



**HAL**  
open science

# La musique des vers vient-elle de la répétition des sons? Ce qu'en dit la statistique.

Étienne Brunet

► **To cite this version:**

Étienne Brunet. La musique des vers vient-elle de la répétition des sons? Ce qu'en dit la statistique..  
Tous comptes faits, Ecrits choisis, tome III, Questions linguistiques, pp.125-144, 2016, 978-2-7453-3553-1. hal-01790570

**HAL Id: hal-01790570**

**<https://hal.science/hal-01790570>**

Submitted on 28 May 2018

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

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

## **La musique des vers vient-elle de la répétition des sons ? Ce qu'en dit la statistique**

Etienne Brunet, BCL(CNRS), Université de Nice Côte d'Azur

Les travaux de Jean Starobinski et Michel Arrivé ont fait connaître une tentative littéraire de Saussure, menée parallèlement à sa théorie linguistique et vouée à un succès moins glorieux. Saussure, pendant des années, s'est acharné à poursuivre l'**anagramme** dans la poésie latine, sans jamais parvenir à la preuve mathématique de ses intuitions. Rappelons que l'anagramme (le mot paragramme serait plus exact) serait un mot-thème (le plus souvent un nom propre, comme Apollon ou Aphrodite) qui inspirerait un poème en y répandant ses lettres, même dans le désordre, comme une signature invisible. Ainsi les deux hémistiches du vers saturnien :

*Donom AmPLOm victOr Ad mea temPLa pOrtatO*

contiennent les lettres constitutives du dieu APOLO. Saussure a beau multiplier les exemples, surtout chez Lucrèce, il échoue à montrer qu'ils sont significatifs et que le hasard ne peut les produire. Au surplus il se heurte à une objection majeure : personne avant lui, ni chez les poètes, ni chez les grammairiens, n'a eu le soupçon de ce processus supposé de production littéraire, auquel Saussure a fini par renoncer.

C'est tout le contraire lorsqu'il s'agit d'un procédé voisin : l'**allitération**, dont parlent sans cesse, sinon les poètes, du moins leurs commentateurs. Il n'est pas un manuel scolaire qui ne cite le vers de Racine :

*Pour qui sont ces serpents qui sifflent sur vos têtes ?*

comme exemple achevé de l'expressivité phonétique, fondée sur l'accumulation et le sifflement sinistre des s. Et de citer les F de Hugo dans :

*Un frais parfum sortait des touffes d'asphodèles,  
Les souffles de la nuit flottaient sur Galgala*

et les v de Verhaeren dans « *Le vent, le vent sauvage de novembre* ». Est-on sûr que ces répétitions de sons soient toujours conscientes et volontaires ? Dans bien des cas le doute est permis. Et l'hésitation s'observe aussi à la réception : un lecteur peut rester sourd à une allitération dans laquelle un autre lecteur voit un effet marqué. On sent bien que la poésie française, à qui manquent les ressources de l'accent, a partie liée avec la répétition, celle des rythmes et celle des sons. La musique répétitive de la rime – dont les poètes latins ou anglais peuvent se passer – lui est indispensable. Pourtant, si la répétition des sons est seule en cause, le hasard aussi pourrait être poète. Car avec un nombre limité de sons, la chaîne acoustique ne peut éviter le retour des mêmes éléments, comme celui des notes sur le clavier du piano.

## 1. Première expérience : calcul des cooccurrences

### 1.1. Le corpus

Pour en débattre, il convient de serrer le jeu et de confronter une œuvre entière à des expériences de laboratoire. Nous prendrons pour exemple un poète virtuose qui a poussé très loin le souci de l'harmonie ou de l'expressivité des sons : Émile Verhaeren. Le corpus n'est pas exhaustif mais il est très largement représentatif, avec 310 poèmes, distribués dans 23 recueils et échelonnés de 1883 (*Les Flamandes*) à 1917 (*Les flammes hautes*).

En 35 ans de production, une évolution se dessine dans l'œuvre du poète, comme en témoigne le graphique 1, fondé sur un calcul de distance intertextuelle<sup>1</sup>.

---

1. Il s'agit ici d'une méthode prônée par Charles Muller dès 1968 sous le nom de « connexion lexicale ». Elle s'appuie sur la loi binomiale, appliquée à toutes les classes de fréquences, et le résultat a une valeur en probabilité (un Chi<sup>2</sup>). Estimée lourde et complexe, elle n'a pas rencontré le succès immédiat. Pourtant son principe est plus solide et ses résultats plus probants que ceux des méthodes pratiquées jusqu'ici, celle de Jaccard et celle de Labbé. Comme le calcul de cet indice n'a rien de rédhitoire, nous l'avons intégré dans la version actuelle du logiciel HYPERBASE. Voir notre article « Muller le lexicomaître », in *Mélanges offerts à Charles Muller pour son centième anniversaire*,

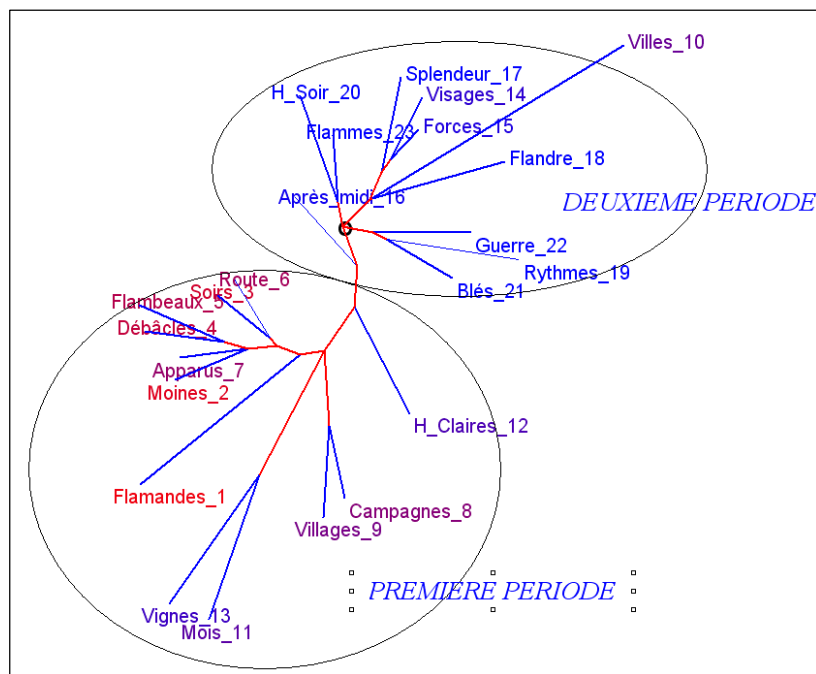


Figure 1. Analyse arborée de la distance intertextuelle (méthode Muller)

Un préalable se dresse à l'entrée : le texte écrit restitue les lettres et non les sons. Il a fallu soumettre le texte à un phonétiseur, que nous avons trouvé au LIMSI<sup>2</sup>. Nous n'avons proposé aucune retouche au programme, même si les *e* muets, dont raffole la poésie, n'ont pas eu un traitement parfaitement adéquat. On verra dans les figures 3 et 4 comment la chaîne graphique du premier poème *Cuisson du pain* a été transcrite. On aurait souhaité un code plus évident pour l'œil humain, celui de l'API. Mais la police appropriée posant des problèmes de transmission, de tri et de traitement à certains réseaux ou environnements

---

Conseil International de la langue française, Paris 2009, p. 99-119 (2009c) (NDÉ : article republié dans le tome II des *Écrits choisis*, chapitre 17).

