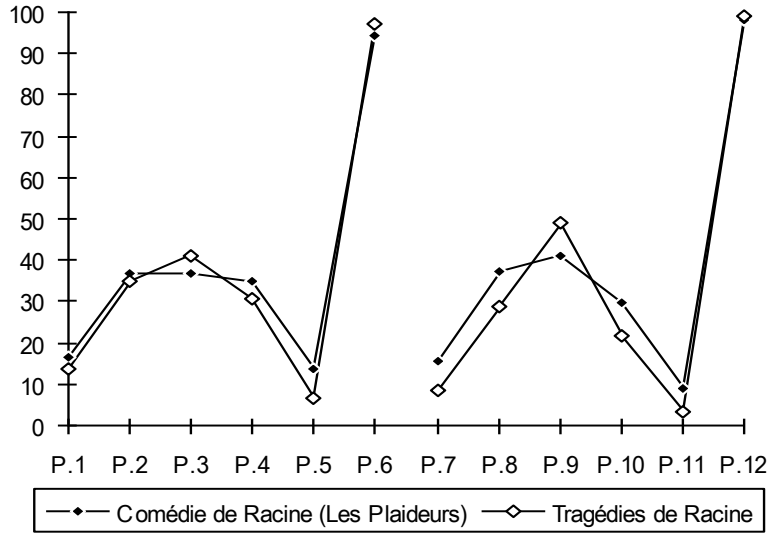
	Tragédies de Corneille	Tragédies de Racine	Comédies de Corneille	Pièces diverses de Corneille	Comédie de Racine	Ensemble
H1_001001	23	**26**	23	23	16	24
H1_010101	37	**39**	36	36	34	37
H1_autre	**40**	35	**41**	**41**	**50**	39
Total	100	100	100	100	100	100
H2_001001	30	**39**	29	30	24	32
H2_010101	**39**	36	39	**40**	36	39
H2_autre	**31**	25	**32**	30	**40**	30
Total	100	100	100	100	100	100

Clef de lecture : 23 % des premiers hémistiches des tragédies de Corneille sont de type H_001001, contre 24 % en moyenne.

Tableau 37. Types d'hémistiches selon les genres et les auteurs

Racine, au contraire, maximise les écarts entre tragédie et comédie, en utilisant dans *Les Plaideurs* un taux de vers « irréguliers » particulièrement élevé (50 % de H1_autres dans sa comédie, contre 35 % dans ses tragédies). La Figure 46 présente les courbes correspondant au marquage accentuel des tragédies et de la comédie de Racine. Les différences les plus fortes apparaissent sur les positions 3 des hémistiches, où l'accentuation est plus systématique (au détriment des positions 2 et 4) dans les tragédies que dans la comédie. L'autre différence rythmique entre genres tient à l'écart entre les cinquième et sixième positions qui est plus important dans les tragédies de Racine que dans sa comédie.

[163] Muller [1967, 1993] a bien montré que l'écriture de Corneille avait évolué, et qu'en particulier le nombre de mots par vers avait sensiblement augmenté au fil du temps. En comparant globalement les comédies et les tragédies, on masque donc des effets de période. Les profils métrico-rythmiques des deux genres seraient sans doute plus différenciés, si des sous-périodes avaient été prises en compte.



Clef de lecture : 16 % des vers des *Plaideurs* ont une position 7 accentuée, contre 9 % des vers de tragédie.

Figure 46. Tragédies versus comédie chez Racine

Quant aux différences d'auteurs, le vers des tragédies de Racine emploie plus fréquemment les modèles rythmiques les plus réguliers (001001 et 010101) : il n'y a que 25 % de seconds hémistiches « irréguliers », contre 31 % chez Corneille. Mais c'est principalement par l'usage de la figure H_001001 que Racine se distingue de Corneille (39 % de H2_001001 chez Racine contre 30 % chez Corneille).

La répartition des trois catégories d'hémistiches (H_001001, H_010101, H_autres) selon les pièces (*cf.* Figure 47) confirme ces résultats. Il apparaît que les comédies de Corneille ont un taux d'hémistiches irréguliers (autres) à peine plus élevé que les tragédies. Le contraste entre le rythme des tragédies de Racine et celle de Corneille, pour les pièces contemporaines, apparaît clairement sur le graphique et plus encore le contraste entre les tragédies de Racine et *Les Plaideurs*, comédie bien plus irrégulière que toutes celles de Corneille.

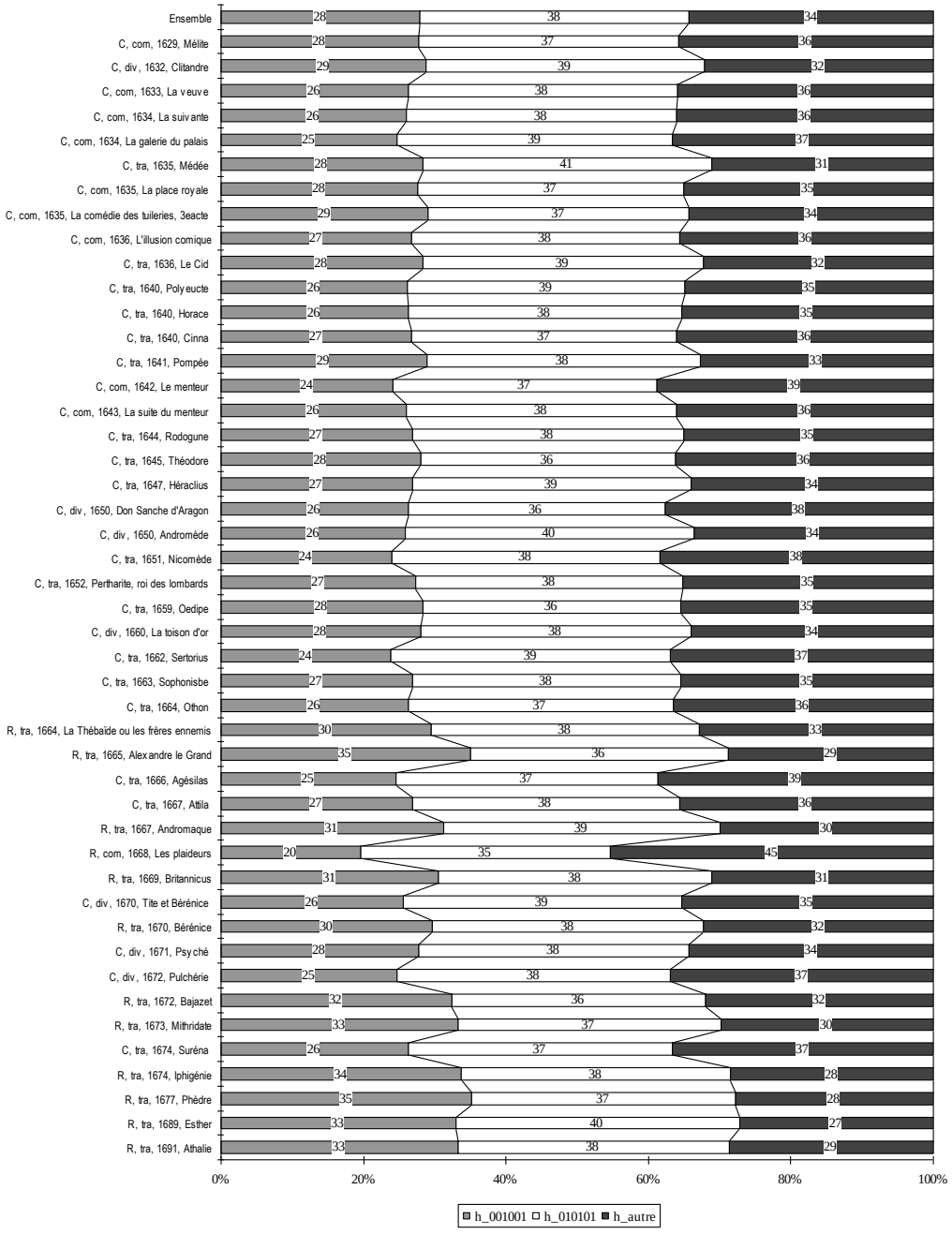


Figure 47. Répartition des types d'hémistiches selon les pièces

En partant sans *a priori* de l'ensemble des pièces pour construire des univers de discours et mettre au jour des liens avec les types de rythmes, nous

avons montré que chaque univers tendait à être en relation avec un genre ou un auteur spécifique, ce qui a conduit à reconsidérer la notion de genre et d'auteur d'un point de vue stylistique. Ainsi, nous venons de voir que la longueur des mots, le type de vocabulaire et le rythme des vers (surtout pour Racine) constituaient des indices stylistiques pertinents pour distinguer les auteurs et les genres.

VIII.2. D'*ANDROMAQUE* A *PHEDRE* : DES TRAGEDIES SEMBLABLES

Dans l'analyse globale précédente, les effets stylistiques de genre et d'auteur brouillent en partie l'identification des liens entre univers de discours et rythme. Nous allons maintenant nous concentrer sur un corpus affranchi de ces limites en choisissant un auteur et un genre : la tragédie chez Racine, et en sélectionnant parmi ces tragédies celles qui se ressemblent le plus.

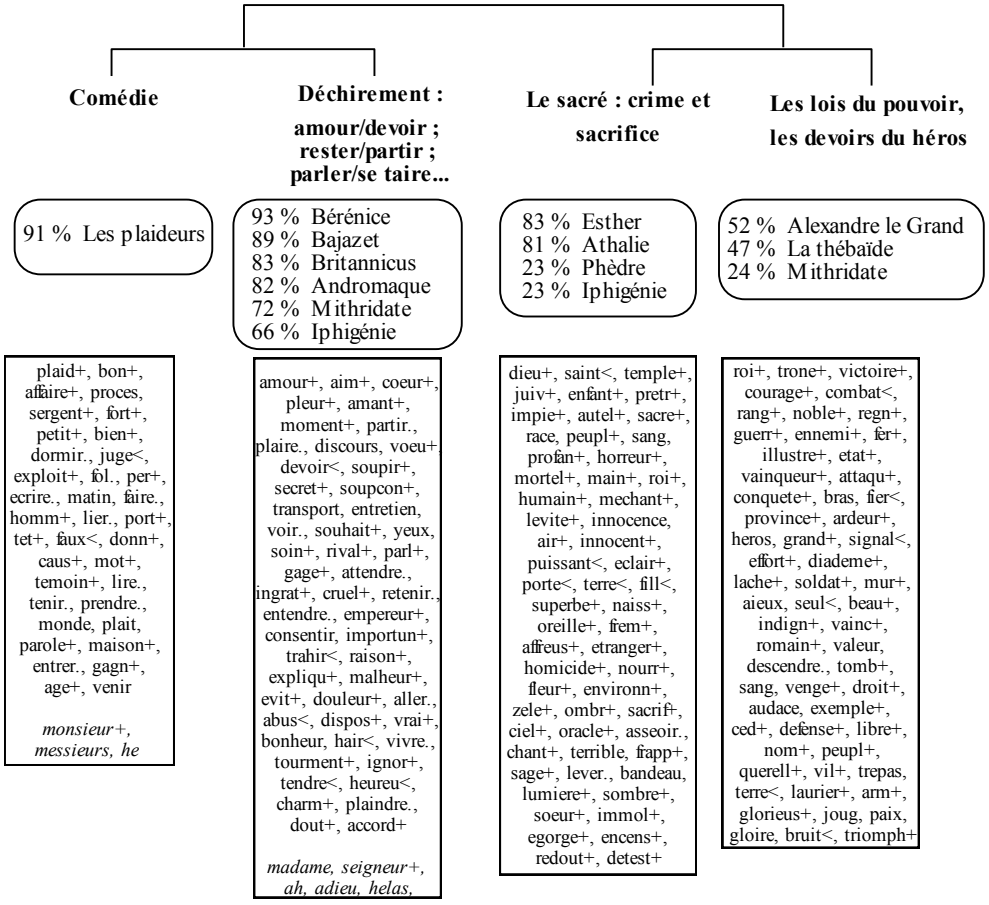
Une première analyse sur les pièces de Racine montre que les tragédies *d'Andromaque* à *Phèdre* ont des profils très similaires. Nous examinerons donc les contenus de ces pièces centrales et nous les croiserons avec les données sur le rythme.

VIII.2.1. Une grande proximité entre les tragédies centrales

Nous avons effectué avec Alceste une analyse des œuvres de Racine qui porte exclusivement sur les vers, sur les paroles effectivement prononcées sur scène. Les mots pleins (noms, verbes, adjectifs et certains adverbes) ont été retenus, à l'exclusion des noms de personnages, de lieux, et de peuples. En effet, nous souhaitons identifier les thématiques des pièces par delà les phénomènes d'actualisation que représentent l'ancrage en un lieu spécifique et les noms donnés aux personnages.

On effectue pour ce faire une classification descendante sur un tableau qui croise les mots lexicaux lemmatisés et des fragments des pièces (unités de contexte) qui comprennent au moins 20 mots actifs. Cette analyse permet de

classer 80 % des unités textuelles dans 4 classes. Le graphique ci-dessous présente pour chaque classe le vocabulaire spécifique et les pièces qui y sont le plus représentées. Les noms qui ont été donnés aux classes relèvent d'un travail d'interprétation par sélection des traits sémantiques qui paraissent le mieux représenter la classe.



Clef de lecture : les « mots » analysés par Alceste ne portent plus d'accent et sont réduits à leur racine (présence d'un signe +, <) selon différents algorithmes.

Figure 48. Champs lexicaux dans les œuvres de Racine

Le graphique suivant présente les profils lexico-sémantiques des pièces : 42 % des fragments de *La Thébaïde* ont été classés dans la classe "Déchirement", 48 % dans la classe "Devoir et pouvoir" et 10 % dans la classe "Sacré".

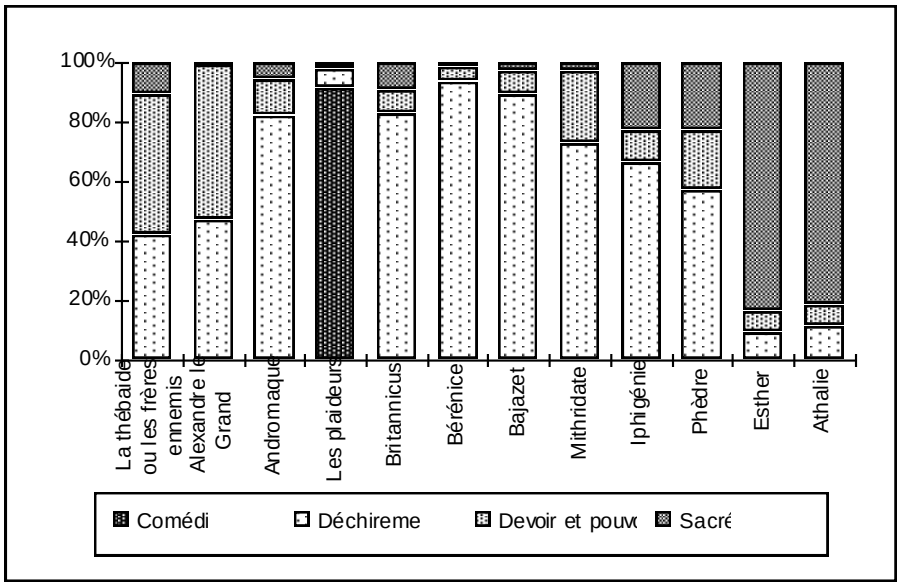


Figure 49. Profils lexico-sémantiques des œuvres de Racine

Les profils lexico-sémantiques permettent d'identifier les ressemblances et différences entre pièces et mettent en évidence deux résultats :

- comédie et tragédie ont des univers sémantiques disjoints ;
- il existe trois périodes dans l'écriture tragique de Racine ;
- le passage d'une période à une autre est amorcé progressivement.

