

LES RYTHMES DANS LE SLAM

[Anne-Catherine Simon](#)

Éditions de la Maison des sciences de l'homme | « Langage et société »

2020/3 N° 171 | pages 139 à 169

ISSN 0181-4095

ISBN 9782735126644

Article disponible en ligne à l'adresse :

<https://www.cairn.info/revue-langage-et-societe-2020-3-page-139.htm>

Distribution électronique Cairn.info pour Éditions de la Maison des sciences de l'homme.

© Éditions de la Maison des sciences de l'homme. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Les rythmes dans le slam

Anne-Catherine Simon

Université catholique de Louvain, Institut Langage et
Communication, Centre VALIBEL/Discours et Variation
anne-catherine.simon@uclouvain.be

À la mémoire de Mathilde Delhal

1. Introduction¹

Cette contribution propose une analyse prosodique d'un corpus de slams pour y rendre compte de la création de styles vocaux particuliers, ou phonostyles, dans la performance orale. Notre approche doit beaucoup à la phonostylistique développée par Léon (2009) et, plus largement, aux linguistes qui ont donné une place première à la signification des faits phonétiques, comme Gumperz (1982), Fónagy (1983) ou Auchlin (2013). La phonostylistique de Léon est apparentée à la sociolinguistique ou à la sociophonétique lorsqu'elle vise à décrire « un répertoire des styles sonores, ou *phonostyles*, tels qu'ils sont perçus en tant que caractéristiques d'un individu (jeune, vieux, homme, femme), d'un groupe social (prolétaire, bourgeois), ou d'une circonstance particulière (discours politique, sermon), etc. » (Léon 1993 : 3). Décrire

1. Mes remerciements vont aux évaluateurs ou évaluatrices de la revue. Par leurs commentaires avisés et constructifs, ils ou elles m'ont permis d'améliorer sensiblement la première version de cet article. Je remercie également Josiane Boutet et Luca Greco pour leur suivi éditorial sans faille.

un phonostyle perçu comme représentatif implique de mesurer la perception de la parole (par des jugements ou des catégorisations) pour comprendre l'effet des variantes et leur signification sociale chez des récepteurs (Candea & Trimaille 2015 : 13). Notre étude ne vise pas à vérifier si le slam forme un phonostyle stable ou facilement identifiable, et distinct du rap par exemple. Le slam étant une forme de poésie orale extrêmement libre et variée, nous ne pensons pas qu'il existe un modèle ou un canon que des études perceptives pourraient mettre en évidence. Si l'on revient à la phonostylistique, on constate que Léon l'associe également à une poétique de l'oralité et cherche à observer « avant tout comment le travail sur la substance et la forme de *l'expression* manifeste des sens nouveaux par rapport au discours ordinaire, considéré seulement comme un instrument de communication » (Léon 1993 : 44). L'analyse phonostylistique vise alors les processus de déconstruction et de reconstruction du signe aux niveaux phonématique et prosodique. Le rôle de la substance phonétique est crucial pour comprendre comment une parole devient expressive, c'est-à-dire une parole centrée sur le locuteur qui lui permet d'exprimer des attitudes ou des émotions ou d'emphatiser certaines parties des énoncés. Mais la parole expressive n'est pas seulement centrée sur le locuteur : elle produit aussi des effets sur les auditeurs. Ces effets sont instanciés par la substance phonétique et font l'objet d'une interprétation expérientielle (Auchlin 2013 : 8) : c'est particulièrement sensible dans la dimension rythmique et temporelle de la parole, qui impose d'une manière directe une vitesse à l'expérience perceptive (on ne peut pas écouter plus vite que la parole n'est produite).

Étudier les phonostyles du slam permet de catégoriser les formes que peut prendre cet art de la parole libre (Vorger 2011 : 65) qui vise avant tout à partager des émotions (Raket & Mezmini 2017 : 6). La définition formelle du slam ne présuppose aucune forme d'écriture ni d'interprétation (Cabot 2017 : 3). L'ensemble des auteurs s'accordent cependant pour dire que le rythme tient une place prédominante dans le slam, rythme créé avec les mots, la voix et le souffle. Notre recherche est empirique et inductive. Nous cherchons d'abord à établir une classification des différents phonostyles, ou styles vocaux, dans un corpus de slam à partir de leurs caractéristiques prosodiques principales et de pallier ainsi l'absence d'analyses prosodiques instrumentées du slam. Cette classification des phonostyles du slam ne peut cependant se fonder uniquement sur des caractéristiques prosodiques, aussi détaillées soient-elles : elle doit tenir compte de leur interaction avec la dimension

verbale et textuelle². Par conséquent, nous investiguons en second lieu les techniques d'écriture et de mise en voix utilisées pour construire ces styles vocaux. Plus précisément, nous cherchons à établir l'interaction entre le matériel verbal (signifiant sonore, lexicale, syntaxe, métrique) et le matériel prosodique (débit, accentuation, scansion, silences) dans la construction des rythmes.