2. Philippe Boula de Mareüil à qui nous avons transmis nos textes les a soumis au logiciel *Graphon* mis au point au LIMSI. Qu'il trouve ici l'expression de nos remerciements pour son obligeance et aussi pour sa prudence. Il nous a prévenu que *Graphon* n'était pas spécialement voué aux textes poétiques : il ne se préoccupe guère de la scansion des vers, omet de traiter les *e* muets, et reste indifférent aux liaisons entre les mots. Sans doute le *Métromètre* de Valérie Beaudouin eût été un meilleur choix, mais la commercialisation de ce logiciel n'est pas assurée.

informatiques, la lecture de nos résultats est rendue moins confortable, tout en restant non ambiguë. Précisons une fois pour toutes le codage un peu maladroit dont nous nous sommes servi et qui aidera à décrypter certains graphiques :

U	=ou (bout)
A	=an (banc)
I	=in (pain)
O	=on (rond)
H	=ch (chat)
W	=oi (bois)
Ï	=yod (piller)
Û	=ui (puis)
E	=eu ouvert (peur)
@	=eu fermé (peut)
O	=o ouvert (port)
Ö	=o fermé (peau)
Q	=gn (bagne)

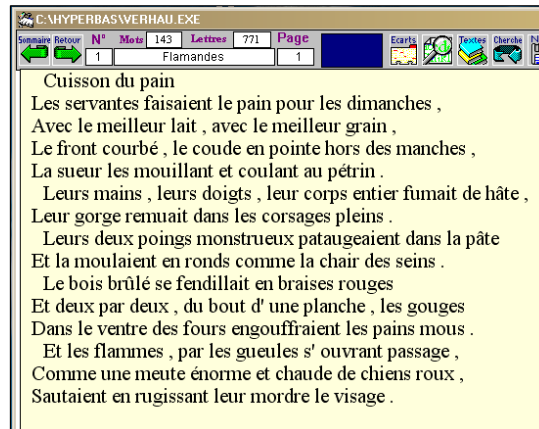


Tableau 2. Codage phonétique

Figure 3. Le premier poème : le sonnet du pain

## 1.2. Les phonèmes

Oublions les graphèmes, mais aussi les mots et les blancs qui les séparent et considérons le poème comme une chaîne continue de phonèmes indépendants (figure 4). Comme les mots ont disparu, il n'y a plus ni syntaxe, ni sémantique, mais un continuum acoustique, seulement ponctué par la prosodie et les marques de fin de poème et de fin de recueil. Par voie de conséquence l'allitération n'est plus identifiable, si on l'entend dans son sens restreint, comme répétition des mêmes consonnes à l'initiale des mots. On l'interprétera au sens large, comme répétition des sons, sans distinguer les consonnes et les voyelles, et sans se préoccuper de la barrière des mots.

Or à partir de ces éléments minimaux un calcul de distance (le même que précédemment) montre qu'une évolution se produit chez Verhaeren dans l'emploi des phonèmes et que la suite chronologique est très reconnaissable dans l'analyse arborée de la figure 5 : les 10 derniers recueils se détachent nettement des 13 premiers. Et le dixième recueil, *Villes*, qui avait déserté son camp dans la figure précédente, rentre ici dans le rang, comme si la chronologie était plus contraignante au niveau des phonèmes qu'au niveau des mots.