La comédie se distingue très nettement des tragédies : la comédie et la tragédie ne partagent absolument pas les mêmes "mondes lexicaux". Elles n'ont rien de commun dans l'organisation du vocabulaire.

Les tragédies, si l'on s'en tient à leurs profils lexicaux, se répartissent en trois grands ensembles de pièces qui correspondent à trois périodes de l'écriture de Racine. L'un regroupe *La Thébaïde* et *Alexandre le Grand*, un autre toutes les tragédies d'*Andromaque* à *Phèdre* (*Bajazet* occupe une place spéciale dans cet ensemble) et enfin un troisième ensemble regroupe *Esther* et *Athalie*. Cette classification effectuée à partir des données lexicométriques met en évidence l'évolution du style de Racine et de ses sources d'inspiration. Les deux premières tragédies sont encore très marquées par l'influence de Corneille

[Bénichou, 1948 ; Benzécri, 1981] qui jouissait alors d'une grande notoriété[164]. Ce n'est qu'à partir d'*Andromaque* que s'affirme avec le plus de vigueur une vision du monde et du tragique très personnelle, que l'on retrouve jusqu'à *Phèdre*, qui clôt la seconde période. Enfin, après un silence de douze années, Racine écrit sur commande ses deux dernières tragédies où la source d'inspiration n'est plus l'histoire antique grecque ou romaine (voire ottomane) mais l'écriture sainte. Cette classification recouvre d'une manière parfaite l'évolution chronologique et recouvre entre autres la distinction entre les tragédies profanes et les tragédies sacrées.

Remarquons aussi que les transitions d'une période à l'autre se font sous le signe de la continuité plus que de la rupture. Le passage aux tragédies sacrées est ainsi amorcé dès *Iphigénie*. On voit en effet apparaître dès cette pièce la thématique du sacré qui s'affirme dans *Phèdre* avant de culminer dans les deux dernières tragédies. Notons enfin que *Phèdre* a le profil « le plus équilibré » : chacune des thématiques tragiques y est assez bien représentée.

Il est intéressant de constater que les tragédies d'*Andromaque* à *Phèdre* ont globalement un profil très proche, ce qui marque une grande homogénéité dans l'écriture et dans la trame tragique, en dépit du fait que les actions se déroulent en des lieux très distincts (Constantinople, Rome, Trézène...) et ne racontent pas les mêmes histoires.

VIII.2.2. L'onomastique comme marqueur stylistique

Les pièces centrales de Racine, d'*Andromaque* à *Phèdre*, celles qui ont le plus contribué à sa gloire, ont donc un profil lexical extrêmement proche, une fois que l'on ne tient plus compte des marqueurs de lieux et des noms de personnages.

C'est un peu comme si chacune des tragédies était la réécriture d'un même modèle, dont le principe de variation serait dicté par l'onomastique. Examinons les facteurs de cette variation.

---

[164] L'analyse globale a déjà permis de mettre ce point en évidence.

Premièrement, les mots les plus spécifiques de chaque pièce correspondent aux noms de lieux (qui sont souvent des actants comme Rome dans *Bérénice*). Ainsi, les mots les plus caractéristiques de *Bérénice* relativement aux autres tragédies (repérés par un module d'Hyperbase [Brunet, 1989]) sont : *Titus*, *Bérénice*, *Paulin*, *Rome*, *Arsace*, *Phénice*, *reine*, *Sénat*, *Orient*, *adieux*, *je*, *déclarer*, *moment*, *partir*, *pleurs*, *adieu*, *amoureux*, *constance*, *moments*, *prince*, *seigneur*, *séparer*[165]. Certes, ces noms n'apparaissent que dans la pièce qui les met en scène (jamais un nom n'est repris d'une pièce à l'autre) et c'est l'une des raisons qui les rend spécifiques d'une pièce donnée. Mais, surtout, ces noms reviennent régulièrement au fil des vers prononcés, plus fréquemment que ne l'imposerait une situation de dialogue « ordinaire ». La situation de représentation théâtrale implique en effet que les noms de lieux et de personnages soient prononcés, en particulier dans les scènes d'exposition, pour que l'ancrage de la pièce et la structure du monde narratif soient compris et intégrés dans l'esprit du public. Toutefois, chez Racine, l'emploi des noms propres dépasse largement cette visée pratique et devient un trait stylistique. Chez Corneille par exemple, la fréquence d'utilisation des noms propres est moindre.

Leo Spitzer [*1931*, 1982] a montré comment l'emploi du nom propre à la place du pronom personnel contribuait à établir un « effet de sourdine », une mise à distance de l'émotion dans l'énonciation. Ainsi, Bérénice s'adressant à Titus dira :

> Dans un mois, dans un an, comment souffrirons-***nous***,
> Seigneur, que tant de mers <u>me</u> séparent de **vous** ?
> Que le jour recommence, et que le jour finisse,
> Sans que jamais **Titus** puisse voir <u>Bérénice</u>,
> Sans que de tout le jour <u>je</u> puisse voir **Titus** ?
>
> Racine, *Bérénice*, vers 1113-1117.

L'usage du nom propre est aussi un moyen de créer, ou plutôt de cristalliser, le mythe autour du nom.

Ce n'est pas seulement pour donner consistance à ces « êtres de papier » que sont les personnages à l'ouverture d'une pièce (ils sont déjà largement

[165] On trouvera dans l'annexe X une liste complète des mots les plus spécifiques de chacune des pièces.

chargés d'histoire), ni pour maintenir une distance ou donner corps au mythe que les noms de personnages reviennent régulièrement. C'est aussi une des composantes rythmiques des pièces. Nous allons voir que les sonorités finales de ces noms constituent, à la rime, la matière sonore spécifique de chaque pièce.

Le travail fait sur la rime (*cf.* chapitre III) a permis de déterminer empiriquement à partir de notre corpus des groupes de mots-rimes. Chaque groupe contient des mots pouvant rimer ensemble, qui partagent le même rimème (la même terminaison). Nous avons ensuite affecté à chaque vers le groupe de mots-rimes qui lui correspondait. Ainsi, le vers

> En quelles mains je dois confier ce trésor,
>
> Racine, *Britannicus*, vers 578.

relève-t-il du groupe de mots-rimes en *or*, qui contient les mots : *Alidor*(4), *Hector*(10), *Médor*(1), *Nicanor*(1), *Rosidor*(2), *encor*(20), *mogor*(1), *or*(3), *trésor*(6). Sur la base de ce classement de toutes les terminaisons de vers, nous avons, à l'aide d'une procédure statistique développée par Christian Derquenne [1992], recherché les groupes de mots-rimes les plus spécifiques de chaque tragédie. Il apparaît d'une manière très claire que les rimèmes qui correspondent aux noms des héros sont significativement plus présents dans chaque pièce (*cf.* la liste des rimèmes spécifiques de chaque pièce dans l'annexe X). Ainsi, pour reprendre l'exemple de *Bérénice*, il y a significativement plus de vers finissant en *is[ə]* dans *Bérénice* que dans les autres tragédies. En général, les personnages principaux de chaque pièce, sauf Phèdre qui n'apparaît jamais à la rime, voient la sonorité finale de leur nom devenir, dans leur pièce, un trait distinctif de la rime. Les rimèmes les plus caractéristiques d'*Andromaque* par ordre de significativité décroissante sont : *or* (*Hector*) ; *one* (*Hermione*, *Cléone*) ; *oie* ; *este* (*Oreste*) ; *aque* (*Andromaque*) ; *vé* ; *is* ; *armes* ; *ade* (*Pylade*) ; *arde* ; *us* (*Pyrrhus*) ; *ui*[166]. Parmi les rimèmes les plus spécifiques de chaque pièce, nous retrouvons ceux qui correspondent aux noms des personnages. Les sonorités qui distinguent chaque pièce correspondent aux finales des noms de personnages. Cela tient à la fréquence d'apparition élevée des noms, mais aussi à la recherche d'une qualité

[166] On a fait figurer entre parenthèses les noms des personnages correspondant au rimème.

phonétique particulière. Ainsi même lorsque l'on ne tient pas compte, dans les calculs, des vers qui contiennent les noms de personnage à la rime, les rimèmes qui correspondent aux noms continuent d'être les plus spécifiques de chaque pièce (sauf dans *Bérénice*).

Le choix des noms par Racine ne relève pas non plus du hasard. Si le nom des héros est imposé par l'histoire, il n'en est pas de même pour les confidents. Or dans quasiment chaque pièce, le nom d'un confident rime avec celui du héros : ainsi en est-il *d'Hermione* et de *Cléone*, de *Bérénice* et de *Phénice*, de *Mithridate* et *d'Arbate*, de *Monime* et de *Phaedime*... Racine évite la facilité qui consisterait à faire rimer deux vers successifs avec cette paire de noms, mais le fait que les confidents aient une sonorité proche de celle des héros renforce la coloration sonore de la pièce.

Il ne s'agit pas d'un effet de sonorité gratuit ; cette composante rythmique renvoie à la sémantique des personnages : le confident est un double du personnage et cette ressemblance est dite aussi par la parenté phonétique des noms. Ces figures du double nous disent que héros et confident sont les deux faces d'une même unité. Le dialogue entre héros et confident permet de rendre explicite, sous forme rhétorique, ce qui ne serait autrement qu'un monologue intérieur du héros[167].

VIII.2.3. Rythme et parties du discours

Dans la première section, nous avons vu la similarité des profils des tragédies centrales de Racine. Identifions maintenant sur cet ensemble homogène les univers de discours. Cette fois, retenons comme unité d'analyse le couple de vers rimant ensemble (nous conserverons uniquement les alexandrins à rimes plates, soit presque l'intégralité des pièces). Nous aurons ainsi considérablement réduit la taille des contextes soumis à l'analyse (deux

[167] C'est à la suite d'une représentation de *Bérénice*, mise en scène par Daniel Mesguish, que nous avons exploré la question des noms. Toute la mise en scène était articulée autour du double : Bérénice et Phénice (tout comme Titus et Paulin) portaient les mêmes costumes, se mouvaient comme si elles (ils) étaient le miroir l'un(e) de l'autre. Costumes, mouvements, jeux de miroirs, écho sonore étaient autant de métaphores du dédoublement.

vers au lieu de dix dans l'analyse globale). Les mêmes catégories de mots sont retenues pour l'analyse.

Le traitement statistique appliqué à ces pièces distingue dans un premier temps deux grands groupes (59 % et 41 %). Comme dans l'analyse globale, et pourtant il ne s'agit ici que de quelques tragédies de Racine et non pas de toutes les pièces, le premier groupe est dominé par le sème /amour/ et le second par le sème /mort/. Ce résultat nous montre, s'il était besoin, que la tragédie mobilise ces deux thèmes. Rappelons que dans l'analyse sur l'ensemble des pièces le pôle dominé par le sème /amour/ apparaissait très peu dans les tragédies, contrairement à ici. Nous voyons bien que le corpus choisi et la manière de le découper (séquences de 9 à 11 vers dans l'analyse globale, de 2 vers dans celle-ci) ont une incidence sur ce que l'on y observe. Finalement, dans la typologie précédente, le domaine de l'amour devait être associé à celui de la mort dans les séquences de dix vers, ce qui est moins vrai en prenant comme empan la paire de vers rimant ensemble.

Dans notre analyse des pièces centrales de Racine, les deux grandes classes ne se limitent pas à une opposition amour / mort. Elles recoupent une opposition formelle entre le dialogue (la confrontation des personnages sur scène) et le récit. Cette opposition est explicite lorsque l'on regarde la répartition dans les deux classes d'un certain nombre de mots grammaticaux qui ne sont pas intervenus dans l'analyse comme variables actives. Commençons par les pronoms personnels.

Dans la classe /amour/, les pronoms personnels de la première et seconde personnes, ceux du dialogue, sont très fréquents, tandis que dans la seconde, domine la troisième personne. Les marques de la fonction conative, de l'adresse à l'autre, sont beaucoup plus fréquentes dans la première classe : *Madame*, *Seigneur* et les formes de l'impératif. Dans la seconde, les noms propres l'emportent. Concernant la question de l'espace, la première classe renvoie à un ici et maintenant de la scène, tandis que la seconde évoque des espaces hors-scène, qui ne peuvent être rendus présents sur scène que par le récit. Tous ces éléments marquent clairement la distinction entre les formes du dialogue et du récit (même inséré dans un monologue ou dialogue). Les verbes modaux, les marqueurs de modalisation et les interjections sont caractéristiques de la classe *dialogue-amour*.

L'examen des mots pleins spécifiques de chaque classe renforce l'opposition entre la logique du dialogue et celle du récit. D'un côté, dans la classe /amour/, le vocabulaire, dominé par la dimension psychologique, renvoie à une mise en scène dans le langage des difficultés de la relation, dans ses aspects de communication entre les êtres. Au cœur de cette classe dominent des dilemmes raciniens sur la parole, se taire ou parler, sur la présence, rester ou partir, et sur la vie, vivre ou mourir. Ces interrogations se font sur un fond de pleurs et de douleur. De l'autre, le vocabulaire renvoie à des univers extérieurs à la scène (décor naturel, monde des dieux et univers des lignées de pouvoir), souvent à des événements passés ou à venir, hors du temps de la tragédie.

Chacune de ces classes (dialogue-amour et récit-mort) est effectivement caractérisée par des rythmes spécifiques ($P < 1.10^{-5}$ avec le test du chi-deux). Dans le pôle du récit, les formes régulières d'hémistiches sont plus fréquentes que dans celui du dialogue. Les différences sont ici plus ténues que dans l'analyse globale lorsque nous comparions des univers de discours propres à la comédie et à la tragédie.

	dialogue-amour (59 %)	**récit-mort (41 %)**	**Ensemble**
H_001001	31%	34%	32%
H_010101	37%	38%	37%
H_autres	31%	29%	30%
Total	100%	100%	100%

	Dialogue-amour			**Récit-mort**			
	Classe 3	Classe 6	Classe 5	Classe 4	Classe 2	Classe 1	total
H_001001	30%	32%	33%	33%	34%	34%	32%
H_010101	37%	37%	37%	37%	38%	39%	37%
H_autres	33%	31%	30%	30%	29%	27%	30%
Total							

Tableau 38. Répartition des types d'hémistiches selon les univers de discours

Nous pouvons détailler les contenus de ces deux grandes classes. En poursuivant l'analyse, chacune conduit à trois sous-classes comme on peut le voir dans le tableau précédent. Pour le pôle dialogue-amour, ce découpage en sous-classes n'introduit pas d'informations particulières qui n'aient été vues

dans la grande classe. Toutes les formes de délibérations (aimer/haïr, rester/partir[168], vivre/mourir, se taire/parler), liées à la confrontation des personnages sur scène sont visibles.

h1_010101 h2_010101	Je n'examine point ma joie ou mon ennui :
h3_autre h4_autre	J'aime assez mon amant pour renoncer à lui.
	Racine, *Bajazet*, vers 835.
h1_autre h2_010101	Adieu. / Prince... / On m'attend, madame, il faut partir.
h3_001001 h4_001001	Mais du moins attendez qu'on vous vienne avertir.
	Racine, *Britannicus*, vers 1561.
h1_autre h2_010101	Quoi ? vous ne perdrez point cette cruelle envie ?
h3_autre h4_001001	Vous verrai-je toujours, renonçant à la vie,
	Racine, *Phèdre*, vers 173.
h1_010101 h2_autre	Je vois que la raison cède à la violence.
h3_autre h4_010101	Puisque j'ai commencé de rompre le silence,
	Racine, *Phèdre*, vers 525.