Enfin, pour établir la signification de ces rythmes et leur contribution à la parole expressive, nous proposons des analyses basées sur l'importance des faits phonétiques concrets (Boutet 2014), la prosodie des émotions (Léon 1993 ; Kraxenberger, Menninghaus, Roth & Scharinger 2018) et l'interprétation iconique de la prosodie (Fónagy 1999 ; Auchlin 2013 ; Prsir & Simon 2013). Par sa fonction iconique, ou imitative, la prosodie peut représenter la forme spatiale ou temporelle d'un signifié. Elle fait ainsi apparaître un signifié dans la parole d'une manière élémentaire, expérientielle, qui donne une impression ou un accès direct à une signification (comme une voix qui descend progressivement vers des tons de plus en plus graves peut donner l'impression d'une chute).

Dans la suite de cette contribution, nous donnons quelques éléments factuels sur le slam (2) avant de présenter la méthodologie d'analyse prosodique du corpus (3). Les quatre phonostyles identifiés sont décrits de manière détaillée dans la section (5) et on montre enfin comment chaque slameur les combine de manière originale (6).

2. Le slam, une pratique poétique libre

Le slam consiste à déclamer en public des textes en les scandant. On slame le plus souvent à l'occasion de scènes ouvertes organisées par des structures associatives ou des bars : chacun est invité à monter sur scène pour dire un texte dont il est l'auteur, sans accessoire ni musique, dans une limite temporelle fixée à trois minutes. Cette pratique remonte aux années 1980, quand Marc Kelly Smith décida d'organiser des tournois de poésie dans un club de Chicago : « Il s'agissait de faire de la poésie quelque chose de vivant, quelque chose de spectaculaire, et de briser la frontière classique qui sépare les artistes du public » (Raket & Mezmiri 2017 : 10). Ces soirées prirent le nom de *Uptown Poetry Slam* et le phénomène allait devenir mondial. En anglais, *to slam* signifie faire

2. Voir des études antérieures où nous avons insisté sur l'importance d'analyser la prosodie dans ses « mises en phase » et « déphasages » avec les autres niveaux de l'analyse linguistique, afin d'en comprendre l'apport (Simon & Auchlin 2001 ; Simon 2004 ; Degand & Simon 2009).

claquer une porte, lancer violemment ou critiquer³ : « Le slam, c'est comme un bruit de porte qui claque, une grosse droite ou un coup de latte qui vous retourne et vous éclate l'âme. » (Petit Poussin⁴)

Le slam permet une expression en prise directe avec la société, traitant de sujets politiques ou intimes, tournés de manière à leur donner une dimension universelle. Ce qui le constitue en genre, ce ne sont pas des contraintes formelles d'écriture ou de déclamation, ni des contraintes thématiques, mais un dispositif situationnel particulier instaurant une série de coupures avec la parole ordinaire ou conventionnelle (Cabot 2017) : le slameur est sur scène et sa voix est amplifiée vers le public par un micro ; il prend la parole pour un temps limité et la possibilité de lui répondre se fait éventuellement par un slam ultérieur ; il s'exprime souvent sous un pseudonyme. On comprend que le slam a « pour enjeu symbolique l'émergence d'une parole partagée, démocratique, démarquée du discours dominant, utilitaire, médiatique, politique, vénal, doxique » (Cabot 2017 : 4-5).

Le slam est un art accessible : la scène est ouverte à toutes et à tous. Le rôle du maître de cérémonie est d'inviter les spectateurs à s'inscrire pour monter sur scène au cours de la soirée. Il n'est pas nécessaire d'être passé par une formation particulière pour venir défendre son texte⁵ et l'accessibilité est à la fois financière (droit d'entrée libre ; souvent, un texte dit donne droit à une boisson), technique (pas de contrainte à respecter dans l'écriture ni dans la déclamation) et sociale (les lieux où s'organisent des scènes slam ne sont pas réservés à certaines catégories sociales ou d'âge). Il y a parfois des tournois visant à élire un gagnant ou une gagnante de la soirée à l'issue de plusieurs manches, mais dans un climat bienveillant.