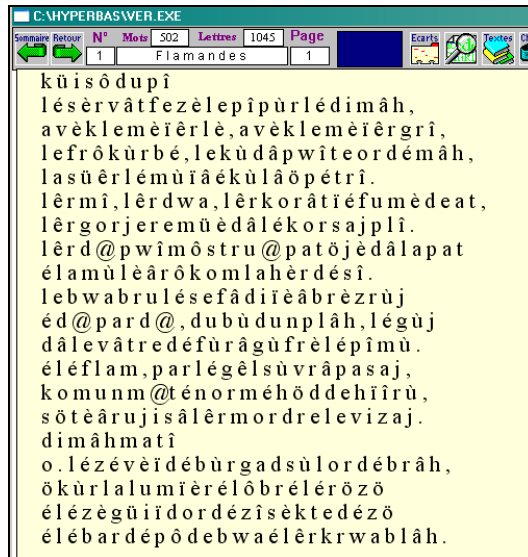


Figure 4. Une chaîne de phonèmes indépendants

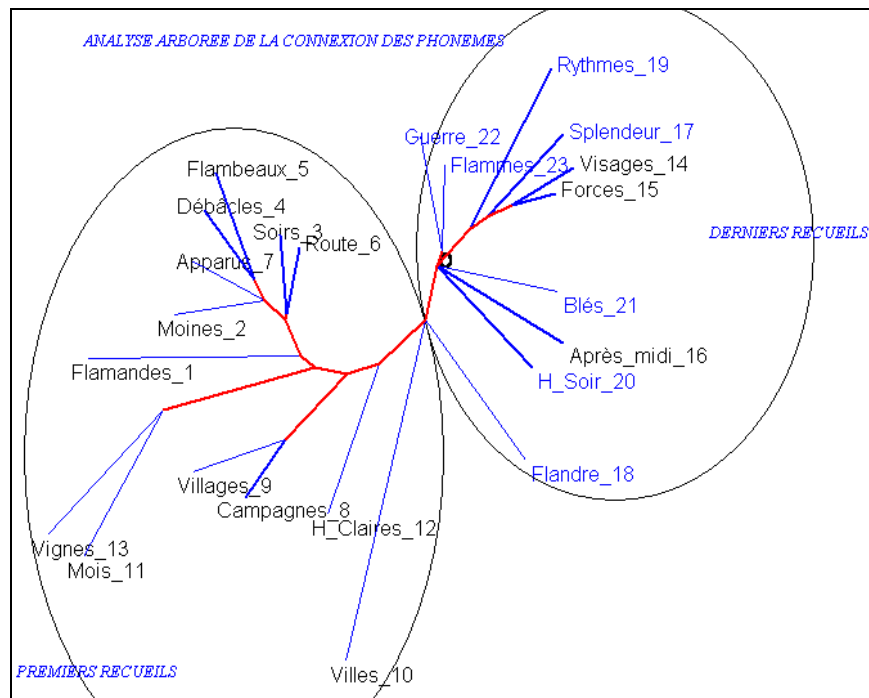


Figure 5. Analyse arborescente des textes fondée sur la connexion des phonèmes

L'évolution constatée dans le dosage des phonèmes reproduit celle qu'on observe parmi les mots, sous leur forme graphique ou phonétique. Le parallélisme des mots et de leurs éléments constituants laisse ainsi en suspens un dilemme bien connu : Où est la cause ? Où est l'effet ? Le mot vient-il parce que le son l'appelle, comme cela se passe à la rime ? Ou bien se présente-t-il d'abord, pressé par le sens et la syntaxe, en imposant les sons qu'il contient ? Sans doute doit-on imaginer un compromis, avec des candidats et une sélection : certains écrivains, même des prosateurs, ont évoqué cet examen de passage, cet oral qui contrôle l'écrit, et qui passe par la bouche – ce que Flaubert appelait le gueuloir.

### 1.3. Les phonèmes adjacents

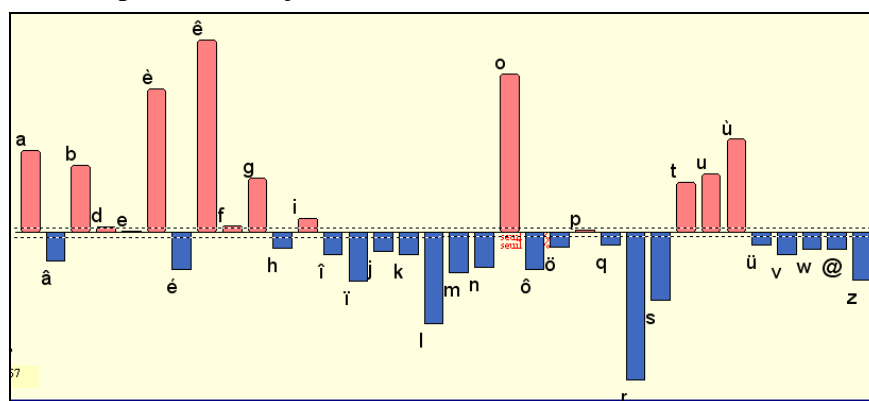


Figure 6. Le voisinage immédiat du phonème r

Or lors de cet examen la répétition immédiate des sons n'est pas ce qu'on recherche mais ce qu'on bannit. Un programme de cooccurrences permet de le démontrer, en analysant chaque phonème dans chacun des vers du corpus et en le comparant au phonème suivant. Les cas où le même phonème est répété ne sont pas rares si le hasard distribue les sons. Mais dans la réalité du corpus ces séquences répétitives sont évitées. Il en est ainsi du phonème le plus souple du système français, le phonème r, qui peut être suivi ou précédé immédiatement de n'importe quelle voyelle, mais aussi de beaucoup de consonnes. Le graphique 6 montre la préférence du r pour le voisinage des voyelles a, e, eu ouvert, i, o ouvert, u, ou et des consonnes b, d, f, g, t. La cooccurrence la plus rare, c'est celle du même au même, celle où deux phonèmes successifs auraient le même son (ici le son r). Le profil du phonème l est à peu près le même. Mais en

dehors des liquides (et des semi-consonnes) les camps sont mieux marqués : les voyelles refusent d'être collées à une autre voyelle et les consonnes à une autre consonne. Mais partout la répugnance la plus vive est celle qui se manifeste avec un phonème identique.

On comprend aisément cette difficulté de la poésie à admettre des phonèmes immédiatement répétés : la langue elle-même tend à l'interdire. Le dictionnaire, sauf celui des noms propres, fournit peu d'exemples du type *Aaron, agréé*. Encore s'agit-il alors de voyelles. Les consonnes redoublées que le français multiplie dans l'écriture se simplifient dans la transcription phonétique et le phonétiseur du LIMSI n'en fait pas état, sauf, curieusement, dans certains cas du L redoublé.

#### 1.4. Décalage de 1, 2, n

Mais si l'on décale le processus, en admettant un élément intercalaire entre deux cooccurrences d'un même phonème, la langue ne s'oppose plus au redoublement du type *midi* ou *maman*. Mais la poésie, au moins celle de Verhaeren, fait sur ce point une distinction entre consonnes (ou semi-consonnes) et voyelles. Le tableau 7 détaille les effectifs et les écarts obtenus dans le corpus pour les redoublements de phonèmes repérés à la distance 2.

<i>Voyelles</i>			<i>Consonnes et semi-consonnes</i>		
A	2916	13.4	B	68	-6.0
AN	906	13.5	D	690	-18.7
E	760	-8.5	F	22	-11.0
E	3718	33.8	G	10	-3.9
EU(R)	6	-7.1	CH	28	-9.6
I	1178	14.3	YOD	112	-1.8
IN	196	7.4	J	28	-5.6
O ouvert	314	2.2	K	226	-16.5
ON	194	2.6	L	1508	-31.7
O fermé	148	1.1	M	780	10.4
U	260	5.2	N	208	-4.8
OU	410	13.9	P	182	-8.8
EU(X)	4	-2.7	GN	0	-0.7
E	1414	16.4	R	1656	-46.7
			S	1222	-6.6
			T	1472	15.9
			UI	0	-3.3
			V	358	4.2
			W (oi)	28	-9.0
			Z	76	-2.6

**Tableau 7. Phonèmes redoublés (avec un élément intercalé)**



À part deux exceptions (E et EU fermé) toutes les voyelles ont de forts excédents alors que le déficit est constant parmi les consonnes, sauf M, T et V. Si l'on allonge la distance entre deux occurrences du même timbre (avec 2, 3, 4 phonèmes intermédiaires ou davantage), les écarts finissent par s'estomper entre phonèmes identiques ou différents. À la distance 4 les répétitions d'un même phonème ont un excédent dans 12 cas sur 34 (notamment pour les consonnes K, L, M, N, T, S, R, Z). À la distance 6, on en compte 18. Il faut donc un certain espace pour que l'allitération puisse se déployer, surtout s'il s'agit de consonnes.

Mais l'alternance entre distances paires et impaires se maintient très longtemps, au moins jusqu'à six éléments intercalés. Quand on a 2, 4, 6 phonèmes intermédiaires, le profil cooccurentiel du phonème retenu est le même que celui de la figure 8, où les excédents sont ceux de la catégorie opposée (ici les voyelles). Inversement quand interviennent 3, 5, 7 éléments entre deux occurrences, le phonème étudié a des déficits dans la catégorie opposée, à l'image de la figure 9, qui représente l'image inversée de la figure 8 et où un seul phonème sépare deux phonèmes cooccurrents.