La dramatisation du surgissement de la parole est une constante des tragédies de Racine : l'aveu se fait attendre, il est fait malgré le personnage et la parole à peine prononcée est déjà regrettée, car elle signe la culpabilité, la faute tragique. La parole de l'aveu est prise dans un réseau d'interdits, qui la rendent problématique. Mais la parole permet aussi l'expression de sentiments feints, permet de tromper :

h1_010101 h2_autre	Combien tout ce qu'on dit est loin de ce qu'on pense !
h3_001001 h4_autre	Que la bouche et le cœur sont peu d'intelligence !
	Racine, *Britannicus*, vers 1523.

Seul le regard permet de révéler la vérité cachée :

h1_010101 h2_001001	Vous n'aurez point pour moi de langages secrets :
h3_001001 h4_010101	J'entendrai des regards que vous croirez muets ;
	Racine, *Britannicus*, vers 681.

Toutes ces scènes de dialogue, de confrontation entre les personnages, se font dans un contexte de douleur et de plainte. La souffrance est mise en scène à travers les références aux larmes, aux pleurs et à la douleur. La plainte est, comme l'écrivait Bénichou [1948], « l'accompagnement de la cruauté », cruauté de l'amour racinien marqué par une agressivité violente à l'égard de l'objet désiré.

[168] C'est aussi une manière de permettre l'organisation des entrées-sorties sur scène.

h1_010101 h2_010101	Quels charmes ont pour vous des yeux infortunés
h3_autre h4_autre	Qu'à des pleurs éternels vous avez condamnés ?

Racine, *Andromaque*, vers 303.

Dans ces scènes de confrontation, la nostalgie d'un passé heureux vient en contrepoint à un présent de souffrance. L'évocation du bonheur, de temps heureux rend plus insupportable le présent.

h1_001001 h2_autre	Ils regrettent le temps, à leur grand cœur si doux,
h3_010101 h4_010101	Lorsque assurés de vaincre ils combattaient sous vous.

Racine, *Bajazet*, vers 47.

La confrontation des personnages sur scène et le déchirement intérieur se traduisent rythmiquement par une proportion d'hémistiches non réguliers plus élevée. Il y a bien dans ce cas une forme d'homologie entre ce qui se dit sur scène et le rythme des vers : le déchirement, la détresse s'expriment dans un vers plus haché, moins structuré. Par ailleurs, la dramatisation de la confrontation des personnages se traduit aussi par des changements de tour de parole accélérés qui entraînent parfois qu'un même vers soit prononcé par plusieurs personnages. Or, sur le corpus de vers dits par plusieurs personnages (3 % des alexandrins), le taux d'hémistiches irréguliers s'élève à 44 %, contre 30 % pour les vers « entiers ».

Revenons à la classe récit-mort. Elle se décompose en trois. La première sous-classe correspond aux moments où l'histoire des personnages est mise en perspective, où les interdits sont énoncés. Il y est question de pouvoir et d'alliance.

h1_autre h2_010101	Voilà par quels exploits il sut se couronner ;
h3_010101 h4_010101	Enfin voilà l'époux que tu me veux donner.

Racine, *Andromaque*, vers 1007.

h1_010101 h2_010101	Mon père la réprouve ; et par des lois sévères
h3_001001 h4_001001	Il défend de donner des neveux à ses frères :

Racine, *Phèdre*, vers 105.

La seconde tourne autour de la mort, des meurtres et des sacrifices, qui se passent ou se sont passés hors scène et qui sont relatés sur scène :

h1_010101 h2_001001 Tandis qu'un autre enfant, arraché de ses bras,
h3_001001 h4_001001 Sous le nom de son fils fut conduit au trépas.
Racine, *Andromaque*, vers 75.

h1_010101 h2_010101 Que vous marchiez au camp, conduit sous mes auspices,
h3_010101 h4_010101 Dans Rome les autels fumaient de sacrifices ;
Racine, *Britannicus*, vers 1187.

La dernière, évocation de l'ailleurs, retiendra notre attention. La plupart des mots spécifiques évoquent des espaces hors de la scène : bords, rives, mer, vent, vaisseaux, voiles, ports, climat, forêts... qui s'opposent par leur ouverture à l'espace clos de la scène. Cette représentation stylisée des paysages a des fonctions dramatiques évidentes. Elle introduit des moments d'apaisement et de respiration au sein de la tragédie, elle permet aux personnages d'échapper par la parole à leur situation sans issue, même si c'est pour y revenir aussitôt. Elle contribue à l'enrichissement des figures de la fuite dans les tragédies. Ainsi Phèdre regrette-t-elle bien vite les trois vers qu'elle prononce :

PHEDRE.
h3_autre h4_010101 Dieux ! que ne suis-je assise à l'ombre des forêts !
h1_001001 h2_001001 Quand pourrai-je, au travers d'une noble poussière,
h3_autre h4_010101 Suivre de l'œil un char fuyant dans la carrière ?
ŒNONE.
h1_001001 h2_010101 Quoi, madame ? / Insensée, où suis-je ? et qu'ai-je dit ?
PHEDRE.
h3_001001 h4_010101 Où laissé-je égarer mes vœux et mon esprit ?
Racine, *Phèdre*, vers 176.

h1_010101 h2_010101 Déjà dans les vaisseaux la voile se déploie,
h3_010101 h4_001001 Déjà sur sa parole ils se tournent vers Troie.
Racine, *Iphigénie*, vers 841-842.

h1_010101 h2_010101 Moi-même, il m'enferma dans des cavernes sombres,
h3_autre h4_001001 Lieux profonds, et voisins de l'empire des ombres.
Racine, *Phèdre*, vers 965-966.

Le rythme des vers qui évoquent cet ailleurs est particulièrement régulier et les hémistiches en 001001 y sont particulièrement bien représentés : cette régularité est cohérente avec l'effet d'apaisement que ces évocations produisent sur les personnages, très éphémère il est vrai.

Dans cette évocation de l'ailleurs, nous retrouvons classée une hypotypose fameuse, celle qui correspond au récit de la mort d'Hippolyte par Théramène où les éléments se déchaînent contre Hippolyte. L'évocation de l'ailleurs est

finalement assez ambivalente chez Racine, puisqu'elle est souvent associée à une menace.

VIII.2.4. Rythme des pièces

Toute tragédie est à l'époque classique composée de cinq actes, elle s'ouvre sur une scène d'exposition et se conclue par un dénouement. On s'est demandé si le rythme des vers ne contribuait pas à renforcer d'une manière ou d'une autre cette structure de la tragédie. Globalement, si l'on considère le taux d'hémistiches réguliers (001001 et 010101) par acte l'allure de la courbe obtenue s'apparente à un W chez Racine, ce qui n'apparaît pas chez Corneille. Le premier acte est le plus régulier (cela est vrai pour toutes les tragédies, y compris celles de Corneille). Dans le second, le drame prend de l'ampleur et le taux de vers réguliers diminue. Le troisième acte propose une forme de pré-dénouement qui se traduit par une remontée du taux d'hémistiches réguliers. Au quatrième acte, la tension tragique reprend toute son ampleur (moins de vers réguliers) avant d'en arriver à l'acte final, où ce taux a presque retrouvé son niveau initial. Les variations sont faibles, mais statistiquement significatives.

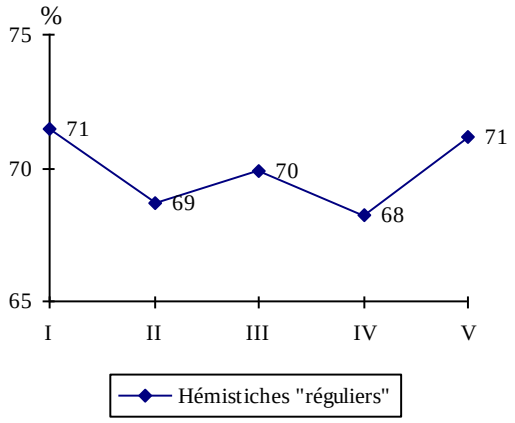


Figure 50 : Taux d'hémistiches réguliers par acte chez Racine

Malheureusement, si l'on descend dans le détail de chaque pièce, seules trois d'entre elles suivent exactement le modèle en W, *Andromaque*, *Bajazet* et

Phèdre. Il serait cependant intéressant d'explorer d'une manière plus fine la manière dont alternent dans les pièces des passages plus ou moins réguliers en termes de rythmes, ce qui montrerait l'existence d'une structure rythmique au niveau de la pièce.

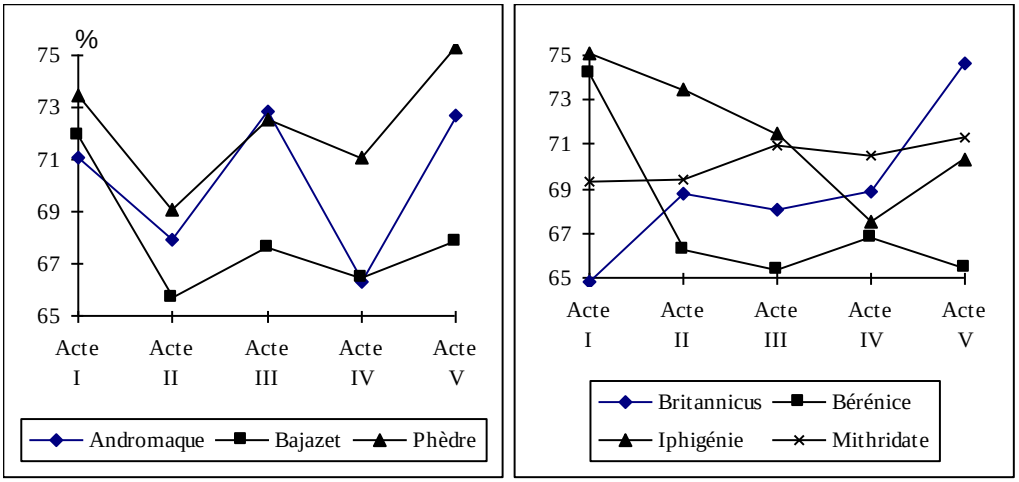


Figure 51 : Taux d'hémistiches réguliers par acte dans les tragédies de Racine

Par ailleurs, la récurrence et l'entrelacement des univers de discours au fil du texte constituent un mode d'organisation rythmique spécifique, qu'il nous faudrait étudier de manière approfondie. En nous limitant à *Phèdre*, il nous a semblé que les deux grandes catégories de discours identifiées (dialogue-amour, récit-mort) alternaient d'une manière régulière au fil de la pièce. Cette observation devra être approfondie et étendue aux autres pièces.

VIII.3. CONCLUSION

Nous avons tenté de montrer, d'une part, qu'il existait des liens entre des univers de discours (c'est-à-dire des séquences de vers mobilisant un ensemble

d'isotopies) et des rythmes spécifiques et d'autre part, que les éléments rythmiques pouvaient contribuer à l'élaboration d'une forme.

Résumons ce qui relie univers de discours et rythme. La thématique de l'amour implique des hémistiches relativement plus irréguliers que celle de la mort, qui rappelons-le est souvent dans la tragédie l'issue du conflit et donc son apaisement. La confrontation des personnages sur scène, la délibération, produisent un vers irrégulier, haché, tandis que l'évocation d'un ailleurs et le récit sont accompagnés d'un vers plus régulier, majestueux. Sur des pièces homogènes en termes de contenu, nous avons pu identifier deux pôles, celui du récit et celui du dialogue, le récit étant caractérisé par un vers plus régulier que le dialogue. Nous retrouvons ici ce que posait A. Ubersfeld :

> Il y a dans le texte deux couches : l'une qui est destinée à être représentée scéniquement, l'autre qui ne renvoie qu'à un hors-scène imaginaire.
>
> [Ubersfeld, *1977*, 1982, p. 174].

Ce hors-scène imaginaire, qui implique le récit, semble être caractérisé par un vers plus régulier.

Concernant le rôle des composants rythmiques dans la constitution d'une forme et d'un style, nous avons vu que chaque genre et chaque auteur avaient une manière spécifique de traiter le vers. Racine est, par exemple, beaucoup plus régulier que Corneille. Nous avons aussi montré que la rime constituait l'une des composantes formelles importantes de la « spécialisation » des pièces de Racine, les sonorités des terminaisons des noms de personnages marquant le trait distinctif des pièces. Enfin, nous avons tenté de montrer le rôle de la structure rythmique des vers dans la construction de l'ensemble des pièces.

Les résultats peuvent paraître bien maigres au regard de l'arsenal qui a été mis en place et après les nombreuses tentatives qui ont été faites. Ils sont de plus à prendre avec prudence et ce pour plusieurs raisons.

Tout d'abord, le marquage du rythme que nous avons choisi est loin d'être parfait, le regroupement en trois catégories (H_001001, H_010101 et H_autres) un peu grossier et nous avons été un peu vite en déclarant que le taux d'hémistiches H_001001 et H_010101 était un indice de régularité du vers. Une meilleure analyse syntaxico-prosodique du vers devrait permettre de proposer un marquage rythmique plus précis.

Ensuite, les méthodes de classification des textes sur la base des mots ne sont pas encore très satisfaisantes et très stables. Nous avons mené des dizaines d'analyses en faisant varier les paramètres pour essayer de repérer ce qui était stable : parfois, comme cela a été le cas pour la classe dialogue-amour, il a été impossible d'interpréter un niveau de découpage plus fin, car ces découpages variaient trop d'une analyse à l'autre. Une annotation enrichie des corpus[169] qui ajouterait aux mots, comme suites de caractères, des informations morphologiques, syntaxiques, sémantiques… devrait permettre d'aboutir à des typologies de fragments plus solides et plus stables. Habert et Salem [1995] ont ainsi montré comment l'introduction d'étiquettes morpho-syntaxiques dans un corpus de réponses à une question ouverte permettait d'enrichir l'interprétation des données.

Enfin, se pose la grande question de la réception : le lecteur-auditeur peut-il percevoir les effets que nous avons identifiés ? Nous avons mis à jour des phénomènes statistiques : sont-ils perceptibles dans la linéarité de la lecture et de l'audition ? S'ils sont perçus, ce n'est sans doute pas en tant que phénomènes statistiques.

Ces résultats, même partiels, remettent à sa place le rôle du rythme dans le style : il s'agit là d'une composante très forte au niveau microstructural, et qui, si elle est effectivement corrélée aux aspects sémantiques, de plus large empan, n'intervient là qu'à titre de soutien.

[169] Cette approche est défendue dans *Les Linguistiques de corpus* [Habert et al., 1997].

CONCLUSION

Au terme de ce parcours, faisons le bilan des acquis et traçons quelques perspectives. Nous passerons d'abord en revue nos principaux résultats : sur l'outil, sur les formes rythmiques et sur la rime. Ensuite, nous ébaucherons une vue plus synthétique de la question du rythme et de ses rapports avec la forme littéraire.