Enfin, le slam est une pratique poétique partagée. La première chose qui est partagée, c'est le temps de parole. La durée des prestations est limitée à trois minutes pour ne pas lasser le public et laisser la possibilité à chacun de s'exprimer. Le strict respect de cette limite est très variable, comme varie la façon d'indiquer la fin du temps de

-
3. Reynolds et Vorger (2014) proposent une analyse détaillée de l'étymologie et des sens actuels du mot anglais et français.
 4. Chaque extrait est identifié par le nom de scène de son auteur. Première phrase d'un slam de Petit Poussin (Simon Raket). Voir en ligne : <<https://www.youtube.com/watch?v=eZKOOZjFkZg>>, page consultée le 27/11/2018.
 5. Cf. la notion d'art moyen (Pierre Bourdieu, *Un art moyen. Essai sur les usages sociaux de la photographie*, Éditions de Minuit, Paris, [1965] 2003, cité par Cabot 2017 : 4).

parole (retentissement d'un gong, lumière rouge, coupure du micro). On partage aussi une attitude de bienveillance envers le slameur : c'est une valeur souvent revendiquée comme constitutive de la scène ouverte. Les auditeurs peuvent manifester leur plaisir ou leur approbation par des claquements de doigts. Enfin, une dimension importante du partage est donnée par le rythme. En scandant son texte, le slameur produit une pulsation rythmique qui est entendue par l'auditoire. Cette pulsation induit une synchronisation sensorimotrice :

Le fait est banal. Qui ne s'est surpris à battre du pied en écoutant un morceau de musique au rythme très marqué ? Cette banalité nous empêche peut-être de nous étonner d'une conduite assez singulière. Non seulement celui qui « accompagne » un rythme exécute un mouvement dont la période coïncide avec celle des sons, mais en outre ses frappes coïncident dans le temps avec les stimulus marqués. En d'autres termes, le mouvement est doublement synchrone. (Fraise 1974 : 63)

La synchronisation motrice n'est pas une simple réponse déclenchée par un battement rythmique : pour se synchroniser rythmiquement, il faut anticiper le battement à venir. Cette synchronisation est donc le signe d'une attention pleine et active, et du partage d'un même mouvement rythmique.

3. Rythme verbal et rythme prosodique

La perception d'un rythme dans la parole implique le retour régulier d'éléments structurés (Auer, Couper-Kuhlen & Müller 1999 : 3). Cette perception peut être le produit du matériel verbal, par la répétition d'une suite de sons, d'un mot ou d'une structure syntaxique. Lorsqu'une suite de sons placés à la fin d'une unité linguistique est répétée (rime) et que l'unité linguistique compte un nombre régulier de syllabes (mètre), cela produit un texte versifié⁶. Dans l'exemple (1), une succession d'assonances viennent scander le déroulement de la phrase [skas, stɛz, stagn, gwas, staz]. Cette régularité sonore est appuyée par la répétition de la structure syntaxique 'se + verbe conjugué' (*se casse, se taisent, se tassent*).

6. Le slam est une pratique orale et il n'est généralement pas possible de se référer à une version écrite, stable, du texte. La régularité métrique observée est basée sur un décompte des syllabes prononcées, et non pas sur des règles de versification de la langue littéraire (Jenny 2003)

(1) [Skash]

Et quand l'espoir se casse	[skas]
mes gars se taisent	[stɛz]
mes gars stagnent	[stagn]
mes gars se tassent	[stas]
avec l'angoisse	[gwas]
en guise de métastases	[staz]

Dans l'exemple (2), la succession de rimes en [ɔl] produit un rythme régulier renforcé par le fait que les trois premiers vers ainsi délimités comptent chacun huit syllabes⁷.

(2) [Petit Poussin]

Ça s'apprend pas dans les écoles	[kɔl]
Sans naphtaline et sans formol	[mɔl]
C'est bien plus grisant que l'alcool	[kɔl]
Ça vole	[vɔl]
Le slam	

La perception d'un rythme régulier peut aussi être le produit de la performance vocale, quand on répète à intervalles réguliers un accent, un contour intonatif ou une pause silencieuse. Dans l'exemple (3), les syllabes accentuées sont régulièrement espacées de manière à produire une scansion dite isochronique. L'intervalle temporel moyen d'une syllabe accentuée à la suivante s'élève à 0,65 seconde, ce qui génère un rythme de 95 battements par minute (bpm)⁸. Chaque syllabe accentuée (forte) est précédée de deux syllabes non accentuées (faibles), ce qui renforce la régularité temporelle par une isométrie, produisant un regroupement ternaire. Ce rythme temporel est soutenu au plan verbal par des assonances en [o] et [ɛm] et par une répétition syntaxique (*c'est... qui nous...*).

7. Le quatrième vers délimité par la rime ne compte que deux syllabes, ce qui produit un effet de syncope par rapport aux octosyllabes qui précèdent.

8. Conventions de notation : les scansions rythmiques sont transcrites avec un retour à la ligne délimitant chaque groupe de syllabes scandé par un accent (symbole '). Les barres obliques représentent les battements. On indique parfois le nombre de syllabes par groupe accentué (isométrie). La durée moyenne des intervalles temporels entre deux syllabes accentuées successives (ou battements) est indiquée en secondes (isochronie). Cette durée permet de calculer le tempo (nombre de battements par minute). Le cas échéant, la durée des pauses silencieuses est indiquée entre parenthèses en secondes. Ce système de notation est adapté de Auer, Couper-Kulhen et Müller (1999).