Précisons que l'ordinateur ignore complètement la différence entre voyelles et consonnes. Si les figures 8 et 9 illustrent une opposition radicale entre les deux espèces, ces deux espèces n'interviennent pas en tant que telles dans le calcul, qui ne connaît que 34 phonèmes sans distinction d'aucune sorte. C'est bien le moins, pensera-t-on, que l'analyse statistique reconnaisse cette distinction qui est le fondement de tout système linguistique articulé. Mais imaginons qu'on ait affaire à une bande sonore recueillie auprès d'une peuplade inconnue. Cette simple partition proposée par le calcul serait d'un grand secours pour le déchiffrement. Pour mettre à l'épreuve la perspicacité de la machine, ajoutons à la distribution des cooccurrents immédiats de *s* (figure 8), celle des 33 autres phonèmes observés dans les mêmes conditions. On obtient un tableau de contingence symétrique, car on a neutralisé l'ordre des éléments et rendu équivalentes les paires *ab* et *ba*. Sur la diagonale, où le même phonème figure en ligne et en colonne, apparaît l'effectif des répétitions, soit  $A_{AA} 62$ ,  $A_N 86$ ,  $B 2$ ,  $D 92$ ,  $E 0$ ,  $E 16$ , etc., dans l'exemple ci-dessous (tableau 11) qui restitue un extrait partiel du tableau.

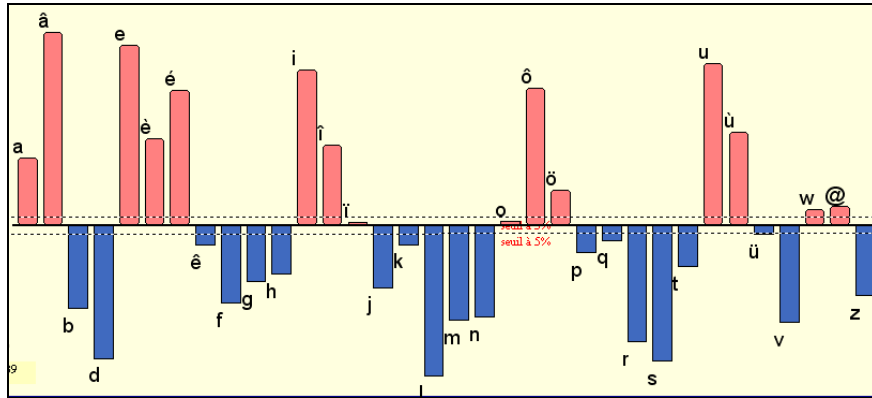


Figure 8. Cooccurrences immédiates du phonème s (sans élément intercalé)

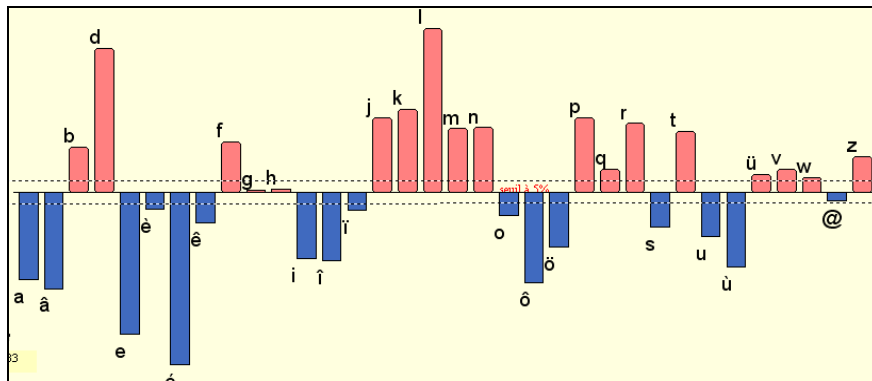


Figure 9. Cooccurrences du phonème s (avec 1 élément intercalé)

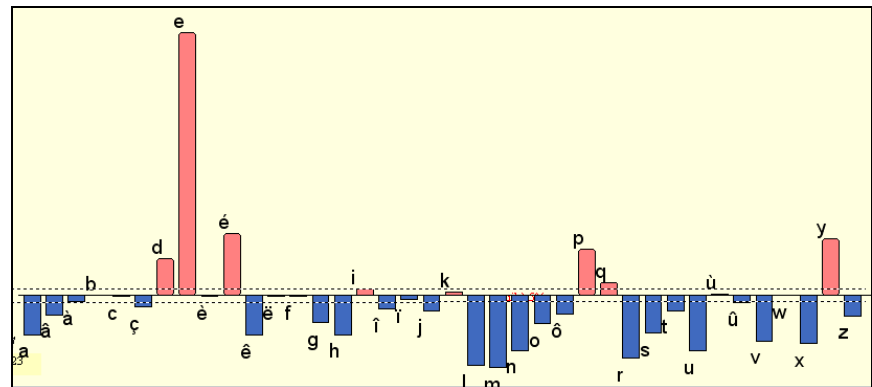


Figure 10. Cooccurrences immédiates du graphème s

**Tableau 11. Tableau de contingence de la cooccurrence immédiate des phonèmes**

	A	A (AN)	B	D	E	E	E	EU(R)	F	G	etc.
A	62	106	1419	1159	26	102	274	2	774	574	
A (AN)	106	86	587	2694	2	90	419	1	386	218	
B	1419	587	2	23	513	445	829	13	3	3	
D	1159	2694	23	92	4786	901	5069	251	32	15	
E	26	2	513	4786	0	20	13	3	832	362	
E	102	90	445	901	20	16	102	2	652	287	

etc.

Les figures 12 et 13 juxtaposent les résultats de la même synthèse réalisée successivement sur 34 phonèmes et 38 graphèmes à partir du même corpus. On voit que le code écrit donne une image dégradée du système. Alors que la division voyelle/consonne est clairement circonscrite parmi les phonèmes (les semi-consonnes formant un groupe intermédiaire, mollement rattaché aux consonnes), divers phénomènes bien connus contribuent à obscurcir cette opposition dans l'écriture où l'étymologie joue un rôle peu discret, comme aussi la syntaxe (désinence des verbes, pluriels). Si dans le graphique 13 l'orientation générale reste perceptible (les consonnes en haut, les voyelles en bas), les particularités de la transcription écrite se reconnaissent dans les graphies *QU* ou *CH* ou se devinent dans les combinaisons où interviennent les voyelles non accentuées. Dans le brouillard du code de l'écriture, on a peine à retrouver les groupements fins que l'analyse distinguait parmi les phonèmes : chuintantes, nasales, labiales. Si la loi des grands nombres arrive encore à percer la brume quand on envisage le tableau entier, la lisibilité devient incertaine quand la focalisation s'exerce sur une ligne, par exemple celle de la lettre *s* illustrée par la figure 10. Alors que dans la figure 8 le phonème *s* manifestait une aimantation radicale pour une voyelle précédente ou suivante, le graphème *s* n'est plus lié qu'aux voyelles *E* et *E* et des consonnes apparaissent parmi les cooccurents privilégiés. Si donc la poésie s'adresse à l'oreille, il y a peu à attendre de l'étude des graphèmes. Revenons donc aux phonèmes.