Le métromètre

Nous disposons désormais d'un outil pour l'analyse des vers classiques, le métromètre. Celui-ci est la concrétisation du projet Dynastie d'analyse du vers classique proposé par Roubaud au début des années 80 ; il a pu voir le jour grâce aux progrès remarquables qu'a connus la linguistique informatique ces dernières décennies. Pratiquement, l'outil repose sur un phonétiseur du français contemporain qui s'appuie lui-même sur un analyseur syntaxique. Le phonétiseur a été adapté aux particularités du décompte du vers classique qui concernent principalement la diérèse, le *e* muet et la liaison. Le métromètre est donc un outil fort précieux. Rappelons que c'est la première fois qu'un outil d'analyse systématique du vers est conçu et fonctionne sur des corpus de vers. Le métromètre a permis en effet d'analyser près de 120 000 vers, dont une grande majorité d'alexandrins : les pièces de Corneille et Racine (sur lesquelles a porté l'essentiel de notre travail), *La Légende des siècles* de Hugo, *Les Fleurs du Mal* de Baudelaire, une partie des poèmes en vers de Mallarmé et de Rimbaud pour ne citer que les plus importants.

Chaque vers est transcrit dans l'alphabet phonétique international et découpé en positions métriques ; l'outil est donc capable de reconnaître tous les types de mètres dans le vers français. Un certain nombre de caractéristiques phonétiques, lexicales, syntaxiques et prosodiques est associé à chacune des

positions métriques. Chaque vers est également caractérisé par des étiquettes « paratextuelles » : qui prononce le vers, dans quel acte, dans quelle scène, de quelle pièce ? ou dans quel poème, de quelle partie, de quel recueil ?

Le travail sur la rime nous a également permis de compléter la description du vers en y ajoutant le mot-rime et le rimème d'appartenance, c'est-à-dire la séquence graphico-phonétique commune aux vers qui riment ensemble. Cette mise en forme du vers sous un format permettant l'analyse statistique n'est pas très spectaculaire. C'est cependant un apport méthodologique et technique qui devrait être utile à d'autres. C'est pourquoi nous avons fourni le détail des formats dans l'annexe IV. La recherche ne peut avancer que si les comparaisons sont possibles entre travaux, et pour cela une certaine normalisation des données est indispensable. Nous espérons y avoir contribué.

La mise au point de l'outil a obligé à préciser les règles de découpage du vers en positions métriques. Cette nécessité formelle nous a fait en passant découvrir certaines tendances intéressantes comme celle qui tend à donner à chaque position une forme creux-sommet (CV), qui subsume les règles du *e* muet, de la liaison, de la synérèse et de l'interdiction du hiatus. Elle nous a également amenée à déterminer empiriquement les noyaux vocaliques élémentaires dans le vers, qui forment un ensemble légèrement distinct des seize phonèmes vocaliques habituellement distingués.

L'outil, et la base de vers qu'il permet de construire, nous ont permis d'étudier quantitativement sur notre corpus de nombreux phénomènes phonétiques, syntaxiques et accentuels, généraux ou très particuliers. Il peuvent encore permettre de valider ou d'invalider bien des hypothèses sur le vers, sur sa forme, sur son évolution, sur ses variantes selon les auteurs, les genres, les époques, les textes...

Les formes du rythme

Le métromètre a été conçu pour analyser les phénomènes métriques et rythmiques dans le vers. Reprenons la distinction entre mètre et rythme. Le mètre désigne les structures qui se répètent systématiquement, tandis que le rythme désigne une organisation, une forme identifiable qui s'inscrit dans un mètre donné. Le modèle métrique de l'alexandrin classique, tel qu'il a été

proposé par Roubaud [*1978*, 1988] comprend plusieurs niveaux hiérarchisés. Partons du vers comme unité métrique et abordons le niveau au-dessus : deux vers unis l'un à autre par la rime constituent une unité d'organisation supérieure. Au-dessus encore, l'alternance en genre des rimes fait du quatrain de vers une unité métrique, dans un contexte d'alexandrins à rimes plates. Le vers lui-même est découpé en deux segments métriques, qui sont chacun constitués de six positions métriques.

Les vers alexandrins des pièces étudiées (dans les séquences d'alexandrins plats) sont tous conformes à ce modèle. Pour qu'il y ait mètre, il faut qu'il y ait périodicité. La position, l'hémistiche, le vers, la paire de vers, le quatrain constituent des unités métriques car ils sont constitués par la répétition d'unités plus petites : une voyelle entourée de non-voyelles pour la position, six positions pour l'hémistiche, deux hémistiches pour le vers, deux vers pour la paire de vers et deux paires pour le quatrain. Mais ces répétitions ne sont encore que du mètre, c'est-à-dire des groupements d'événements répétés à l'identique.

Nous avons montré comment la matière de la langue (son, syntaxe et accent) donne forme à ce modèle, comment elle se moule dans les positions métriques ; et réciproquement comment elle marque les différents niveaux d'organisation métrique. Pour cela, nous avons adopté l'hypothèse forte que le mètre, en dehors du nombre qui le définit, est vide de forme, c'est-à-dire qu'il n'y a pas d'autres formes que celles découlant de la répartition des constituants linguistiques sur l'échelle du mètre. Autrement dit, il n'y a pas d'accent métrique : le fait d'être accentué ou non ne dépend pas de la position occupée dans le vers. Si la fin d'hémistiche est accentuée, c'est parce qu'il s'y trouve une syllabe qui serait naturellement accentuée en discours. La langue, en informant le modèle métrique, introduit du rythme, et c'est ce dernier que nous avons cherché à décrire. Mais aussi, en plaçant ses accents et contrastes naturels à intervalles réguliers dans le vers, elle trace les contours du mètre.

Nos résultats confirment que notre hypothèse est suffisante pour rendre compte, et des résultats classiques, et d'une forme du vers, que nos calculs font émerger d'une manière claire et indiscutable. Les résultats sont solides vu la taille du corpus. Ils sont cohérents avec le savoir commun sur le vers, ce qui nous permet d'affirmer avec plus d'assurance les résultats nouveaux.

Les positions métriques

Afin de découper la matière des vers selon les positions métriques, nous avons examiné dans le détail les règles de correspondance entre le vers et le modèle, entre les syllabes de langue et les positions métriques ; et en particulier nous avons porté notre attention sur les questions de la diérèse, du *e* et de la liaison. Nous avons décrit tous les cas de diérèse et de synérèse sur notre corpus, c'est-à-dire la manière dont les groupements vocaliques commençant par une voyelle haute (*i*, *u* et *ou*) sont comptés dans le vers selon les mots dans lesquels ils figurent (diérèse si le groupement occupe deux positions métriques, synérèse s'il n'en occupe qu'une). Nous avons décrit le fonctionnement du *e* muet (les cas de maintien et d'élision principalement) et de la liaison dans le vers. Cet effort de formalisation visait l'introduction de règles de phonétisation dans l'outil d'analyse du vers que nous avons construit.

Cet examen détaillé des correspondances entre les « exemples de vers » et le « modèle de vers » [Jakobson, 1963] nous a montré qu'un certain nombre de phénomènes manifestent dans le vers une préférence pour des positions métriques organisées en CV (consonne suivie d'une voyelle qui définit une séquence creux-sommet). Autrement dit, le vers tendrait vers une forme qui serait la concaténation de douze séquences CV.

Si la préférence pour CV était effective, il devrait y avoir deux phonèmes par position métrique. Nous avons vu que le nombre de phonèmes variait selon les positions métriques : en positions 1 et 7, il est inférieur à 2, signalant le déficit d'attaque. En revanche, il est proche de 3 pour les positions 6 et 12. Les variations du nombre de phonèmes selon les positions jouent aussi un rôle rythmique dans le vers, au niveau de l'hémistiche. Le nombre de phonèmes est un marqueur de la forme de l'hémistiche qui, statistiquement, commence par une position avec peu de phonèmes et finit par une position riche en phonèmes.

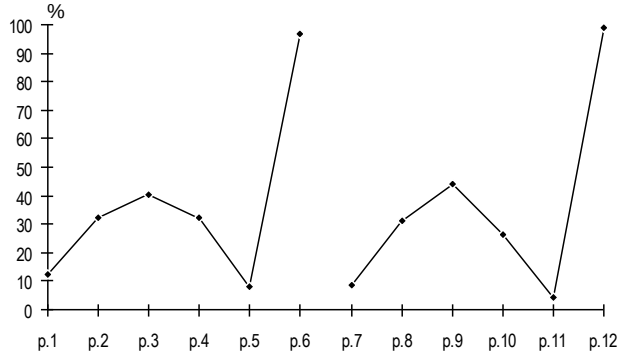
Le vers tend vers le modèle CV au niveau des règles de correspondance entre le vers et le modèle, mais les niveaux supérieurs du modèle, plus précisément celui de l'hémistiche, imposent au contraire une différentiation de la structure syllabique selon les positions. Le contenu des positions est le résultat d'un compromis entre ces deux tendances.

Cet examen nous fait toucher du doigt un point important sur lequel nous reviendrons : il n'y a pas indépendance entre les aspects phonétiques, syntaxiques et prosodiques. On ne peut dissocier les composantes de la langue. Ce point trivial prend toute son importance quand on réalise que le rythme est produit par la concordance métrique de la répartition des marqueurs des différentes composantes.

La forme de l'hémistiche

Il est admis depuis son origine que l'alexandrin est un vers constitué de deux hémistiches. En revanche, la nature de la césure a fait couler beaucoup d'encre. Nous avons essayé de préciser empiriquement ce qu'est la forme de l'hémistiche et quelle est la nature de la césure.

Nous avons examiné comment se répartissaient sur les douze positions métriques un certain nombre d'indicateurs linguistiques : les phonèmes vocaliques, les catégories syntaxiques, les fins de mots et les marques accentuelles. Tous ces indicateurs délimitent dans le vers deux formes similaires que la tradition désigne sous le terme d'hémistiche. C'est le cas du marquage accentuel, représenté dans le graphique suivant :



L'analyse des courbes de répartition pour une trentaine de marquages montre que *la longueur de la période est la même quel que soit l'indicateur retenu, soit six positions*. C'est cette longueur commune qui définit le mètre. Au niveau statistique, une « forme » spécifique se dessine sur l'hémistiche pour chacun de ces marquages, forme similaire sur le premier et le second hémistiche. Ces formes varient selon les marquages mais elles ont comme point

commun de présenter un écart important entre la dernière et la première position. Plutôt que de se concentrer sur la nature de la césure : est-ce un silence, une pause, un accent [*cf.* les controverses évoquées par Lote, 1949, p. 167 et sq.], nous regardons le profil de chaque hémistiche : la césure ne devient alors qu'une frontière entre deux formes similaires.

On remarquera que, du coup, le mètre apparaît clairement, même à qui ignorerait son existence. Construit sur le mètre, le rythme à son tour le met en évidence en tant que période de ses régularités. Mais si le mètre s'inscrit dans des mesures fixes, le rythme est une forme.

Un certain nombre de marquages (fin de mot, marque accentuelle, mots pleins) ont pour point commun de marquer quasiment systématiquement la fin du segment métrique et de constituer un faisceau de clôture. Ces mêmes indicateurs qui marquent la clôture sont au contraire peu présents en début d'hémistiche. Ainsi, il y a 62 % de polysyllabes en fin d'hémistiche contre 26 % en début ; la dernière position est quasiment toujours occupée par un mot plein, la première n'est occupée par un mot-plein que dans 31 % des cas ; la dernière position est toujours accentuée, la première ne l'est que dans 11 % des cas ; le nombre de phonèmes de la dernière position est proche de trois, il est inférieur à deux dans la première. De ce fait, le contraste entre la fin d'un hémistiche et le début d'un autre est très grand. Et il apparaît mieux encore quand on opte pour un marquage global qui combine les marquages spécifiques. La forme émerge du contraste. Comme toujours, c'est par les frontières que se dessinent les contours des formes, ici l'hémistiche.

Pour ces mêmes marquages, la saillance de la dernière position est renforcée par le très faible marquage de la précédente (positions 5 ou 11).

L'hémistiche existe en tant que forme parce que sa longueur de 6 se reproduit à l'identique de vers en vers pour tous les indicateurs linguistiques que nous avons utilisés, parce que chacun de ces indicateurs adopte au niveau statistique une forme spécifique sur les différentes positions de l'hémistiche, similaire sur les deux hémistiches, et enfin parce que la frontière est fortement marquée au niveau statistique par le contraste entre la fin d'un segment et le début du suivant.

Si l'on revient à la marque accentuelle, qui peut être considérée comme une marque fondamentale pour l'étude du rythme, et à sa répartition sur les deux hémistiches, on observe que les séquences d'hémistiches les plus fréquentes sont : H_010101 (qui recouvre les formes H_000101, H_010101 et H_010001) et H_001001 qui couvrent 65 % des hémistiches. Quasiment toutes les autres combinaisons de 0 et de 1 sur six positions sont attestées mais à des fréquences basses (moins de 10 %), ce qui nous a conduit à les regrouper sous la catégorie H_autres. Il existe donc un niveau rythmique intermédiaire entre la position et l'hémistiche : les positions des hémistiches sont dans la majorité des cas (65 %) organisées en groupements de 2, 3 ou 4 positions.

Le vers en tant que forme

Si la période dans le vers est bien l'hémistiche, il existe cependant des différences entre le premier et le second hémistiche. Les formes (courbes de répartition des marqueurs) ont statistiquement une allure semblable, mais celle du second hémistiche est plus contrastée et plus saillante. Ainsi, la douzième position est plus marquée que la sixième quel que soit le marquage étudié, la onzième nettement moins marquée que la cinquième... Si nous considérons le schème accentuel des hémistiches, nous voyons que les formes les plus régulières (H_001001 et H_010101) sont nettement plus fréquentes sur le second hémistiche que sur le premier (70 % d'hémistiches réguliers sur le second hémistiche contre 61 % sur le premier). En ce sens, l'unité du vers, assurée par son autonomie sur la ligne, par son nombre de positions, par la rime, est également marquée par le fait que le second hémistiche amplifie et exacerbe la forme du premier et qu'il n'en est pas une simple reproduction.

La rime

L'unité du vers est également assurée par la rime qui relie quelques vers entre eux (deux pour les alexandrins à rimes plates) grâce à une terminaison commune. Nous avons cherché à voir ce qui était nécessaire pour que deux vers riment ensemble. Nous avons construit à partir des couples de mots-rimes attestés des groupes de mots-rimes, qui rassemblent les mots qui peuvent rimer ensemble. Un travail manuel sur ces groupes nous a permis d'identifier le

segment commun à tous les mots de chaque groupe que nous désignons par le terme rimème. Il n'y a pas moins de 500 rimèmes différents.

Pour qu'il y ait rime, et c'est un résultat classique, il faut au moins l'équivalence de la dernière voyelle métrique du vers et de tout ce qui la suit, soit l'équivalence de la séquence VCT[170]. Nous avons démontré qu'un phonème d'appui est indispensable quand la voyelle est un /e/ et quand la séquence VCT correspond dans le couple de mots-rimes à un même morphème grammatical, ce deuxième résultat étant bien connu. En dehors de ces cas extrêmes, nous avons découvert que les éléments nécessaires pour qu'il y ait rime variaient selon le type de terminaison et nous avons essayé de cerner les critères qui régissent ces variations.

Nous avons pu construire une hiérarchie des voyelles en fonction de la fréquence avec laquelle elles mobilisent une consonne d'appui pour faire la rime. Cette hiérarchie est corrélée avec le nombre de mots-rimes attestés dans Corneille et Racine pour chacune de ces voyelles. En bref, plus une voyelle est rare, moins elle a besoin d'appui pour rimer.

Nous avons montré que la. dimension graphique jouait aussi un rôle important. Plus une terminaison est constituée d'un nombre élevé de graphèmes, moins elle a besoin d'une consonne d'appui.