(3) [Gaëtan Sortet]

	Nb de syl.	Intervalle moy.	Tempo
C'est le 'beau /	3		
qui nous 'sauve /	3		
de nous- 'mêmes. /	3		
(0,82 sec)			
C'est la 'peau /	3		
qui nous 'chauffe /	3		
et nous 'aime. /	3	0,65 s	92 bpm

Ces exemples illustrent les deux niveaux constitutifs du rythme : le niveau verbal (préparé par l'écriture) et le niveau paraverbal (réalisé dans la performance). Le niveau verbal, ou textuel, a été amplement étudié par Vorger (2011, 2016) qui a relevé dans des slams les jeux de rimes, d'assonances, d'allitérations ou de paronomases qui actualisent la fonction poétique du langage :

Dans le slam, les mots sont travaillés dans leur matérialité et leur musicalité [...]. Ce flot de mots n'est endigué par aucun rythme préétabli, à la différence du rap [...]. La musique tend à sourdre des mots, de leurs sonorités et de leur succession. (Vorger & Abry 2009 : 64)

Quant au rythme prosodique réalisé au cours de la performance orale, il repose sur une analyse auditive ou instrumentale d'enregistrements sonores. En effet, le slameur est libre d'organiser temporellement sa parole pour lui imprimer un rythme particulier, par son débit d'articulation, l'insertion de pauses et l'accentuation de certaines syllabes. Dans ce sens, le rythme est « l'organisation de la parole par un sujet » (Dessons & Meschonnic 1998 : 28). Le texte fournit des potentialités rythmiques qui seront actualisées par le slameur au moyen de choix concernant le débit d'articulation, les syllabes accentuées et l'isochronie ou l'isométrie de ces scansionnements accentués.

La méthode d'analyse que nous proposons vise à rendre compte de la construction de rythmes dans la performance par la prosodie et le matériel de la structure linguistique signifiante.

4. Méthode

Nous présentons le corpus de 12 slams à la base de cette étude, ainsi que les méthodes d'annotation et les outils d'extraction de mesures prosodiques utilisés. Ces outils sont distribués librement et utilisables sur n'importe quel corpus d'enregistrements.

4. 1. Sélection des enregistrements

Le corpus comprend 12 slams performés par des artistes (11 hommes et 1 femme) qui se produisent régulièrement à la scène ouverte de La Zone à Liège en Belgique. Il s'agit d'enregistrements réalisés en studio par un ingénieur du son dans le cadre d'un projet de diffusion et pour lesquels nous avons reçu un accord d'analyse à des fins scientifiques⁹. Habituellement, la durée d'un slam est limitée à trois minutes, mais on voit que cette règle n'a pas été respectée strictement pour ces enregistrements.

Tableau 1 : Présentation du corpus : titre des morceaux, auteurs (nom de scène), durée, proportion de pauses silencieuses par rapport au temps total de parole, débit moyen d'articulation et proportion de syllabes accentuées.

Titre	Auteur-tinterprète	Durée	Taux de pause (%)	Débit (syll/s)	Syllabes accentuées (%)
Slam	Petit Poussin	2 m 35 s	31 %	4,0	19 %
Je ne suis pas parano	Philippon	2 m 20 s	13 %	4,0	30 %
Métastases	Skash	3 m 08 s	35 %	4,4	22 %
Une rencontre	NK	2 m 35 s	16 %	4,1	31 %
Dandy	L'Ami Terrien	1 m 47 s	7 %	4,4	10 %
Wallifornie	Renard de Feu	3 m 19 s	28 %	4,0	19 %

9. Ces enregistrements ont été réalisés par Ludo et rassemblés sous le titre : *23 textes, 23 voix, 23 poètes actifs à La Zone*. Ils sont disponibles sur le site web de La Zone (<<https://www.lazone.be/slam/>>, page consultée le 14/04/2020). Nous remercions Maxime Deflandre, responsable de cette scène, qui a sollicité l'accord des auteurs pour l'exploitation scientifique de ces enregistrements. La Zone accueille des projets culturels alternatifs axés sur « la résistance créative ouverte à toutes les formes d'expression et toujours réceptive aux cultures émergentes dans des rapports de tolérance, de reconnaissance des désirs affinitaires et de garantie d'accès pour chacun ».