### 1.5. Cooccurrences généralisées

Envisageons cette fois tous les vers du corpus, tous les phonèmes, et toutes les places où peut se produire une cooccurrence, du début à la fin du vers. Ce faisant, nous généralisons les procédures, la répétition n'étant qu'un cas particulier de la cooccurrence, et la rencontre des phonèmes adjacents un cas particulier de cette cooccurrence générale.

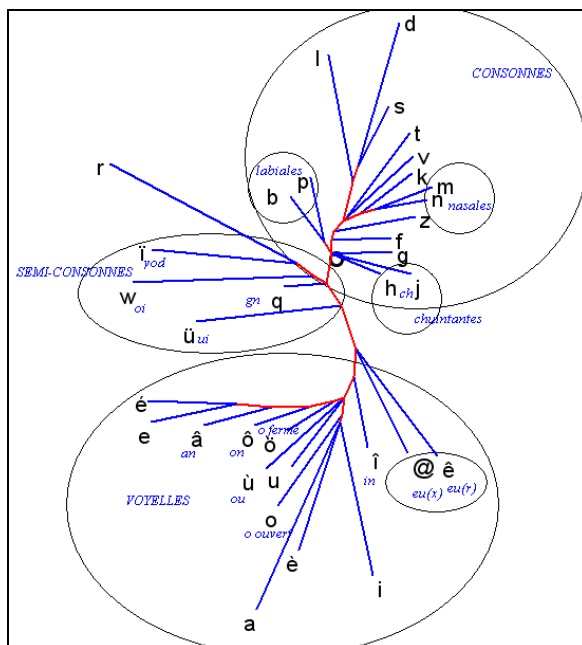


Figure 12. Analyse arborée du tableau des phonèmes immédiatement cooccurrents

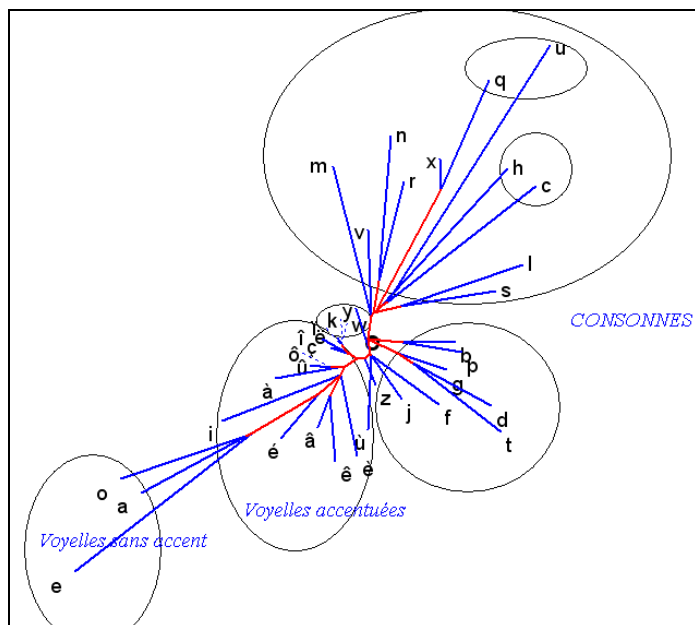


Figure 13. Analyse arborée du tableau des graphèmes immédiatement cooccurrents

Les effectifs sont alors multipliés puisque, pour 30 phonèmes par vers en moyenne, on compte  $(30 \times 29)/2 = 435$  combinaisons. Voici les trois premières lignes du tableau des cooccurrences :

```
28176 15561 8043 22246 17483 20369 28595 4768 7646 3717 3283 16782
6049 7031 6067 15989 37309 14509 10360 9486 7341 6845 11979 546
44608 24205 20019 7995 8983 1889 11641 8110 2310 5978 , 445918 a
15561 8536 4407 13312 9967 10818 15720 2682 4009 2174 1937 9571
3558 3799 3164 8360 19885 8314 5692 5670 4016 3686 6231 241 22909
14304 11916 4393 5034 1182 6051 3061 1329 3198 , 244687 â
8043 4407 1650 5774 4790 4912 7654 1150 1668 724 830 4211 1714 1813
1224 3599 9554 3127 2418 2709 2298 2182 2118 102 11017 5577 4536
2344 2491 520 2239 1319 660 1576 , 110950 b
```

Précisons ce qu'on entend ici par cooccurrence<sup>3</sup> : dans le cadre du vers un balayage de gauche à droite fixe le premier élément auquel est comparé le second élément repéré par un second balayage qui commence immédiatement à droite du premier. Si donc un vers compte 4 fois le phonème A, on enregistrera les rencontres successives :  $a_1-a_2$ ,  $a_1-a_3$ ,  $a_1-a_4$ ,  $a_2-a_3$ ,  $a_2-a_4$  et  $a_3-a_4$ , soit 6 cooccurrences sur un total de  $(n^2-n)/2$  cooccurrences ( $n$  étant la taille du vers exprimée en nombre de phonèmes). En cumulant tous les vers on obtient le tableau qui précède.

La carte des phonèmes obtenue en fin de compte fixe l'architecture vocalique et consonantique du système français, au moins tel qu'il apparaît dans la conscience et l'usage d'un poète (figure 14).

Là encore la diagonale totalise les répétitions observées dans les vers du corpus : A 28 176, AN 8 536, B 1 650, etc. Comme on peut rapporter cet effectif au total des cooccurrences pour chaque phonème (respectivement 445 918, 244 687, 110 950), on a le moyen de mesurer la part que joue la répétition dans la musique des vers. La conclusion est surprenante et négative. À presque tous les degrés de la gamme des sons, le déficit s'installe largement :

A	-11.7	AN	-5.7	B	-5.0	D	-15.3	E	-8.8	E	-14.1	E	-4.1
EU(R)	7.5	F	-3.2	G	-3.5	CH	6.0	I	-2.7	IN	1.6	YOD	-5.7
J	-0.1	K	-11.0	L	-22.1	M	-3.2	N	-1.0	O	-2.5	ON	-2.7
O fermé	-2.2	P	-7.9	GN	2.2	R	-33.4	S	-10.9	T	-3.6	U	-0.2
OU	7.5	UI	0.4	V	-2.7	OI	-8.4	EU(X)	1.0	Z	0.5		

3. En d'autres circonstances la cooccurrence peut être entendue comme une coprésence, lorsqu'on comptabilise l'apparition simultanée de deux objets linguistiques dans le segment considéré, sans se préoccuper de la répétition éventuelle de l'un ou de l'autre. En procédant ainsi, le calcul met en relief les associations privilégiées et dissymétriques où un phonème n'apparaît qu'en présence d'un autre, ce qui est le cas des demi-consonnes dans les associations : w et A, (BOIS), w et I (COIN), î et @ (CIEUX), ü et I (LUI), î et E (PIERRE), î et E (MEILLEUR).