Ce travail empirique nous conduit donc à revoir les critères de rimes suffisante, riche ou pauvre. Il ne peut y avoir de définition absolue de ces catégories, celles-ci varient selon le rimème considéré. Ainsi, pour certains rimèmes, l'identité du phonème d'appui sera majoritaire et l'on devra alors considérer que cela constitue la rime suffisante, on aura une rime pauvre quand ce phonème d'appui ne sera pas identique et une rime riche quand un phonème supplémentaire sera identique. Inversement, si pour un autre rimème, l'identité de la consonne est rare, le cas sera considéré comme une rime riche : ainsi une terminaison commune en *-er* ne peut constituer une rime, une terminaison en *-ir* est une rime pauvre tandis qu'une terminaison en *-or* constituera une rime suffisante. Il ne peut donc y avoir d'évaluation de la qualité d'une rime en dehors du contexte du rimème. C'est en comparant rimème par rimème le

[170] Voyelle - groupe Consonantique – Terminaison.

traitement de la rime au XIXe siècle à celui du XVIIe siècle que nous serons en mesure d'évaluer s'il y a eu ou non enrichissement de la rime. Beaucoup de travail reste à faire sur la rime, mais le RIMARIUM qui présente, par rimème, l'ensemble des couples de mots-rimes classés par fréquence décroissante devrait être un outil utile pour la poursuite des investigations. Nous l'avons intégralement fourni en annexe (XI).

Ce travail sur la rime présente aussi l'intérêt de donner une réponse empirique à la question des noyaux vocaliques dans le vers. Que retenir comme noyaux vocaliques pour les marquages ? La liste des phonèmes vocaliques semblait s'imposer, et cela fut notre point de départ. Grâce à l'analyse empirique de la rime, qui révèle des classes d'équivalences vocaliques, nous avons modifié notre position initiale. La distinction actuelle entre *a* ouvert et fermé ne vaut pas pour le vers du XVIIe siècle : nous n'avons pas distingué ces deux phonèmes vocaliques. En revanche, nous avons ajouté trois noyaux vocaliques constitués d'une semi-voyelle et d'une voyelle (/w a/, /w ɛ̃/ et /ɥi/). Ce résultat est un autre exemple de la contribution que peut apporter une analyse statistique sur de grands corpus.

Rythme et sens

Nous avons enfin cherché à voir dans quelle mesure le rythme pouvait jouer un rôle d'indice stylistique et comment il contribue, au niveau macrostructural, à l'élaboration d'une forme. Si dans les analyses de poèmes ou de passages en vers l'analyse stylistique sait bien intégrer dans l'interprétation du texte les phénomènes de versification, il n'en est pas de même, et pour cause, quand il s'agit d'analyser une pièce ou les œuvres d'un auteur. Nous avons cherché à établir des corrélations entre des champs ou univers lexicaux et des rythmes particuliers.

En découpant l'ensemble des pièces de Corneille et Racine en unités d'une dizaine de vers et en les classant avec un outil de statistique textuelle en fonction de la similarité des univers lexicaux mobilisés dans les fragments, nous avons identifié deux pôles dans ce corpus. Le premier sera résumé par *amour*, le second par *mort*, qui sont respectivement les mots les plus spécifiques de chaque pôle. Or, en mettant en relation ces deux pôles avec les figures rythmiques des hémistiches, nous voyons que la forme H_001001 est

significativement plus fréquente dans le pôle *mort* et la forme H_autres dans le pôle *amour*, tandis que la forme H_010101 ne varie guère entre les deux. L'amour apparaît donc comme un fond thématique plus propice que la mort à l'expression de la variété du rythme.

Cette distinction recouvre en partie des effets de genre et d'auteur. Nous voyons que, par rapport à Corneille, Racine privilégie nettement dans ses tragédies les rythmes les plus réguliers (H_001001 et H_010101). Si chez Corneille les différences rythmiques entre les genres sont faibles, chez Racine elles sont maximales : son unique comédie est beaucoup plus irrégulière en termes de rythme que celles de Corneille. Le rythme, tout comme le type de vocabulaire, le nombre de mots par vers, contribuent à différencier les genres et les auteurs. L'examen des tragédies centrales de Racine, qui sont très homogènes en termes de contenu, a permis de montrer que l'alternance des moments de récit et de dialogue s'accompagnait d'un changement dans le rythme des vers, les vers de récit étant plus réguliers que ceux du dialogue ; le dialogue correspondant en général à des moments de confrontation des personnages sur scène au cours desquels le vers devient plus haché.

Rythme et forme

Le rythme apparaît donc comme une structure feuilletée. Structure, car il s'inscrit dans un cadre métrique rigide défini par le nombre et délimité par de forts contrastes (fin d'hémistiche, de vers) sur les plans phonétique, syntaxique et prosodique. Feuilletée, car les courbes de répartition des marqueurs montrent tous une caractéristique commune, celle de la répétition entre les deux hémistiches (un peu accentué sur le second). Chaque composante de langue présente, de manière parallèle, ce même profil répétitif : le rythme est la superposition de toutes ces formes qui se répètent. Et les répétitions sont synchrones.

Par conséquent, on peut émettre l'hypothèse que les différentes composantes sont en phase : pour chacune, les contrastes sont au même endroit (positions 6 et 12). Ce fait est, comme on l'a dit, d'autant plus remarquable qu'il n'y a pas indépendance entre les différents niveaux de langue, et que ces dépendances ne sont en aucune manière simples ou linéaires. Nous voyons

concrètement comment le rythme est à la fois forme et répétition, et surtout, comment il structure de manière cohérente tous les niveaux de langue à la fois.

Rappelons que l'effet, pour évident qu'il apparaisse sur nos graphiques est en partie statistique. Il s'agit de règles pour les positions 6 et 12, mais pour le reste il s'agit de tendances et c'est bien en ce sens qu'il y a rythme : ce qui fait le charme et la part libre et singulière du feuilleté, c'est justement ces variations locales par rapport à la tendance générale.

On peut considérer nos résultats empiriques à la fois comme une confirmation, une illustration et un enrichissement de la théorie du rythme de Lusson et des recherches sur le vers de Roubaud. Confirmation éclatante notamment sur la symétrie des hémistiches et sur le fait que le vers est fait pour un œil-oreille. Illustration, car elle montre, à chaque niveau, la forme exacte du vers. Enrichissement enfin, dans la mesure où clarifiant la distinction entre rythme et mètre, elle donne chair à une microstructure feuilletée et synchrone. Ces résultats devraient aussi permettre d'avancer sur la théorie de la rime, qui n'est pas encore traitée par la théorie du rythme.

Ces deux dernières caractéristiques suggèrent que les différents niveaux de langue sont en phase (et nous n'avons examiné que les niveaux phonétique, syntaxique et lexical) et qu'il peut donc y avoir une amplitude en réception bien plus forte que celle que révèle l'examen des vers isolés. On sait bien, en effet, que l'amplitude d'un phénomène s'amplifie par des stimulations répétées aux bons intervalles.

Notons que, si le mètre est fortement marqué et se distingue aisément sur un vers isolé, l'effet individuel de chaque marquage rythmique, que nous avons repéré de manière statistique, semble trop faible pour être repérable par l'esprit humain. Mais nous avons montré que toutes les composantes étaient en phase. C'est peut-être la coïncidence des marquages qui explique pourquoi une structure rythmique est perceptible sur seulement quelques vers. Nous entrevoyons par le caractère synthétique et multidimensionnel de ces phénomènes de résonance la complexité perceptive de ce subtil composé de forme et de mesure dont l'expérience subjective reste en grande partie indicible.

BIBLIOGRAPHIE

Bibliographie critique

ANGOUJARD, J. P. (1992). « L'interprétation linguistique du rythme », *Percevoir, raisonner, agir. Articulation de modèles cognitifs*, Actes du Cinquième colloque de l'ARC, Nancy, pp.57-67.

AROUI Jean-Louis (1996). « Rimbaud : les rimes d'une *Larme* », *Parade sauvage*, n°13, pp. 24-44.

ARRIVE Michel, GADET Françoise, GALMICHE Michel (1986). *La grammaire d'aujourd'hui*, Paris, Flammarion, 720 p.

BARTHES Roland (1960). *Sur Racine*, Seuil, Coll. « Points Essais », 1979, 160 p.

BEAUDOUIN Valérie (1993). « Stylistique et analyse lexicale : Corneille et Racine ». In : S. J. Anastex (eds.), (1993). *JADT 1993 : Actes des Secondes journées internationales d'analyse statistique de données textuelles*, Paris, TELECOM, pp. 51-62.

BEAUDOUIN Valérie, YVON François (1994). « Le Métromètre et les douze positions alexandrines », *Mezura*, n°32, Paris, Publications Langues'O, 33 p.

BEAUDOUIN Valérie, YVON François, (1996). « The Metrometer : a Tool for Analysing French Verse », *Literary & Linguistic Computing*, vol. 11, n°1, Oxford University Press, 1996, pp. 23-32.

BEAUMATIN Eric (1999). « Pour l'analyse métrique du Laberinto de fortuna : un réexamen des propositions de Jacques Roubaud sur l'Arte Mayor », *Les langues néo-latines*, pp. 65-87.

BEHAR Henri (1994). « Le traitement du plein texte ». In : Martin Eveline (1992). *Traitements informatisés de corpus textuels*, Paris, Didier Érudition, pp. 13-29.

BENICHOU Paul (1948). *Morales du grand siècle*, Paris, Gallimard, 317 p.

BENVENISTE Émile (*1951*, 1966). « La notion de "rythme" dans son expression linguistique », Journal de psychologie, 1951. In : *Problèmes de linguistique générale*, Paris, Gallimard, pp. 327-335.

BENZECRI Jean-Paul (1966). *Imago,* Présentation du laboratoire de calcul de l'institut de statistique de l'Université de Paris, sans numérotation.

BENZECRI Jean-Paul (1968). « La place de l'a priori, "Organum" », *Encyclopedia Universalis*, pp. 11-24.

BENZECRI Jean-Paul éd., (1981). « Introduction I », *Pratique de l'analyse des données, Linguistique et lexicologie,* Paris, Dunod.

BENZECRI Jean-Paul et coll. (1981). *Pratique de l'analyse des données, Linguistique et lexicologie,* Paris, Dunod.

BERNARDY Michel (1988). *Le jeu verbal. Traité de diction française à l'usage de l'honnête homme*, Paris, L'aube poche, 215 p.

BERNET Charles (1983). *Le vocabulaire des tragédies de Racine*, Paris-Genève, Slatkine-Champion, 385 p.

BERRETI Jany (1992). « De deux mots sans importance. I. Place de et dans l'alexandrin. II. Roman d'un mot du français : de », *Cahiers de poétique comparée*, n°20, INALCO, pp. 41-87.

BILLY Dominique (1984). « La nomenclature des rimes », *Poétique*, n°57, pp. 64-75.

BILLY Dominique, CORNULIER Benoît de, GOUVARD Jean-Michel (éds) (1993). Métrique française et métrique accentuelle, *Langue française*, Larousse, n°99, 127 p.

BONHOMME Patrice (1995). *La Poétique du vers. Jorge Guillén :* Aire nuestro, Presses Universitaires de Bordeaux, 352 p.

BONNARD Henri (1981). *Code du français courant*, Paris, Magnard.

BRUNET Étienne (1989). « L'exploitation des grands corpus : le bestiaire de la littérature française ». *Literary and Linguistic Computing*, t. 4, n°2, pp. 121-132.

BRUNET Étienne (1993). «Un hypertexte statistique : Hyperbase». In S. J. Anastex (eds.), (1993). *JADT 1993*, Paris, TELECOM.

CERCLE POLIVANOV (1969). *Bulletin 0.*

CERQUIGLINI Bernard (1991). *La naissance du français*, PUF, Que sais-je ?, 128 p.

CERQUIGLINI Bernard (1995). *L'accent du souvenir*, Paris, Les éditions de Minuit, 168 p.

CHOMSKY Noam, HALLE Morris (1968, 1973). *Principes de phonologie générative*, (trad. de *Sound Pattern of English*), éd. du Seuil, 1973, 351 p.

CHRISTOPHE Anne (1993). *Rôle de la prosodie dans la segmentation en mots*, Thèse EHESS, sous la direction de J. Mehler, 253 p.

CONSTANT Patrick (1991). *Analyse syntaxique par couche*, Paris, Doctorat ENST.

CONSTANT Patrick (1995). « L'analyseur syntaxique SYLEX », Ecole d'été du CNET, Lannion.

CONSTANT Patrick (1996). « Reducing the Complexity of Encoding Rule-Based Grammars », Rapport interne n°9602 du laboratoire de Sémiotique Informatique, Paris, La Défense.

CORNULIER Benoît de (1977). « Le vers français classique », *Le français moderne*, 1977, n°2, pp. 97-125.

CORNULIER Benoît de (1981). « La rime n'est pas une marque de fin de vers », *Poétique*, n°46, pp. 247-256.

CORNULIER Benoît de (1982a). « La cause de la rime », *Poétique*, n°52, pp. 499-508.

CORNULIER Benoît de (1982b). *Théorie du vers. Rimbaud, Verlaine, Mallarmé*, Paris, Éditions du Seuil, 1982, 381 p.

CORNULIER Benoît de (1985). « Rime "riche" et fonction de la rime », *Littérature*, n°59, pp. 115-125.

CORNULIER Benoît de (1995). *Art Poëtique. Notions et problèmes de métrique*, Lyon, Presses Universitaires de Lyon, 1995, 300 p.

COUPRIE Alain (1994). *Lire la tragédie*, Paris, Dunod, 262 p.

DELL François (1984). « L'accentuation dans les phrases en français ». In : Dell, François, Hirst, Daniel, Vergnaud, Jean-Roger (eds), *Forme sonore du langage. Structure des représentations en phonologie*, Paris, Hermann, 1984, pp. 65-122.

DELL François (1985). *Les règles et les sons*, Paris, Hermann, 2 édition, 287 p.

DELL François, HIRST Daniel, VERGNAUD Jean-Roger (eds), *Forme sonore du langage. Structure des représentations en phonologie*, Paris, Hermann, 1984.

DERQUENNE Christian (1992). *Caractérisation statistique d'une typologie : principales méthodes et présentation du programme « Qualicls »*, note technique, EDF-DER, 29 p.

DEVAUX Patrick (1993). *La Comédie-Française*, Que sais-je ?, Paris, PUF, n°2736, 128 p.

DIDAY E. et LEBART L. (1977). «L'analyse des données», *La Recherche*, n°74, pp. 15-25.

DINU Mihai (1993). « Structures accentuelles de l'alexandrin chez Racine », *Langue française*, n°99, pp. 63-74.

DOMINICY Marc (1992). « On the Meter and Prosody of French 12-syllable Verse », *Empirical Studies of the Arts*, vol 10(2), pp. 157-181.

DOMINICY Marc, NASTA Mihaï (1993). « Métrique accentuelle et métrique quantitative », *Langue française*, n°99, pp. 75-96.

DUPOUX Emmanuel, MEHLER Jacques (1992). « La segmentation de la parole », Le courrier du CNRS, n°79, p. 10

DUPRIEZ Bernard (1984). *Gradus, Les procédés littéraires*, Union Générale d'édition, 10/18.

DURAND Jacques (1990). *Generative and non-linear phonology*, Longman, Londres et New-York, 338 p.

EVRARD Etienne, PURNELLE Gérald (1998). « Rime et action dans la tragédie française : Racine, Corneille, Hugo », *Revue Informatique et Statistique dans les Sciences humaines*, pp. 71-106.

ELWERT W. Theodor (1965). *Traité de versification française. Des origines à nos jours*, Paris, Klincksieck, 210 p.