L'être et le non-être	7 ^e Œil	2 m 37 s	31 %	3,9	18 %
La vache	Gaëtan Sortet	1 m 54 s	54 %	2,1	22 %
Sur le fleuve Ramour	Mineige (femme)	3 m 53 s	25 %	3,4	24 %
L'amour n'est jamais propre	Volauvent	3 m 34 s	49 %	2,8	22 %
Inspiration nocturne	Josh	3 m 47 s	36 %	3,7	17 %
Cordiales vitupérations	PiY	5 m 15 s	30 %	4,0	17 %
Durée totale du corpus		36 m 42 s			

4. 2. Extraction des mesures prosodiques

Les enregistrements sonores ont été transcrits orthographiquement¹⁰ en utilisant le logiciel Praat pour aligner ces transcriptions sur les sons (Boersma & Weenink 2018). Les transcriptions orthographiques ont été converties en transcriptions phonétiques, vérifiées manuellement (pour les élisions les liaisons et autres modifications par rapport à la prononciation standard), puis alignées au niveau du mot, de la syllabe et du phone à l'aide du plug-in EasyAlign (Goldman 2011). Cet alignement a été vérifié manuellement.

La transcription phonétique alignée sur le son est à la base des analyses prosodiques. Le script Prosogram (Mertens 2004) a été utilisé pour styliser la courbe de fréquence fondamentale et représenter

10. Pour 4 textes sur les 12, nous avons eu accès à une version écrite produite par les auteurs, ce qui a permis de trancher entre des transcriptions multiples possibles : les textes « Dandy » et « La vache » sont publiés dans l'ouvrage *Slam. Poésies et voix de Liège* (Raker & Mezmizi 2017). Les textes « Métastases » et « Une rencontre » nous ont été communiqués par leurs auteurs à l'occasion d'une rencontre avec les étudiantes et étudiants en lettres à l'UCLouvain en mars 2019.

visuellement la voix parlée comme sur une partition musicale. Les trois couches d'annotation dans la partie inférieure du schéma (Figure 1) correspondent aux syllabes, aux mots et aux énoncés. Pour chaque syllabe articulée, un trait noir représente le noyau de la syllabe dont la hauteur mélodique (ou fréquence fondamentale) a été stylisée (Mertens 2004).

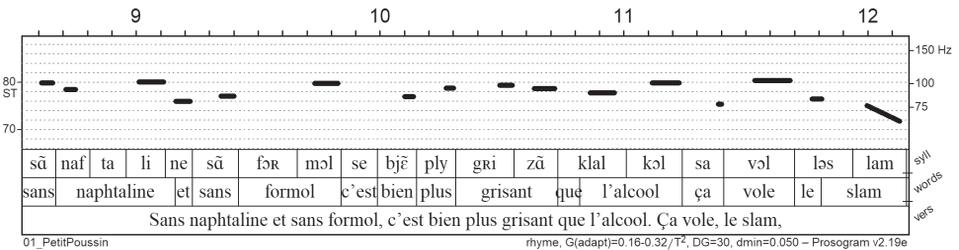


Figure 1 : Prosogramme (script de Mertens 2004) d'un extrait du slam de Petit Poussin

L'analyse s'est concentrée sur les variables temporelles. À l'aide du plug-in ProsoBox distribué librement (Goldman & Simon 2020), nous avons extrait automatiquement les mesures suivantes :

- détection des syllabes proéminentes selon la méthode présentée dans Simon, Avanzi et Goldman (2008). Cette détection automatique, basée sur des critères acoustiques de durée et de hauteur, a été vérifiée auditivement en tenant compte d'autres facteurs pouvant induire une perception proéminence, comme la catégorie grammaticale (Lacheret, Simon, Goldman & Avanzi 2013) ;

- extraction de mesures prosodiques temporelles (décrites dans Grosjean & Deschamps 1972) : débit d'articulation (nombre de syllabes articulées par seconde) ; durée des pauses silencieuses ; temps total des pauses ; proportion du temps de pause par rapport au temps d'articulation ;

- extraction de mesures accentuelles : densité accentuelle (proportion de syllabes accentuées par rapport au nombre total de syllabes) ; proportion des syllabes accentuées en position initiale de mot par rapport aux syllabes accentuées en position finale de mot.

L'exemple (4) et la figure 2 illustrent les mesures temporelles exploitées dans notre étude.

(4) [Skash ; chaque syllabe accentuée est précédée du symbole ']

c'est peut-être le 'fait qu'on 'ti'tube qui nous 'consti'tue

En haut à droite du graphique sont affichées des mesures prosodiques sur l'énoncé : la durée (2,9 s) est calculée à l'aide de la graduation sur la partie supérieure du schéma ; $n=13$ correspond au nombre de syllabes articulées ; $f_0 = 141$ correspond à la hauteur mélodique moyenne exprimée en Hertz (141 Hz) et $\Delta 71$ Hz mesure la différence entre la hauteur minimale et maximale, correspondant au registre local¹¹ ; le débit d'articulation est calculé en syllabes par seconde (4,5 syll./s) ; enfin, la proportion 6/13 indique le nombre de syllabes détectées comme proéminentes sur le nombre total de syllabes, donnant une indication de la densité accentuelle.