Quelques sons toutefois sont épargnés : EU(R), EU(X), CH, GN, OU, UI. Nous allons voir dans une seconde approche que ce privilège leur est maintenu, sans qu'on sache pour l'instant ce qui leur a valu cette grâce : choix du poète ou tendance du genre poétique.

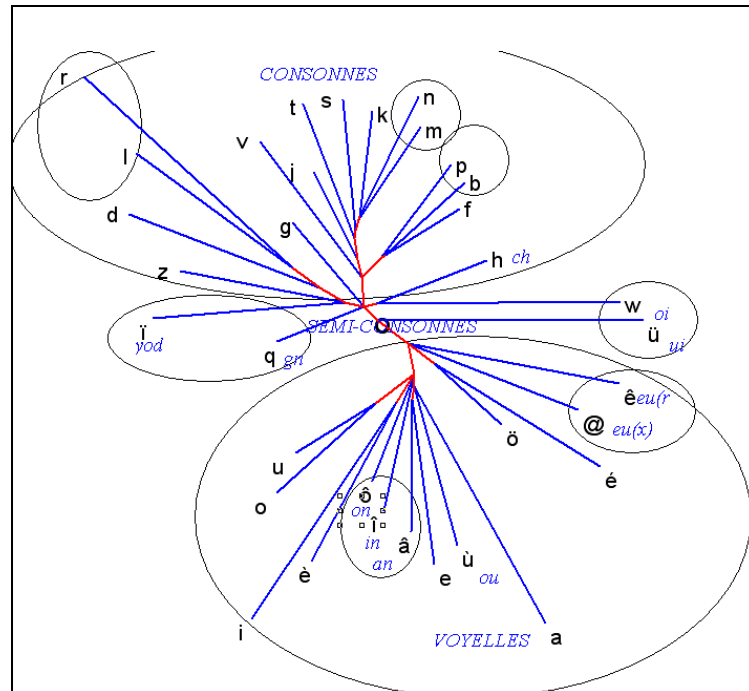


Figure 14. Analyse arborée du tableau des cooccurrences généralisées

## 2. Deuxième expérimentation : le hasard poète

L'examen qui précède est *a posteriori*. Le hasard n'est invoqué qu'à la fin du jeu, dans un rôle de juge. Mais on pourrait lui donner un rôle actif, de simulation, au cours du processus de composition. Voyons donc ce que ferait un poète mécanique qui, comme certains surréalistes, épingle au hasard, dans le dictionnaire, les mots de son texte. Il y aurait nécessairement des sons répétés. Plus ou moins que chez les vrais poètes ?

### 2.1. L'observation

On a mis en couleur les répétitions de sons qui sont significatives, compte tenu de la taille du vers et de la fréquence des sons dans le

corpus. Si tous les sons avaient la même fréquence, dans un vers où 30 places sont disponibles, on aurait à peu près un représentant de chacune des 34 variétés de phonèmes. Mais comme le son *a* est très fréquent, sa répétition n’a rien de surprenant et l’on en compte jusqu’à quatre dans le même vers sans ébranler l’hypothèse nulle et justifier la couleur. Le rouge est requis lorsque l’écart est patent ( $z > 4$ ), le magenta notant des écarts plus faibles ( $z$  entre 3 et 4). Au total sur les 370 phonèmes du premier poème représenté ci-dessous (figure 15), 19 révèlent un redoublement marqué.

The screenshot shows the HYPERBASAPHONEM.EXE application window. The title bar reads "C:\HYPERBASAPHONEM.EXE". The menu bar includes "Sommaire", "Retour", "N°", "Mots", "144", "Lettres", "538", and "Page". The toolbar contains various icons for navigation and analysis. The main window is divided into two panes. The left pane displays the original text of a poem in Flemish, with words highlighted in red and magenta. The right pane displays the same text with words in a random order, also with red and magenta highlights. On the right side of the window, there are two small data tables. The top table shows the following data: b 3, h 2, e 5, e 5, u 3, @ 4. The bottom table shows: 370, 22, 0.06. A second table below it shows: 370, 19, 0.05.

Figure 15. Observation et hasard : à gauche ordre des mots normal, à droite ordre aléatoire

## 2.2 Simulation aléatoire

Mais qu’aurait fait le hasard avec les mêmes éléments : mêmes mots, mêmes phonèmes dans le poème, mêmes proportions des phonèmes dans le corpus, même nombre et même taille des vers, en changeant seulement aléatoirement la place des mots, et donc des phonèmes dans le poème, comme fait Monsieur Jourdain avec la *marquise* et ses *beaux yeux*. Ce tirage au sort détruit naturellement la syntaxe et la sémantique et il faudrait bien de la patience à un oulipien, et de multiples essais, pour obtenir un sens. Mais les faits de répétition, soumis aux mêmes tests, se montrent aussi fréquents que dans le texte original : on en compte même un peu plus : soit 22 contre 19.

Or ce qu'on observe dans ce premier poème se reproduit plus largement dans le recueil et même dans le corpus entier. Résultat surprenant : si l'allitération est un procédé cher à la poésie, on s'attendrait qu'elle y soit plus fréquente que dans le simple jeu du hasard. Comment expliquer que sur 288 828 phonèmes relevés dans le corpus Verhaeren, les répétitions de phonèmes dans le même vers (effectif de 165 316, soit 57 %) soit moins nombreuses que dans une distribution aléatoire où les mêmes mots seraient déplacés au hasard dans le poème (effectif 173 216, soit 60 %). Plutôt que de renoncer trop vite à l'allitération volontaire, on invoquera une force contraire qui s'oppose au rapprochement des semblables. Pour des raisons syntaxiques, sémantiques et parfois seulement stylistiques, la phrase répugne à utiliser deux fois de suite, ou à courte distance, le même mot ou deux mots de la même famille. Ainsi on évitera de dire : « la figure de de Gaulle » ou « allonger la longueur ». Le hasard, lui, n'a pas de ces pudeurs et ne s'attache pas à cultiver la variété ou à bannir la cacophonie. Pareillement en peinture l'artiste tend à éviter les télescopes trop brutaux, ton sur ton, rouge sur rouge, noir sur noir, ce qui ne l'empêche pas de proposer des rappels de teintes. Tout est affaire de distance : trop rapprochée la répétition est perçue comme un heurt, une incivilité, trop éloignée elle cesse d'être sensible et perçue.