ENCREVE Pierre (1988). *La liaison avec et sans enchaînement*, Paris, Editions du Seuil, 319 p.

FABRE Cécile, HABERT Benoît (1998). « Acquisition de relations entre mots pour une lecture sémantique de corpus ». In : S. Meillet (éd.). *JADT 98*, 4èmes journées internationales d'analyse statistique des données textuelles, Nice, pp. 273-282.

FISHER Sophie (1995). « L'*exemple* est-il l'objet de la linguistique », *Mélanges de l'école française de Rome*, tome 107, 2, pp. 458-471.

FOUCHE Pierre (1959, 1969). *Traité de prononciation française*, Paris, Klincksieck, 528 p.

FRAISSE Paul (1956). *Les structures rythmiques*, Louvain, Publications Universitaires de Louvain, 127 p.

GARRETTE Robert (1995). *La phrase de Racine. Étude stylistique et stylométrique*, Toulouse, Presses Universitaires du Mirail, 333 p.

GASPAROV M. L. (1987). « A Probability Model of Verse (English, Latin, French, Italian, Spanish, Portuguese) », New metrics, *Style*, vol. 21, n°3, pp. 322-358.

GAUTHIER Michel (1974). *Système euphonique et rythmique du vers français*, Paris, Klincksieck, 165 p.

GENETTE Gérard (1972). *Discours du récit. Figures III*, Paris, Seuil, Collection Poétique, pp. 67-282.

GENETTE Gérard (1982). *Palimpsestes. La littérature au second degré*, Paris, Seuil, « Poétique », 579 p.

GETZLER Pierre (1977). « Éléments d'un bibliographie de l'alexandrin », *Cahiers de poétique comparée*, Vol. 2-fasc.4, Publications Langues'O, pp. 7-95. [bibliographie très exhaustive de 1920 à 1975]

GOUVARD Jean-Michel (1994). *Recherches sur la métrique interne du vers composé dans la seconde moitié du dix-neuvième siècle. Pour une analyse distributionnelle systématique*, thèse de doctorat, Centre d'études métriques, Université de Nantes.

GOUVARD Jean-Michel (1995). « Métriques de l'alexandrin chez Mallarmé et Dario : du vers français au vers espagnol », Littérature et nation, n°15, pp. 139-156.

GOUVARD Jean-Michel (1996). « Le vers français : de la syllabe à l'accent », *Poétique*, n°106, avril 1996, Seuil, pp. 223-247

GRAMMONT Maurice (1904, 1947). *Le vers français, ses moyens d'expression, son harmonie*, Paris, Delagrave, 2eme édition.

GRAMMONT Maurice (*1908*, 1965). *Petit traité de versification française*, Paris, Armand Colin, 1965, 158 p.

GRAMONT Ferdinand de (1876). *Les vers français et leur prosodie*, Paris.

GREIMAS Algirdas Julien (1986). *Sémantique structurale*, Paris, PUF, 263 p.

HABERT Benoît, NAZARENKO Adeline, SALEM André (1997). *Les linguistiques de corpus*, Paris, Armand Colin/Masson, 240 p.

HABERT Benoît, SALEM André (1995). « L'utilisation de catégorisations multiples pour l'analyse quantitative de données textuelles », *T. A. L.*, 36, n°1-2, pp. 249-275.

HALLE Morris, KEYSER Samuel Jay (1966). « Chaucer and the Study of Prosody », College English, n°28, pp. 187-219.

HALLE Morris, KEYSER Samuel Jay (1971). *English Stress : Its Form, Its Growth, and Its Role in Verse*, New York Evanston London, Harper & Row, Publishers, 206 p.

HAMBURGER Kate (1957, 1986 pour la traduction française). *Logique des genres littéraires*. traduction P. Cadiot, Préface G. Genette, Paris, Seuil, Collection Poétique, 1986.

HAYWARD Malcom (1991). « A Connectionist Model of Poetic Meter », *Poetics*, North-Holland, pp 303-317.

HAYWARD Malcom (1996). « Analysis of a corpus of poetry by a connectionist model of poetic meter », *Poetics*, 24, The Netherlands, p. 1-11.

HAYWARD Malcom (1996). « Applications of a Connectionist Model of Poetic Meter to Problems in Generative Metrics », *Research in Humanities Computing 4*, Oxford, Clarendon Press, pp.185-192.

INGENIA S.A. (1995). *SYLEX version 1.5. Manuel d'utilisation*, Paris.

JAKOBSON Roman, HALLE Morris, CHOMSKY Noam, FAYE Jean-Pierre, PARIS Jean, ROUBAUD Jacques, RONAT Mitsou (1972). *Hypothèses. Trois entretiens et trois études sur la linguistique et la poétique*, Paris, Seghers-Laffont, Change, 224 p.

JAKOBSON Roman. (1960, 1963). « Poétique », *Essais de linguistique générale*, Paris, Ed. de Minuit, pp.207-248 (Trad. Nicolas Ruwet).

JOUET Jacques, ROUBAUD Jacques (1993). « [ə] », *La Bibliothèque Oulipienne*, n°64, pp. 5-34.

KAYE Jonathan D. , LOWENSTAMM Jean (1984). « De la syllabicité ». In : DELL François, HIRST Daniel, VERGNAUD Jean-Roger (eds) *Forme sonore du langage. Structure des représentations en phonologie*, Paris, Hermann, pp. 123-160.

KIPARSKY Paul (1975). « Stress, Syntax and Meter », *Language*, 51, pp. 576-616.

KIPARSKY Paul (1977). « The Rythmic Structure of English Verse », *Linguistic Inquiry 8*, pp. 189-247.

KIPARSKY Paul, YOUMANS Gilbert (eds) (1989). *Phonetics and Phonology : Rythm and Meter*, San Diego, Academic Press, 399 p.

KLEIN Marc (1991). *Vers une approche substantielle et dynamique de la constituance syllabique. Le cas des semi-voyelles et des voyelles hautes dans les usages parisiens*, Thèse de doctorat, Paris VIII, 342 p.

KLEIN Marc (1993). « La syllabe comme interface de la production et de la réception phoniques ». In : LAKS Bernard, PLENAT Marc (ss dir.), *De natura sonorum*. Essais de phonologie, Saint-Denis, PUV Saint-Denis, pp. 101-142.

LABBE Dominique (1993). «Un modèle d'analyse du vocabulaire», *Secondes journées internationales d'analyse statistique de données textuelles*, Montpellier, ENST-TELECOM, pp.103-114.

LABBE Dominique, HUBERT Pierre (1998). « La connexion des vocabulaires ». In : S. Meillet (éd.). *JADT 98*, 4èmes journées internationales d'analyse statistique des données textuelles, Nice, pp. 361-369.

LAHLOU Saadi (1995). « Vers une théorie de l'interprétation en analyse statistique des données textuelles ». In : S. Bolasco, L. Lebart, A. Salem (eds) (1995). *JADT 1995 : III Giornate internazionali di Analisi Statistica dei Dati Testuali.* Roma : CISU, vol. I, pp. 221-228.

LAHLOU Saadi (1998). *Penser manger. Alimentations et représentations sociales*, Paris, PUF, 241 p.

LAKS Bernard, PLENAT Marc (ss dir.), *De natura sonorum*. Essais de phonologie, Saint-Denis, PUV Saint-Denis.

LAWDER Bruce (1993). *Vers le vers*, Librairie Nizet, 286 p.

LE TELLIER Hervé (1997). *Les amnésiques n'ont rien vécu d'inoubliable*, Le Castor astral, Bordeaux, 135 p.

LE VOT Gérard, LUSSON Pierre, ROUBAUD Jacques (1979), « La chanson de "L'*Amour de Loin"* de Jauffré Rudel. Essai de lecture rythmique », *Mezura*, n°3, pp. 3-92.

LEBART Ludovic, SALEM André (1994). *Statistique textuelle*, Paris, Dunod, 342 p.

LOTE Georges (*1913*, 1975). *Études sur le vers français. L'alexandrin d'après la phonétique expérimentale*, tomes 1 et 2, Paris ; Genève : Slatkine, 1975.

LOTE Georges (1949). *Histoire du vers français*, tome I, Éditions Boivin et Cie, 362 p.

LUSSON Pierre (1973). « Notes préliminaires sur le rythme », *Cahiers de poétique comparée*, vol I - fascicule 1, Paris, Publications Langues' O, pp. 30-54.

LUSSON Pierre (1982). « Isomorphismes rythmiques dans les récitatifs de l'*Alceste* de Lulli », *Cahiers de poétique comparée*, n°6, Publications Langues'O, pp. 7-80

LUSSON Pierre (1998). « Une méthode d'analyse des rapports texte/musique : application d'une théorie générale du rythme », *Mezura* n°13, pp. 7-45.

LUSSON Pierre, ROUBAUD Jacques (1974). « Mètre et rythme de l'alexandrin ordinaire », *Langue française*, n°23, septembre 1974, pp. 41-53.

LUSSON Pierre, ROUBAUD Jacques (1980). « Lectures rythmiques XI - XII, deux sonnets en vers imbriqués d'Étienne Jodelle », *Mezura*, Paris, n° 10, 1980.

LUSSON Pierre, ROUBAUD Jacques (1981). « Sur la devise de “Nœu et de feu” un sonnet d’Étienne Jodelle », *Langue française*, Paris, n° 49, 1981, pp. 49-67.

LUSSON Pierre, ROUBAUD Jacques (1983). « Le sonnet de Veniero et ses traductions : une lecture rythmique » *Cahiers de poétique comparée*, n°7, Publications Langues’O, pp. 7-32.

MALMBERG Bertil (1954). *La phonétique*, Paris, Que sais-je, PUF, 128 p.

MARTIN Éveline (1991). « FRANTEXT, base de données textuelles : une aide nouvelle à l'exploration des textes », *Bulletin de l'Institut d'histoire du temps présent*, 1991, n°46, pp. 17-22

MARTIN Eveline (1997). « Le traitement des corpus textuels à l’Institut National de la Langue Française (CNRS) », *Literary and Linguistic Computing*, vol. 12, n°1, Oxford University Press, 1996, pp. 37-44.

MARTINET André (1970, 1991), *Éléments de linguistique générale*, Paris, Armand Colin.

MARTINON Philippe (1905, 1962). *Dictionnaire méthodique et pratique des rimes françaises*. Précédé de *Versification française (Théorie et pratique)*, Paris, Librairie Larousse, 1962, 288 p.

MAZALEYRAT Jean (1963). *Pour une étude rythmique du vers français moderne. Notes bibliographiques*, Lettres Modernes, Paris, 127 p.

MAZALEYRAT Jean (1974). *Eléments de métrique française*, Paris, Armand Colin, 232 p.

MELESE Pierre (1951). *Racine*, in Théâtre de Racine, Bibliothèque des éditions Richelieu, 1951. tome 1, pp.13-150.

MELHER Jacques, CHRISTOPHE Anne (1995). « Maturation and Learning of Language in the First Year of Life », In : GAZZANIGA Michael S. (ed), The Cognitive Neurosciences, pp. 943-954.

MESCHONNIC Henri (1970). *Pour la poétique I*, Paris, NRF, Gallimard, 180 p.

MESCHONNIC Henri (1982). *Critique du rythme*, Paris, Editions Verdier, 735 p.

MESGUICH Daniel (1993). « “Andromaque” de Jean Racine ». In : BOUGNOUX Daniel, Sciences de l’information et de la communication, Paris, Larousse, p. 748

MILNER Jean-Claude (1974). « Réflexions sur le fonctionnement du vers français », *Cahiers de poétique comparée*, vol I, fasc. 3, Paris, Publications Langues' O, pp 2-21, repris dans *Ordres et raisons de langue*, Paris, Seuil, 1982, pp. 283-301.

MILNER Jean-Claude (1987). « Accent de vers et accent de langue dans l'alexandrin classique », *Cahiers de poétique comparée*, n°15, Publications Langues'O, pp. 31-77.

MILNER Jean-Claude (1989). *Introduction à une science du langage*, Paris, Éditions du Seuil, 713 p.

MILNER Jean-Claude, REGNAULT François (1987). *Dire le vers. Court traité à l'intention des acteurs et des amateurs d'alexandrins*, Paris, Seuil, 188 p.

MIQUEL Jean-Pierre (1997). *Propos sur la tragédie*, Actes Sud, 111 p.

MOLINIE Georges (1986). *Éléments de stylistique française*, Paris, PUF, 213 p.

MOLINO Jean, Tamine Joëlle (1982). « Des rimes, et quelques raisons... », *Poétique*, n°52, pp. 487-498

MORIN Yves-Charles (1993). « La rime d'après le *Dictionnaire des rimes* de Lanoue (1596) ». In BILLY Dominique, CORNULIER Benoît de, GOUVARD Jean-Michel (éds), Métrique française et métrique accentuelle, *Langue française*, n°99, pp. 107-123.

MOUGIN Pascal (1995). « Mondes lexicaux et univers sémantiques : le logiciel ALCESTE au service de l'étude de l'imaginaire simonien, à partir du traitement de *La Route des Flandres* », *Literary & Linguistic Computing*, vol. 10, n°1, Oxford University Press, 1995, pp. 59-68.

MULLER Charles (1967, 1992*). Étude de statistique lexicale. Le vocabulaire du théâtre de Pierre Corneille*, Paris, Larousse, 1967, réimpression aux éditions Slatkine, 1979, 1992 382 p.

MULLER Charles (1977, 1992). *Principes et méthodes de statistique lexicale*, Larousse, 1977, réimpression Champion-Slatkine, 1992, 211p.

PIGAMO Frédéric (1990). *Outils de traitement sémantique du langage naturel*, Paris, Doctorat ENST, 242 p.

PRIJOVIC Sacha, SOLA Agnès (1983). « Le tétramètre ïambique de Pouchkine dans *Eugène Onéguine* », *Cahiers de poétique comparée*, n°8, Publications Langues'O, pp. 7-40.

QUICHERAT Louis (1838). *Petit traité de versification française*, Hachette.

QUILIS Antonio (1968). *Métrica española*. Edicion corregida y aumentada, Barcelona, Editorial Ariel, 1984, 213 pp.

RASTIER François (1987). *Sémantique interprétative*, Paris, PUF, 279 p.

RASTIER François (1989). *Sens et textualité*, Paris, Hachette, 287 p.

REGNAULT François (1996). *La Doctrine inouïe. Dix leçons sur le théâtre classique français*, Paris, Hatier, 314 p.

REINERT Max (1983). « Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte », *Les cahiers de l'analyse des données*, Vol VIII, n° 2, p 187-198.

REINERT Max (1987). « Classification descendante hiérarchique et analyse lexicale par contexte : application au corpus des poésies d'Arthur Rimbaud », *Bulletin de Méthodologie Sociologique*, n°13.

REINERT Max (1990). « ALCESTE : Une méthodologie d'analyse des données textuelles et une application : Aurélia de Gérard de Nerval », *Bulletin de méthodologie sociologique*, n°26, pp. 24-54.

REINERT Max (1993). « Les "mondes lexicaux" et leur logique ». *Langage et société*, Paris, Maison des Sciences de l'Homme, n°66, pp. 5-39.

RONAT Mitsou (1975). « Métrico-phono-syntaxe : le Vers Français Alexandrin », *Cahiers de poétique comparée*, vol II - fascicule 2, Paris, Publications Langues' O, 1975, pp. 11-46.

RONAT Mitsou (1975b). « Scrambling peut-elle être française ? », *Cahiers de poétique comparée*, vol II - fascicule 2, Paris, Publications Langues' O, pp. 47-92.

ROUBAUD Jacques (1969). « Thèses sur la poétique (notes partielles) », *Bulletin 0*, Cercle Polivanov.