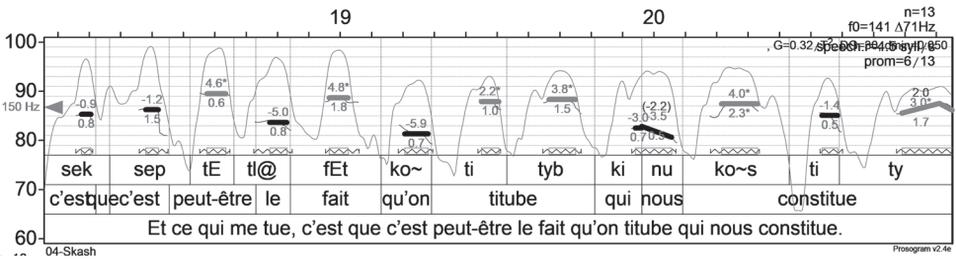


Figure 2 : Représentation visuelle de l'exemple (4) sous la forme d'un prosogramme (Mertens 2004) enrichi des mesures de ProsoBox (hauteur relative, durée relative, mouvement mélodique). La transcription phonétique est présentée en alphabet SAMPA

En outre, chaque syllabe est accompagnée de mesures locales¹² : si l'on considère la dernière syllabe [ty] du mot *constitue*, sa durée relative est de 1,7, ce qui signifie que cette syllabe est allongée ; la hauteur relative est de 3 demi-tons, ce qui signifie que cette syllabe est perçue comme haute ; la variation mélodique interne est de 2 demi-tons, ce qui signifie que cette syllabe est perçue comme montante. Les syllabes proéminentes successives peuvent donner lieu à des scansions. Ces scansions rythmiques ont été annotées manuellement.

11. Les mesures mélodiques ont été extraites mais n'ont pas fait l'objet d'une analyse systématique.
12. Ces mesures indiquent la saillance de la syllabe dans son contexte local. Elles sont calculées à l'aide de l'outil de détection automatique des syllabes proéminentes ProsoProm (Simon *et al.* 2008) intégré dans la ProsoBox. L'empan local utilisé pour mesurer la hauteur et la durée relatives est de 4 syllabes : la syllabe cible, les deux qui précèdent et celle qui suit. Cet empan est réduit à trois syllabes si une pause silencieuse supérieure à 0,25 s précède ou suit la syllabe cible. Une syllabe dont la durée relative = 2 est donc deux fois plus longue que la durée moyenne des 3 syllabes environnantes.

4. 3. Annotation des scansions rythmiques

Des scansions sont perçues si l'organisation des syllabes accentuées est régulière. L'isochronie, ou régularité temporelle, provient de ce qu'une syllabe accentuée revient à intervalles temporels perçus comme identiques. L'isométrie, ou régularité métrique, provient de ce que les syllabes accentuées sont séparées par un nombre régulier de syllabes inaccentuées. On identifie un passage isochronique dès qu'on peut accompagner la scansion en tapant en rythme dans les mains. L'isométrie requiert un décompte des syllabes prononcées et ne correspond pas nécessairement aux règles classiques de la poésie métrée.

Les passages scandés sont identifiés auditivement et annotés selon la méthode présentée dans Auer, Couper-Kuhlen et Müller (1999) et adaptée au français par Simon et Grobet (2005). L'annotation vise à objectiver et à valider la perception à l'aide de mesures. Elle consiste à identifier les syllabes accentuées qui constituent les battements d'une scansion, à mesurer les intervalles temporels entre ces battements et à compter le nombre de syllabes non accentuées (faibles) entre deux syllabes accentuées (fortes) afin de vérifier si la scansion possède également une régularité métrique. L'annotation se fait dans un fichier Textgrid du logiciel Praat, comme illustré par la figure 3.