### 2.3. Allitérations et assonances

On note dans les relevés que l'allitération intéresse moins les consonnes que les voyelles et semi-voyelles. De là est née l'idée d'un calcul séparé pour ces deux groupes, d'où il ressort que les voyelles et semi-voyelles se prêtent mieux que les consonnes aux effets de répétition. Dès que l'écart réduit franchit le seuil de 2, l'effectif observé (20 807) des voyelles répétées dépasse celui du hasard (20 207). Peut-être conviendrait-il de réactiver la distinction traditionnelle qui survit encore dans beaucoup d'esprits et dans les définitions des dictionnaires entre l'allitération et l'assonance. La première concernerait les consonnes, quand la seconde s'intéresserait aux voyelles. S'y ajoute une différence de position : l'allitération porte sur l'attaque du mot, la seconde sur la finale. Il n'est pas certain que l'oreille fasse la distinction quand elle perçoit des échos dans le vers. C'est comme si l'on devait distinguer à l'écoute les touches blanches et les touches noires du piano. Mais l'allitération n'est pas indifférente à l'accent des mots, et son effet est beaucoup plus sensible quand il porte sur une syllabe accentuée. Il convient donc de rétablir la barrière qui sépare les mots, ce qu'on vient de faire dans cette seconde approche. Sans doute aussi pourrait-on distinguer



les rencontres intra et interlexicales. On a vu que la langue répugne à la répétition immédiate du même son, sur le modèle CREE. Mais dans la combinatoire du discours, des rencontres peuvent se produire où la finale d'un mot a le même son que l'initiale du mot suivant. Sans doute les principes de versification sont hostiles au heurt ainsi produit, au moins lorsqu'il s'agit de voyelles. Mais rien n'interdit des séquences du type *la mer recommencée* ou *toujours recommencée*. Le tableau 16 fait le relevé de telles rencontres qu'elles soient à l'intérieur des mots (colonne intra) ou à la jonction de deux mots (colonne inter). Les exemples internes sont proscrits. On n'en trouve aucun qui concerne une voyelle. Les rares cas que l'on relève parmi les consonnes ou semi-consonnes sont liés aux options contestables du logiciel *Graphon*, soit que l'effacement des E muets ait provoqué des télescopages injustifiés du R dans ENTERR(E)RA et IMPLOR(E)RA, ou du T dans AGIT(EN)T-ILS OU MONT(EN)T-ILS, soit qu'on ait affaire à la consonne complexe X (dans EXSANGUE) ou au yod redoublé dans VOYIONS, GROSEILLER OU MARGUILLER. Au total ces bizarreries se comptent sur les doigts, n'était le fort effectif des L redoublés (du type ILLUMINE) qui n'ont pas vraiment leur place ici<sup>4</sup>. Quant aux répétitions externes ou de position, leur nombre 798 paraît élevé pour un poète dont l'oreille est si sensible et les principes si rigoureux. En réalité la plupart viennent d'un artefact introduit par le phonétiseur : en ignorant un E muet intermédiaire, le logiciel a rapproché des sons qui n'étaient pas immédiatement contigus. Ainsi le groupe LA MER(E) REBELLE a été traité comme LA MER REBELLE et cela fausse les résultats observés pour les consonnes D, K, L, M, N, R, S, T, mais aussi pour les voyelles I et E.

Ce traitement inadéquat n'empêche pourtant pas une conclusion certaine : les répétitions immédiates sont pourchassées par le poète. Quand au contraire le hasard entre seul en jeu, il leur fait bon accueil et les effectifs recueillis dans la colonne aléatoire sont trois fois plus élevés. Il en est ainsi plus largement des hiatus ou rencontres de voyelles différentes sans consonne intermédiaire. Là aussi le traitement est approximatif, car les liaisons n'ont pas été prises en compte par le phonétiseur en dehors des s du pluriel, ce qui gonfle indûment l'effectif des hiatus. Mais ce gauchissement s'exerçant pareillement dans les données réelles et aléatoires, la disproportion demeure entre la

---

4. On peut imaginer que cette exception (le phonétiseur a simplifié à juste titre les autres consonnes redoublées de l'écriture) est due au souci de distinguer dans la graphie ILL les cas où l'on entend un L (ILLUSTRE) et ceux où on entend un yod (FILLE).

répugnance du poète et le laxisme du hasard (2 656 et 9 366 exemples respectivement).

**Tableau 16. Les redoublements intra et interlexicaux. Les hiatus**

	<i>Répétitions immédiates</i>				<i>Hiatus</i>				
	<i>CONSONNES et SEMI-CONSONNES</i>				<i>VOYELLES</i>			<i>VOYELLES</i>	
	<i>inter</i>	<i>intra</i>	<i>inter(hasard)</i>		<i>inter</i>	<i>intra</i>	<i>inter (hasard)</i>	<i>observé</i>	<i>aléatoire</i>
B	1	0	4	A	29	0	247	344	1211
D	45	0	183	A	43	0	131	440	1067
F	2	0	6	E	0	0	2	105	1896
G	2	0	1	E	179	0	721	665	2664
H	0	0	4	E	8	0	32	310	526
J	0	0	14	E	0	0	0	3	2
K	22	0	96	I	12	0	38	451	737
L	160	46	554	I	3	0	43	201	540
M	16	1	68	ö	7	0	73	132	576
N	14	0	38	o	0	0	2	1	26
P	1	0	4	o	2	0	11	2	111
Q	0	0	0	U	2	0	4	2	4
R	144	2	205	@	0	0	1	0	1
S	71	4	179	U	0	0	5	0	5
T	23	2	85						
V	5	0	12						
Z	7	0	11						
W	0	0	0						
ï	0	4	1						
ü	0	0	0						
				<i>Total répétitions (consonnes + voyelles)</i>				<i>Total hiatus</i>	
				<i>Total</i>	798	59	2775	2656	9366

## 2.4. Les allitérations volontaires

Reste à poser des filtres pour décanter les redoublements fugitifs et les allitérations insistantes. L'écart réduit qui mesure cette insistance avait servi à mettre en relief (en rouge ou magenta) les rencontres de la figure 15. À l'échelle du corpus, le tableau 17 montre le renversement qu'on observe lorsque le filtrage est plus ou moins sévère.

**Tableau 17. Nombre de redoublements en fonction du seuil choisi pour les retenir**

	<i>aucun seuil</i>	<i>z &gt; 2</i>	<i>z &gt; 3</i>	<i>z &gt; 4</i>	<i>z &gt; 5</i>
<i>hasard</i>	173 216	40 126	11 856	3 334	898
<i>observation</i>	165 316	38 724	11 672	3 621	1 023

Quand le seuil n'est pas pris en compte, ou quand il est faible, le hasard génère la répétition plus souvent que le poète. Mais là où le seuil est le plus sévère (à partir d'un écart réduit supérieur à 4), le hasard perd

pied devant l'observation, tandis que s'affirment les choix volontaires du poète.