ROUBAUD Jacques (1971). « Mètre et vers : Deux applications de la métrique générative de Halle-Keyser », *Poétique*, n°7, pp. 366-387.

ROUBAUD Jacques (1977). « Une vie : Synopsis en douze branches. Première partie », *Cahiers de poétique comparée*, vol III - fascicule 2, Paris, Publications Langues' O, pp. 3-85.

ROUBAUD Jacques (*1978*, 1988). *La vieillesse d'Alexandre*, Paris, Éditions Ramsay, (éd. François Maspero 1978), 219 p.

ROUBAUD Jacques (1979). « Description du projet », *Mezura*, Publication Langues'O, Paris, n° 9, 95 p.

ROUBAUD Jacques (1983). « Le pentamètre ïambique dans mille sonnets élizabéthains », *Cahiers de poétique comparée*, n°8, Publications Langues'O, pp. 41-112.

ROUBAUD Jacques (1986). « DYNASTIE : études sur le vers français, sur l'alexandrin classique », Première partie, *Cahiers de poétique comparée*, n° 13, Paris, Publications Langues' O, pp. 47-109.

ROUBAUD Jacques (1988). « DYNASTIE : études sur le vers français, sur l'alexandrin classique », Deuxième partie, *Cahiers de poétique comparée*, n° 16, Paris, Publications Langues' O, pp. 41-60.

ROUBAUD Jacques (1990). La Forme du sonnet français de Marot à Malherbe, *Cahiers de poétique comparée*, n°17-18-19, Publications Langues'O, 386 p.

ROUBAUD Jacques (1991). "T.R.A.(M,m) (question d'une poétique formelle, I)", Mezura, n° 24, Paris, Publications Langues' O, 26 p.

ROUBAUD Jacques (1995). *Poésie, etcetera : ménage*, Paris, Stock, 284 p.

ROUBAUD Jacques (1998). *La ballade et le chant royal*, Les Belles Lettres, 400 p.

SAS Institute Inc. (1989). *SAS Procedures Guide*, Version 6, Third Edition Cary, NC. , Cary, NC : SAS Institute Inc.

SAS Institute Inc. (1990). *SAS language : Reference*, Version 6, First Edition, Cary, NC : SAS Institute Inc.

SHANE S. (1968). *French Phonology and Morphology*, Cambridge, Mass. : MIT Press.

SHAPIRO Michaël (1974). « Sémiotique de la rime », *Poétique*, n°20, 1974 , pp. 501-519.

SPITZER Léo. (*1931*, 1982). « L'effet de sourdine », *Études de style*, Paris, Tel, Gallimard, pp. 208-335.

STAROBINSKI Jean (1961). *L'Œil vivant*, Paris, Gallimard, 259 p.

TARLINSKAJA Marina (1987). « Rythm and Meaning : "Rythmical Figures" in English Iambic Pentameter, Their Grammar and Their Links with Semantics », *Style*, vol. 21, n°1, pp. 1-35.

TESNIERE Lucien (*1959*, 1982). *Éléments de syntaxe structurale*, Paris, Klincksieck, 1959, 674 p.

TOBLER Adolphe (1885, 1972). *Le vers français ancien et moderne*, Paris, Viewveg, rééd. Slatkine.

TOGEBY Kund (1968). « Histoire de l'alexandrin français ». In : *Mélange pour Togeby*, Copenhague, pp. 208-234

TOURNIER Maurice (1980). « D'où viennent les fréquences de vocabulaire ? », *Mots*, Paris, Presses de la fondation nationale des sciences politiques, n° 1, pp. 189-209.

UBERSFELD Anne (*1977*, 1982). *Lire le théâtre*, Paris, Éditions Sociales.

VAISSIERE Jacqueline (1991). « Rhythm, accentuation and final lengthening in French ». In : SUNDBERG Johan, NORD Lennart, CARLSON Rolf (eds). *Music, Language, Speech and Brain*. Wenner-Gren International Symposium Serie, vol. 59, pp.108-120.

WAGNER R.-L., PINCHON J. (1962). Grammaire du français classique et moderne, Paris, Hachette, 1962, 648 p.

WIOLAND François (1985). *Les structures syllabiques du français. Fréquence et distribution des phonèmes consonantiques. Contraintes idiomatiques dans les séquences consonantiques*, Genève-Paris, Slatkine-Champion, 356 p.

YOUMANS Gilbert (1986). « Iambic Pentameter. Statistics or Generative Grammar ? », *Language and Style*, Queen College Press, vol. 19, n°4, pp. 388-404.

YOUMANS Gilbert (1989). « Introduction : rythm and meter ». In : KIPARSKY Paul, YOUMANS Gilbert (eds) (1989). *Phonetics and Phonology : Rythm and Meter*, San Diego, Academic Press, pp. 1-14.

YVON François (1993). « A Tidy Rule Based Graphem to Phonem Transcriber », Proceedings of ONOMASTICA Research Colloquium, London, 1993, pp. 47-56.

YVON François (1996). *Prononcer par analogie : motivation, formalisation et évaluation*, Thèse de l'ENST, 278 p.

YVON François et al. (1998). « Objective evaluation of grapheme to phoneme conversion for text-to-speech synthesis in French », *Computer Speech and Language*, 12, pp. 393-410.

ZIPF G. K. (1936, 1974). *La psychobiologie du langage : une introduction à la philologie dynamique*, Paris, RETZ-CEPL, 1974, 232 p.

Œuvres en vers

Textes informatisés

Corneille, Pierre (1629-1674). *Œuvres théâtrales*, Ed. Ch. Marty Laveaux, Paris, Hachette, 1862, (Bibliothèque Nationale).

Racine, Jean (1664-1691). *Œuvres théâtrales*, Ed. Paul Mesnard, Paris, Hachette, 1885, (Bibliothèque Nationale).

Malherbe, François de (1627). *Les poésies*, Ed. J. Lavaud, Paris, Droz, 1936-1937, (Frantext).

Rotrou, Jean de (1649). *Cosroès*, Ed. J. Scherrer, Paris, Didier, 1950, (Frantext).

Boileau-Despreaux, Nicolas (1674-1683). *Le lutrin*, Œuvres complètes, T.2, Paris, Les Belles Lettres, 1939, (Frantext).

Boileau-Despreaux, Nicolas (1694). *Satire X*, In Œuvres complètes, T.1, Paris, Les Belles Lettres, 1934, (Frantext).

Voltaire (1761). *Tancrède*, In Théâtre. Paris, Garnier, 1868, (Frantext).

Hugo, Victor (1856). *Les Contemplations* (Livre I), Ed. J. Vianey, Paris, Hachette, 1922, (Frantext).

Hugo, Victor (1859-1883). *La Légende des siècles*, Ed. P. Berret, Paris, Hachette, 1920-1927, (Frantext).

Baudelaire, Charles (1861). *Les Fleurs du Mal*, Paris, J. Corti, 1942, (Frantext).

Mallarmé, Stéphane (1898). *Poésies*. In Œuvres Complètes. Paris, Gallimard, la Pléiade, 1965 1945 (Frantext).

Rimbaud, Arthur (1870-1871). *Poésies*, Paris, Poésie / Gallimard, 1973, (Max Reinert).

Éditions consultées

Boileau Despreaux, Nicolas, *Œuvres*, Paris, Firmin Didot et Cie, 1928.

Corneille, Pierre, *Comédies*, Paris, GF-Flammarion, 1968, 695 p.

Corneille, Pierre, *Œuvres Complètes*, Présentation et notes de André Stegmann, Paris, Ed. du Seuil, L'Intégrale, 1963, 1133 p.

Corneille, Pierre, *Théâtre II*, Paris, GF-Flammarion, 1980, 603 p.

Hugo, Victor, (1856), *Les Contemplations*, Préface, Commentaires et notes de Jean Gaudon, Paris, Le Livre de Poche, 1985, 576 p.

Hugo, Victor, *La Légende des siècles*, In Œuvres Complètes, Poésie II et Poésie III, Paris : Robert Laffont, Bouquins, 1985, p. 564-833 et p. 182-712.

Le roman d'Alexandre (éd. Laurence Harf-Lancner), Lettres gothiques, Le Livre de poche, Paris, 864 p.

LY Nadine (éd.), *Anthologie bilingue de la poésie espagnole*, Bibliothèque de la Pléiade, Gallimard, Paris, 1995.

Mallarmé, Stéphane, *Œuvres complètes*, Bibliothèque de la Pléiade, Gallimard, Paris, 1977.

Mallarmé, Stéphane, *Vers et Prose*, Paris, Garnier-Flammarion, 1977.

Racine, Jean, *Théâtre complet*, texte établi et annoté par Pierre Mélèse, Bibliothèque des éditions Richelieu, 1951, 5 volumes.

Rimbaud, Arthur, *Œuvres complètes*, Bibliothèque de la Pléiade, Gallimard, Paris.

Rimbaud, Arthur, *Poésies*, Paris, GF-Flammarion, 1989.

Rimbaud, Arthur, *Poésies*. Une saison en enfer. Illuminations, Paris, Poésie/Gallimard, 1973, p. 17-97.

Ronsard, *Œuvres complètes*, Bibliothèque de la Pléiade, Gallimard, Tomes I & II, 1994.

SCHMIDT Albert-Marie (éd.). *Poètes du XVIe*, Bibliothèque de la Pléiade, Gallimard, Paris, 1953.

Représentations de Corneille et Racine (vues ou entendues)

Corneille, *Suréna*, mise en scène d'Anne Delbée, Vieux-Colombier, 1998, avec François Beaulieu (Orode), Dominique Constanza (Eurydice), Michel Favoy (Sillace), Eric Ruf (Suréna), Nathalie Nerval (Ormène), Clotilde de Bayser (Palmis), Christian Gonon (Pacorus).

Corneille, *Rodogune*, mise en scène de Jacques Rosner, Comédie-Française, 1998, avec Catherine Samie (Laonice), Martine Chevallier (Cléopâtre), Jean-Pierre Michaël (Antiochus), Cécile Brune (Rodogune), Pierre Vial (Timagène), Laurent d'Olce (Séleucus), Roger Mollier (Oronte).

Corneille, *Clitandre*, mise en scène par Muriel Mayette, Comédie-Française, 1997, avec Anne Kessler (Dorise), Jean-Pierre Michaël (Clitandre), Eric Ruf (Pymante), Samuel Le Bihan (Rosidor), Florence Viala (Caliste)...

Corneille, *Horace*, mise en scène par Jean-Pierre Miquel, Comédie-Française, 1971-1972, enregistrement INA-Comédie-Française, Film Office SA, VHS, Michel Etcheverry (Le vieil Horace), Simon Eine (Tulle), Jean-Noël Sissia (Valère), François Beaulieu (Horace), Nicolas Siberg, Jean-Luc Boutté (Curiace), Claude Winter (Julie), Christine Fersen (Sabine), Ludmila Mikaël (Camille).

Corneille, *Horace*, mise en scène par Marion Bierry, Théâtre de l'œuvre, 1998, avec Delphine Rich, Elizabeth Vitali, Isa Mercure, Claude Giraud, Stéphane Bierry, Jean-François Vlerick, Hervé Falloux, Thomas Cousseau, Thierry Courtois, Olivier Sabin.

Corneille, *Le Cid*, enregistrement sur la scène du TNP, 1955, Auvidis-Hachette, Théâtre, CD, avec Sylvia Monfort (Chimène), Mona-Dol (Elvire), Monique Chaumette (L'infante), Laurence Constant, George Wilson (Le Comte), Jean Deschamps (Don Diègue), Gérard Philippe (Don Rodrigue), Philippe Noiret, Jean Vilar (Le Roi Don Fernand), Roger Mollien, Georges Riquier...

Corneille, *Tite et Bérénice*, mise en scène par Patrick Guinand, Vieux Colombier, 1996, avec Claude Mathieu (Bérénice), Michel Favory (Tite),

Isabelle Gardien, Jean-François Rémi, Olivier Dautrey, Alain Lenglet, eric Génovèse, Coraly Zahonero (Domitie).

Racine, *Bajazet*, mise en scène par Eric Vigner, Vieux-Colombier, 1995, avec Bérengère Dautun (Zatime), Martine Chevallier (Roxane), Véronique Vella (Zaïre), Jean Dautremay (Acomat), Isabelle Gardien (Atalide), Eric Ruf (Bajazet), Alain Lenglet (Osmin).

Racine, *Bérénice*, mise en scène par Daniel Mesguish, (La Métaphore), Maison des Arts de Créteil, 1996, avec Sandy Boizard (Bérénice), Hervé Furic(Antiochus), Benoît Guibert (Arsace), Didier Kersten (Paulin), Matthieu Marie (Titus), Sarah Mesguich (Phénice).

Racine, *La Thébaïde*, mise en scène par Yannis Kokkos, Comédie-Française, 1995, avec Catherine Samie (Jocaste), Jean-Yves Dubois (Polynice), Michel Favory (Créon), Anne Kessler (Antigone), Redjep Mitrovitsa (Etéocle), Malik Faraoun, Alexandre Pavloff.

Racine, *Mithridate*, mise en scène par Daniel Mesguish, Vieux-Colombier, 1996, ace Simon Eine (Mithridate), Claude Mathieu (Monime), Olivier Dautrey (Arbate), Eric Génovèse (Xipharès), Samuel Le Bihan (Pharnace)...

Racine, *Phèdre*, mise en scène par Marie Bell, Théâtre du Gymnase-Marie Bell, 1965, enregistrement, ADES, Collection « Théatre », CD, avec Marie Bell (Phèdre), Jacques Dacqmine (Thésée), Henriette Barreau (Œnone), Raymond Gérôme (Théramène), Claude Giraud (Hippolyte), Danielle Volle (Aricie)...

Corneille, *Le Cid*, mise en scène par Thomas le Douarec, 1997, Théâtre Madeleine, Marigny.