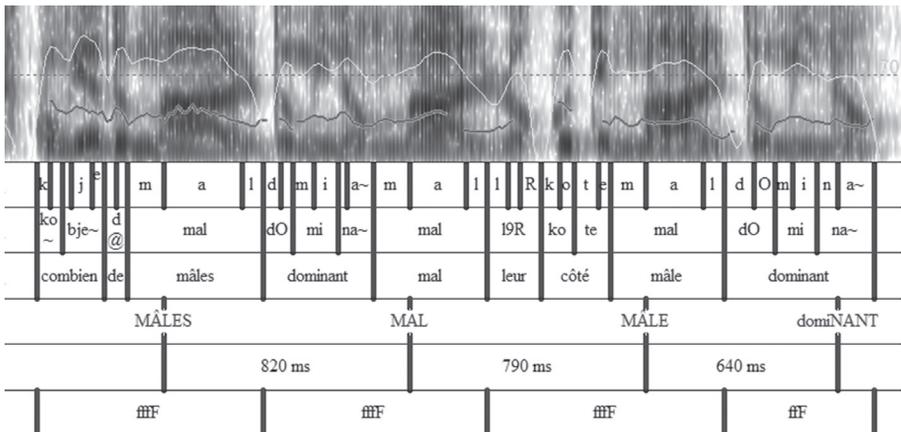


Figure 3 : Annotation d'une scansion rythmique sous Praat avec les couches d'annotations suivantes : phones (en alphabet phonétique SAMPA), syllabes (en SAMPA), mots ; attaque de la voyelle des syllabes accentuées (battements) ; intervalles temporels entre deux syllabes accentuées mesurés en millisecondes (pour vérifier l'isochronie) ; décompte des syllabes faibles (f) et fortes (F) dans chaque groupe accentué (pour vérifier l'isométrie)

Dans l'exemple *combien de mâles / dominant mal / leur côté mâle / dominant*, nous percevons quatre battements régulièrement espacés qui créent trois intervalles temporels mesurables (820 ms, 790 ms et 640 ms). La durée moyenne de ces intervalles s'élevant à 0,75 sec, le tempo de cette scansion est de 80 battements par minute (bpm). Pour représenter une scansion à l'écrit (exemple 5), on utilise les conventions de notation présentées ci-dessus (note 8). La syllabe accentuée est finale de groupe (quand elle est délimitée par un accent final démarcatif primaire en français) ou initiale de groupe (ce qui produit un effet de regroupement différent). Les travaux des psycholinguistes sur la synchronisation motrice avec un rythme perçu auditivement nous informent de plusieurs éléments importants (Fraisse 1974 ; Couper-Kuhlen 1993). Un intervalle temporel est perçu comme identique au précédent malgré des différences de durée qui peuvent aller jusqu'à 20 % (voire 30 % dans certains cas). Ainsi, le raccourcissement de 4 % ou 8 % des intervalles successifs dans l'exemple (5) ne dégrade la perception de la régularité temporelle de la scansion.

(5) [Petit Poussin]

	Nb. syl.	Intervalle moy.	Diff. de durée relative
combien de 'mâles /	4		
dominant 'mal /	4	0,82 s	
leur côté 'mâle /	4	0,79 s	-4%
domi'nant /	3	0,64 s	-8%

D'autre part, lorsqu'on se synchronise avec une pulsation rythmique, il a été montré que le geste (par exemple le claquement de doigts) est synchronisé avec l'attaque de la voyelle de la syllabe accentuée plutôt qu'avec l'attaque de la syllabe : c'est le *P-centre* (*perceptual center*, voir Auer, Couper-Kuhlen & Müller 1999 : 53). C'est la raison pour laquelle nous mesurons les intervalles de durée entre deux attaques de voyelles. Dans cette scansion formée de quatre battements, les trois premiers groupes accentués comptent chacun 4 syllabes (*com-bien de mâles / do-mi-nant mal / leur côté mâle*) tandis que le dernier en compte 3 (*do-mi-nant*). L'isométrie est quasiment atteinte (4-4-4-3) et renforce le caractère régulier de la scansion. Cette régularité métrique est aussi appuyée par la répétition phonique et lexicale (*mâle/mal/mâle, dominant/dominant*). L'ensemble de ces éléments a été annoté pour chaque scansion : isochronie, isométrie, nombre de battements, tempo, présence ou non d'éléments verbaux (rimes, assonances, allitération, répétitions lexicales ou syntaxiques).

4. 4. Paramètres primaires et secondaires et typologie des styles vocaux

Pour conclure cette partie méthodologique, nous présentons un tableau des paramètres primaires et secondaires retenus pour l'analyse des styles vocaux dans le slam. Les paramètres primaires sont ceux qui contribuent prioritairement à créer des rythmes car ils sont responsables de la segmentation et de la scansion du texte. Les paramètres secondaires viennent renforcer ces rythmes ou les colorer.

Tableau 2 : Paramètres verbaux et prosodiques impliqués dans les rythmes du slam¹³

Paramètres verbaux	Paramètres prosodiques
Paramètres primaires	
rimes ou assonances (répétition de syllabes, de voyelles ou de consonnes en fin de vers)	scansions isochroniques
régularité métrique (nombre de syllabes)	scansions isométriques
Paramètres secondaires	
allitérations (répétition de consonnes à l'initiale ou dans le mot)	type d'accentuation (finale ou initiale)
homophonie ou paronomase (répétition de mots phoniquement identiques ou proches ¹⁴)	densité accentuelle (proportion de syllabes accentuées)
répétitions lexicales ou syntaxiques	débit d'articulation (en syllabes par seconde)
etc.	pauses silencieuses

Les rimes permettent d'identifier une segmentation du texte en vers¹⁵ tandis que les scansions opèrent une segmentation du texte par les battements rythmiques réguliers. À partir de là, nous définissons

13. Dans le cadre de cet article, nous nous limitons aux variables prosodiques temporelles. Les variables mélodiques sont également très importantes : contours intonatifs, registre mélodique restreint ou élargi, etc.