On n'insistera pas sur les sons que les itérations privilégient, car on ne sait si ces préférences sont propres à Verhaeren ou si d'autres poètes (de la Belgique, de la campagne ou de la campagne belge) les partagent. Parmi les premiers de la liste quelques phonèmes se distinguent : EU ouvert 59, EU fermé 43, CH 44, UI 28, J 17, G 16, GN 14. Ce sont ceux que distinguait notre première approche. La liste 18 relève certains exemples qui franchissent le seuil de 5.

**Tableau 18. Quelques exemples d'allitération relevés par le calcul ( $z > 5$ )**

<i>foyer fourbi de naphte et de phosphore</i> (F),	<i>les chercheurs de chimères</i> (CH),
<i>chaque pêcheur pêche pour soi</i> (CH et P),	<i>Bleu comme le bleu des cieux</i> (EU),
<i>si vos bois, vos vallons, vos plaines et vos grèves</i> (V),	<i>à suivre un vol d'oiseau de suie</i> (UI),
<i>ses longs chemins jonchés de branches mortes</i> (CH),	<i>courbe et ligne maligne</i> (GN),
<i>et ses rancœurs et ses douleurs et ses erreurs</i> (EUR),	<i>l'acharnement à mieux chercher</i> (CH),
<i>bataille d'ombre et d'or, bouge dans les ténèbres</i> (B),	<i>regard d'orgueil, regards de proie</i> (G),
<i>d'heure en heure, à neuf, leurs appétits</i> (EUR),	<i>ce Londres de fonte et de bronze</i> (ON),
<i>et des gueules d'égout engloutissant la nuit</i> (G),	<i>les gens branlant, les gens bêlant</i> (AN),
<i>de grands bœufs noirs beuglaient</i> (B et G),	<i>gel, givre et neige</i> (J),
<i>l'heure qui meurt sur les demeures</i> (EUR),	<i>langue sèche, regards aigus</i> (G),
<i>d'autres regrets que des regrets fugaces</i> (G),	<i>sans feu ni lieu</i> (sans foi ni loi),
<i>et leur joie et leur rire et leurs cris et leur fièvre</i> (R),	<i>se pose au sommet des roseaux</i> (Z),
<i>de mes marches, de roche en roche</i> (CH),	<i>les horloges que j'interroge</i> (J)
<i>les chiens du noir espoir ont aboyé ce soir</i> (OI),	<i>la vierge, de berge à berge</i> (J), etc.

On remarquera que l'allitération est parfois obtenue à bon compte, par l'effet d'une répétition de type rhétorique, le cas limite étant le vers :

*Vivre et vivre et vivre et se sentir meilleur.*

Comme le hasard n'a ni goût, ni dégoût, ni habitudes, il est impuissant à produire les séries faciles de l'anaphore et il s'étonne naïvement devant les vers qui les multiplient :

*Et leur cœur chante, et leur cœur pleure, et leur cœur bout...*

*Soit ton pouvoir, soit ton tourment, soit ton effroi...*

*J'aime mes yeux, mes mains, ma chair, mon torse...*

## 2.5 Motivation du signe

On a beaucoup glosé sur la motivation des signes et des sons, sur les *s* qui sifflent, les *F* et les *v* qui soufflent. Et Verhaeren exploite volontiers les opportunités du lexique français quand elles se présentent. Ainsi par chance le vent commence par un *v* qui semble souffler à son oreille et le poète en profite pour déchaîner les *v* en rafales dans le poème bien connu :

*Le vent, le vent sauvage de novembre*

En réalité la composante sémantique de l'allitération est essentielle : en d'autres occasions on peut multiplier les *v*, sans entendre aucun souffle. Le *v* peut alors s'attacher au mol abandon du vol, comme dans le vers :

*L'oisive rêverie au vol aventureux.*

Qu'on supprime le verbe *SIFFLER* dans le vers toujours cité de Racine et le sifflement cesse, même si les *s* abondent. Comme J.M. Viprey l'a montré, la simple substitution d'un autre animal au serpent détruit l'effet de sifflement, et plus encore si l'on garde tous les *s* en changeant le reste :

*Pour qui ce saucisson qui s'offre à ma narine ?*

On a beau tendre l'oreille : le saucisson est un serpent qui ne siffle pas.

Il semble donc que l'allitération soit un procédé plus complexe qu'il n'y paraît. Son effet n'est guère senti que lorsque le sens s'accorde avec le son. En dehors des jeux verbaux des oulipiens, il est plutôt rare en poésie. Si le hasard tenait la plume, il en fabriquerait davantage, mais de faible qualité et sans signification. C'est peut-être sa rareté qui fait son intérêt et son succès un peu surfait dans les commentaires.

Mais l'allitération ne saurait fonder à elle seule la musique du vers. La statistique des sons répétés ne permet guère que de cerner vaguement la tonalité, alors que c'est la modulation des sons qui est en cause. La poésie, qui se déroule dans le temps, comme la musique, a besoin d'une description dynamique, sensible aux multiples jeux où la dissonance danse avec l'accord, la transition et la syncope avec le tempo, la variation avec la répétition et la symétrie. Cette approche exige des algorithmes plus complexes et plus fins où l'oreille puisse expliquer ce qu'elle sent à la raison. Les compositeurs et les musicologues ont montré la voie, au moins pour les faits acoustiques. Les images et les sentiments, que la musique, sans le secours du langage, véhicule sourdement, sont plus lisibles et perceptibles en poésie. Mais les règles de l'harmonie, du rythme et de l'effcience y sont plus diffuses : en poésie, l'oreille a ses raisons que la raison ne connaît pas.

## **Références**

V. Beaudouin, *Mètre et rythmes du vers classique. Corneille et Racine*, Paris : Champion, 2002.

B. de Cornulier, *Théorie du vers. Rimbaud, Verlaine, Mallarmé*, Paris : Éditions du Seuil, 1982.

- B. de Cornulier, *Art poétique. Notions et problèmes de métrique*, Lyon : Presses universitaires de Lyon, 1995.
- É. Évrard, « Quelques variations qualitatives dans l'hexamètre latin », dans *Méthodes quantitatives et informatiques dans l'étude des textes*, Genève-Paris : Slatkine-Champion, 1986 (1986c).
- Ph. De Marëuil, « From dilatation to coarticulation. Is there vowel harmony in French ? », *Studies in the Linguistic Sciences*, vol. 46 n°2, 2002, p. 1-21.
- G. Purnelle, « La place de l'e post-tonique dans l'octosyllabe de quelques poètes », *Actes des 4<sup>e</sup> Journées internationales d'Analyse Statistique des Données Textuelles (JADT 1998)*, Nice : UNSA, 1998 (1998a), p. 529-544.
- J. Roubaud, *La vieillesse d'Alexandre*, Paris : Éditions Ramsay, 1978.
- J.M. Viprey, « Pour un traitement textuel de l'allitération », *Semen*, vol. 12, 2000.