TABLE DES MATIERES

INTRODUCTION **9**

I. DONNEES HISTORIQUES ET THEORIQUES **21**

I.1. BREVE HISTOIRE DE L'ALEXANDRIN 22

I.2. LE VERS FRANÇAIS : SYLLABIQUE OU SYLLABO-TONIQUE ? 26

I.2.1. Versification quantitative, accentuelle ou syllabique *26*

I.2.2. L'accent dans le vers français ? *28*

I.2.3. Une tentative de résolution empirique *33*

I.3. LA METRIQUE GENERATIVE AMERICAINE ET SES SUITES 37

I.3.1. Les apports de la théorie générative *37*

I.3.2. L'approche connexionniste *41*

I.4. APPLICATION DE LA THEORIE DU RYTHME DE PIERRE LUSSON A L'ALEXANDRIN ... 44

II. LE CORPUS **51**

II.1. PIECES DE CORNEILLE ET RACINE 52

II.2. PREPARATION DU CORPUS 55

II.2.1. Mise en forme du paratexte *56*

II.2.2. Reconstitution et correction des vers *58*

II.3. PREDOMINANCE DE L'ALEXANDRIN CHEZ CORNEILLE ET RACINE 62

III. LA RIME **69**

III.1. LA RIME SELON QUELQUES THEORICIENS 73

III.1.1. La fonction de la rime *73*

III.1.2. La rime : une identité sonore ? *75*

III.1.3. Principes d'identité et d'opposition *78*

III.1.4. Les nomenclatures de rimes *78*

III.2. EXAMEN EMPIRIQUE DE LA RIME 82

III.2.1. Identification des mots-rimes et de leur fréquence *83*

III.2.2. Constitution des paires de vers rimant ensemble *88*

III.2.3. Constitution de groupes et de réseaux de mots-rimes *92*

III.3. QUELQUES ELEMENTS SUR LE FONCTIONNEMENT DE LA RIME DANS LE VERS CLASSIQUE 98

III.3.1. Les principes d'équivalence et d'opposition *99*
III.3.2. Le rôle de la graphie *108*
III.4. LE ROLE DE LA CONSONNE D'APPUI 110
III.4.1. Hiérarchie des voyelles *112*
III.4.2. Principe de consistance graphique minimale *119*
III.4.3. Les morphèmes grammaticaux *121*
III.4.4. Incidence de la richesse lexicale des rimes *125*
III.4.5. Les glissantes dans la rime : diérèse et hiatus *128*
III.5. DEFINITION CONTEXTUELLE DE LA RIME PAUVRE, SUFFISANTE ET RICHE 133
III.6. CONCLUSION 136

IV. DIERESE, E "MUET", LIAISON : TROIS LIEUX DE VARIATION PHONOLOGIQUE 139

IV.1. LA DIERESE 145
IV.1.1. L'interprétation classique de la diérèse *150*
IV.1.2. Examen empirique de la diérèse chez Corneille et Racine *155*
IV.1.3. Tableaux récapitulatifs des règles de la diérèse *157*
IV.1.4. Formalisation *164*
IV.2. LA QUESTION DU *E* MUET 165
IV.2.1. Le décompte du e muet dans le vers *167*
IV.2.2. Formalisation *173*
IV.3. LA LIAISON : INCIDENCE SUR LE DECOUPAGE SYLLABIQUE 175
IV.3.1. La liaison dans le vers *177*
IV.3.2. Liaison avec / sans enchaînement ou forte /faible ? *180*
IV.3.3. Formalisation *183*
IV.4. SYNTHESE ET CONCLUSION 186

V. MARQUAGES POSSIBLES, MARQUAGES CHOISIS 193

V.1. POSITIONS METRIQUES : VOYELLES ET SYLLABES PHONETIQUES 197
V.2. MARQUAGE DES FINS DE MOT 203
V.3. NATURE SYNTAXIQUE DES POSITIONS 205
V.3.1. Quels marquages syntaxiques pourraient être envisagés ? *206*
V.3.2. Le marquage syntaxique choisi *210*
V.4. MARQUAGE "ACCENTUEL" 212
V.4.1. Existe-t-il un accent en français ? *214*

V.4.2. Marquages prosodiques dans le vers *218*
V.4.3. Une autre approche de l'accent, le marquage métrico-syntaxique *221*
V.5. SYNTHESE SUR LES MARQUAGES ADOPTES 222

VI. LE METROMETRE : L'OUTIL ET SES LIMITES 225

VI.1. OUTILS DE LINGUISTIQUE INFORMATIQUE 226
VI.1.1. L'analyse syntaxique par couche *226*
VI.1.2. La transcription graphème-phonème *228*
VI.1.3. Déroulement pratique *232*
VI.2. LES LIMITES ACTUELLES DU METROMETRE 237
VI.2.1. De l'écrit à l'oral : problèmes liés à l'orthographe *237*
VI.2.2. Les limites de l'analyse syntaxique *239*
VI.2.3. Les limites des marquages *252*
VI.2.4. Pas de règles pour l'identification d'un « bon » alexandrin *258*

VII. SON, SYNTAXE ET PROSODIE DE L'ALEXANDRIN « ORDINAIRE » . 263

VII.1. LES VOYELLES METRIQUES DANS LE VERS 264
VII.1.1. Répartition des phonèmes vocaliques *266*
VII.1.2. Variété vocalique dans les vers *280*
VII.2. FRONTIERES DE MOTS ET POSITIONS METRIQUES 284
VII.3. REPARTITION DES CATEGORIES SYNTAXIQUES 288
VII.3.1. Catégories syntaxiques et positions métriques *289*
VII.3.2. Mots-outils en fin d'hémistiches : quelques difficultés d'analyse syntaxique 296
VII.3.3. Conclusion *307*
VII.4. LA MARQUE ACCENTUELLE 309
VII.4.1. Le marquage accentuel des positions *311*
VII.4.2. Les profils prosodiques des hémistiches *316*
VII.4.3. Existe-t-il des régularités dans l'enchaînement des hémistiches ? *320*
VII.4.4. Comparaison avec d'autres résultats *323*
VII.4.5. Conclusion *328*
VII.5. CONCLUSION SUR LA FIGURE DU VERS CLASSIQUE 328
VII.6. APPENDICE SUR L'EVOLUTION DU VERS 332

VIII. LA CONTRIBUTION DU RYTHME A LA FORME POETIQUE 341

VIII.1. UNIVERS DE DISCOURS ET RYTHME : RETOUR SUR LES NOTIONS D'AUTEUR ET DE GENRE 352

VIII.1.1. Univers lexico-sémantiques et rythme *352*

VIII.1.2. Les genres : des indices convergents de spécificité *369*

VIII.2. D'*ANDROMAQUE* A *PHEDRE* : DES TRAGEDIES SEMBLABLES 380

VIII.2.1. Une grande proximité entre les tragédies centrales *380*

VIII.2.2. L'onomastique comme marqueur stylistique *383*

VIII.2.3. Rythme et parties du discours *386*

VIII.2.4. Rythme des pièces *392*

VIII.3. CONCLUSION 393

CONCLUSION 397

BIBLIOGRAPHIE 409

TABLE DES ILLUSTRATIONS

Figure 1. Fonctionnement du métromètre et représentation métrique 14
Figure 2. Le modèle de l'alexandrin classique à rimes plates alternées.....46
Figure 3. Niveaux métriques abstraits de l'alexandrin à rimes plates.....46
Figure 4. Courbe de répartition des mots-rimes.....86
Figure 5. Exemple de groupe de mots-rimes 94
Figure 6. Réseau de mots-rimes ous 98
Figure 7. Le réseau des mots-rimes en u.....115
Figure 8. Rôle de la consonne d'appui en fonction de la richesse lexicale des terminaisons.....118
Figure 9. Le réseau des mots-rimes en der-ler-ner.....122
Figure 10. Réseau des rimes en ent-ant.....124
Figure 11. Groupes de mots-rimes pour chaque voyelle.....126
Figure 12. Corrélation entre nombre de groupes et richesse lexicale 127
Figure 13. Groupe de mots-rimes en eux/ieux.....131
Figure 14. Richesse de la rime.....133
Figure 15. Le modèle de la syllabe d'après Kaye et Lowenstamm [1984].....140
Figure 16. Constituants syllabiques de école.....141
Figure 17. Part du vocabulaire affecté par la diérèse.....149
Figure 18. Recherche de vers contenant le groupement « ia » (extrait).....156
Figure 19. Tendance à la synérèse selon les glissantes.....164
Figure 20. Positions respectives des différentes voyelles [Arrivé et al, 1986, p. 514].....200
Figure 21. Fonctionnement du métromètre et représentation métrique 235
Figure 22. Répartition des positions métriques selon les phonèmes vocaliques.....268
Figure 23. Répartition des voyelles entre prose et vers 269
Figure 24. Répartition des e muets sur les douze positions 272
Figure 25. Répartition des voyelles métriques sur les douze positions.....275
Figure 26. Nombre moyen de phonèmes par position métrique 278
Figure 27. Répartition des vers selon le nombre de phonèmes différents.....280
Figure 28. Distribution des frontières de mots sur les douze positions 285
Figure 29. Répartition des alexandrins selon le nombre de mots 286

Figure 30. Mots pleins dans le vers 290
Figure 31. Répartition des catégories syntaxiques sur les douze positions.....294
Figure 32. Tableau des pronoms [Arrivé et al., 1986, p. 499] 300
Figure 33. La marque fondamentale sur les mille premiers vers de Bérénice [Roubaud, 1988, p. 49].....311
Figure 34. Répartition des « marques accentuelles » sur CORRAC 312
Figure 35. Poids accentuels obtenus avec la position 12 comme étalon de marquage 315
Figure 36. Marquage de Dinu [1993] et marquage du métromètre.....324
Figure 37. Évolution de la proportion de e muets dans le vers.....334
Figure 38. Répartition des fins de mots selon les positions 335
Figure 39. Proportion de mots-outils en sixième et douzième positions 336
Figure 40. Marquages accentuels les plus différenciés.....338
Figure 41. Arbre de la classification 354
Figure 42. Rythme et univers lexicaux chez Corneille et Racine 363
Figure 43. Profils lexico-sémantiques des pièces 366
Figure 44. Premier plan de l'analyse factorielle 368
Figure 45. Vocabulaire spécifique des différents genres 374
Figure 46. Tragédies versus comédie chez Racine 378
Figure 47. Répartition des types d'hémistiches selon les pièces 379
Figure 48. Champs lexicaux dans les œuvres de Racine 381
Figure 49. Profils lexico-sémantiques des œuvres de Racine 382
Figure 50 : Taux d'hémistiches réguliers par acte chez Racine.....392
Figure 51 : Taux d'hémistiches réguliers par acte dans les tragédies de Racine 393

TABLE DES TABLEAUX

Tableau 1. Répartition des types d'hémistiches [Gasparov, 1987, table 19, p. 356]....35
Tableau 2. Les pièces de Corneille....53
Tableau 3. Les pièces de Racine....54
Tableau 4. Règles de codage....57
Tableau 5. Répartition des vers de CORRAC selon le nombre de positions métriques....63
Tableau 6. Diversité métrique chez Corneille et Racine....63
Tableau 7. Diversité métrique selon les pièces....66
Tableau 8. Les mot-rimes les plus fréquents sur CORRAC....87
Tableau 9. Genre des terminaisons en ent....91
Tableau 10. Extrait de la liste des groupes de mots-rimes....95
Tableau 11. Extrait du RIMARIUM....96
Tableau 12. Voyelles cœur de rime....102
Tableau 13. Présence de la consonne d'appui pour les séquences V....113
Tableau 14. Présence de la consonne d'appui pour les séquences Ve....116
Tableau 15. Présence de la consonne d'appui pour les séquences Ves....117
Tableau 16. Présence de la consonne d'appui pour les séquences Vs....117
Tableau 17. Terminaisons de l'infinitif....121
Tableau 18. Les consonnes de liaison : réalisation, exemples....177
Tableau 19. Le traitement de la liaison....182
Tableau 20. Noyaux vocaliques des positions métriques....199
Tableau 21. Le traitement des mots composés....219
Tableau 22. Dérivations phonologiques....232
Tableau 23. Répartition des voyelles en fonction des positions métriques....271
Tableau 24. Répartition globale des catégories syntaxiques....290
Tableau 25. Répartition des catégories syntaxiques sur les douze positions....292
Tableau 26. Répartition des mots-outils en positions 6 et 12....306
Tableau 27. La marque de ponctuation sur les mille premiers vers de Bérénice [Roubaud, 1988, p. 52]....310
Tableau 28. Répartition des catégories syntaxiques en position 12....314

Tableau 29. Figures accentuelles des premier et second hémistiches.....319
Tableau 30. Les profils prosodiques les plus fréquents 319
Tableau 31. Figures des segments métriques sur les couples de vers 322
Tableau 32. Un échantillon plus « anapestique ».....325
Tableau 33. Figures des segments métriques sur les paires de vers.....327
Tableau 34. Répartition des vers selon les auteurs dans « TOT » 332
Tableau 35. Répartition des types d'hémistiches selon les auteurs 339
Tableau 36. Nombre de mots par vers selon les genres 372
Tableau 37. Types d'hémistiches selon les genres et les auteurs.....377
Tableau 38. Répartition des types d'hémistiches selon les univers de discours 388

TABLE DES ANNEXES (SECOND TOME)

ANNEXE I. **UNE APPLICATION DE LA THÉORIE DU RYTHME À L'ANALYSE DU VERS FRANÇAIS (PAR PIERRE LUSSON) 5**

ANNEXE II. **LE PROGRAMME DYNASTIE : DE LA SÉQUENCE GRAPHIQUE À LA SÉQUENCE MÉTRIQUE 13**

ANNEXE III. **SORTIES DU MÉTROMÈTRE 25**

ANNEXE IV. **STRUCTURE DE LA BASE DE VERS 61**

ANNEXE V. **SPÉCIFICITÉS DES POSITIONS MÉTRIQUES ET DES HÉMISTICHES 63**

ANNEXE VI. **ANALYSE D'AUTRES CORPUS (HUGO, BAUDELAIRE, MALLARMÉ, RIMBAUD…) 85**

ANNEXE VII. **MOTS-OUTILS EN SIXIÈME ET DOUZIÈME POSITIONS SUR 109 000 VERS DE CORNEILLE À MALLARMÉ 103**

ANNEXE VIII. **STATISTIQUE TEXTUELLE : UNE APPROCHE EMPIRIQUE DU SENS À BASE D'ANALYSE DISTRIBUTIONNELLE 123**

ANNEXE IX. **ANALYSE LEXICALE DE L'ENSEMBLE DES PIÈCES DE CORNEILLE ET RACINE 131**

ANNEXE X. **TRAGÉDIES CENTRALES DE RACINE : *D'ANDROMAQUE* À *PHÈDRE* 134**

ANNEXE XI. **LE RIMARIUM : DICTIONNAIRE ORGANISÉ DES RIMES CHEZ CORNEILLE ET RACINE 151**

RESUME en français :

Ce mémoire présente une approche empirique du rythme et de la rime du vers alexandrin classique. Un outil d'analyse systématique du vers, le métromètre, a été mis au point. Il transcrit le vers dans l'alphabet phonétique, le découpe selon les positions ou syllabes métriques, en respectant les règles de la versification (diérèse/synérèse, décompte du *e* muet et liaison) ; finalement, il attribue à chacune des positions un certain nombre de marquages ou étiquettes d'ordre exclusivement linguistique : noyau vocalique et syllabe de la position, catégorie morpho-syntaxique, situation dans le mot et présence d'une marque accentuelle sur la position. L'outil repose sur un analyseur syntaxique développé par P. Constant et un phonétiseur développé par F. Yvon. Les outils ont été adaptés aux particularités du vers.

Le métromètre a été utilisé pour analyser les 80 000 vers de Corneille et Racine. Il en résulte que la structure rythmique du vers est bel et bien le résultat d'un agencement spécifique des éléments de la langue (il n'y a pas « d'accent de vers »). Nous montrons que tous les marquages, qui relèvent de composantes *a priori* indépendantes de la langue, contribuent chacun à leur manière à définir la forme rythmique de chaque hémistiche et du vers. Les corrélations entre ces niveaux *a priori* indépendants semblent être constitutives du rythme.

Parallèlement, l'examen systématique de la rime chez Corneille et Racine a permis de montrer l'existence de régularités, non attestées dans les traités de métrique (la voyelle [e] a, par exemple, toujours besoin d'une consonne d'appui) et de proposer une définition contextuelle des concepts de rime riche, suffisante ou pauvre, lesquels varient selon le type de terminaison examiné.

Enfin, nous avons tenté de montrer en quoi la structure rythmique et rimique des vers pouvait contribuer à enrichir l'analyse stylistique : nous avons ainsi mis à jour des corrélations entre des univers de discours et des rythmes spécifiques.

TITRE en anglais :

Rhythm and rhyme in French Verse. Empirical study about 80 000 verses by Corneille and Racine.

DISCIPLINE :

Sciences du langage

MOTS-CLEFS :

Rythme, mètre, rime, métrique, versification, alexandrin, linguistique informatique, statistique textuelle, Corneille, Racine.

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE :

CELITH – EHESS
54, boulevard Raspail
75006 Paris