14. Ce paramètre est en interaction avec les rimes.

15. Par défaut le slam n'existe pas par écrit. La segmentation en vers se base sur les rimes et ne correspond pas nécessairement à la manière dont les auteurs auraient opéré le découpage graphique de leur texte (par retour à la ligne après chaque vers).

quatre phonostyles en fonction de la combinaison des paramètres verbaux et prosodiques : style parlé (non rimé, non scandé), style accentué (non rimé, scandé), style poétique (rimé, non scandé) et style élaboré (rimé, scandé).

Tableau 3 : Définition des quatre phonostyles de slam identifiés selon les critères verbaux (+/- rimé) et prosodiques (+/- scandé)

	- scandé	+ scandé
- rimé	style parlé	style accentué
+ rimé	style poétique	style élaboré

Chaque style s'actualise de manière spécifique en fonction des paramètres secondaires, donnant lieu à des variantes et à des modulations analysées dans la section suivante.

5. Analyse des phonostyles

On peut comparer les styles vocaux dans le slam au flow dans le rap. Le flow décrit la technique utilisée pour produire le flux de parole en le scandant sur une musique donnée. Un flow est linéaire lorsqu'on établit une correspondance entre les accents de la voix et les accents musicaux et complexe lorsque s'établissent des déphasages entre les accents vocaux et musicaux (Mahiou 2015 : 33 ; Migliore & Obin 2018). Dans le slam, la voix est nue et le flow se construit à partir de la combinaison de paramètres verbaux (les rimes dans le texte) et de paramètres prosodiques (les scansion dans la performance orale), donnant lieu à 4 phonostyles (figure 4).

Le style parlé ne fait usage ni de rimes ni de scansion prosodiques : il correspond à une parole modérément accentuée, sans régularités rythmiques ou sonores particulières, proche de ce qu'on peut observer dans la parole quotidienne. Ce style parlé représente 21 % du temps de parole dans le corpus (section 5. 1.). Le style poétique contient des rimes et le texte présente une structure métrique, mais la performance n'actualise pas de scansion rythmiques régulières. Ce style représente 16 % du temps de parole dans le corpus (section 5. 2.). Le style accentué caractérise des passages à forte densité accentuelle, avec une présence accrue d'accents en position initiale de mot. Ces accents ne sont pas produits de manière régulière et le texte, bien que non rimé, peut

contenir des allitérations. Ce style accentué représente 19 % du temps de parole dans le corpus (section 5. 3.). Enfin, le style élaboré résulte d'une combinaison de texte rimé ou assonancé performé avec des scansions prosodiques régulières. Il représente 44 % du temps de parole dans le corpus (section 5. 4.).

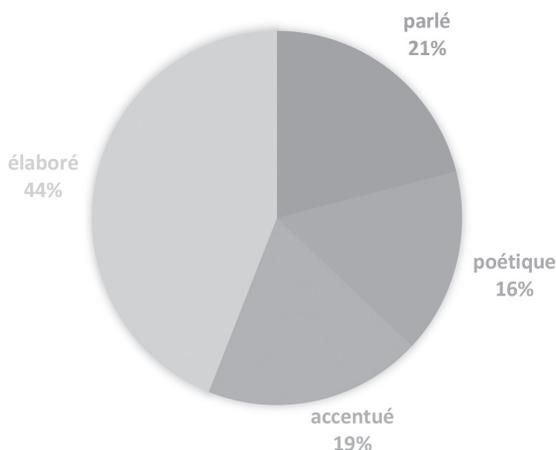


Figure 4 : Distribution (en temps de parole) des 4 phonostyles dans l'ensemble du corpus

Nous décrivons chaque style et ses variantes avant d'analyser comment ces styles se combinent chez les différents slameurs (section 6).

5. 1. Le phonostyle parlé

Le style parlé représente 21 % du temps de parole dans le corpus. Par ses paramètres prosodiques, le style parlé ressemble à une parole publique prononcée à un débit normal (environ 4 syllabes par seconde) et avec une densité accentuelle assez faible (moins de 20 % des syllabes sont accentuées). Deux slams représentatifs de ce style démarrent l'un comme une prise de parole politique (exemple 6) et l'autre comme un cours de philosophie (exemple 7) :

(6) [Renard de Feu]

J'avais quatorze ou quinze ans début des années quatre-vingt quand j'ai reçu un autocollant « Brabant wallon un indivisible ». Comme c'est risible.

(7) [7^e Œil]

Lorsque je dis « je pense donc je suis », cela revient à dire « je ne pense pas donc je ne suis pas » dans le sens où émettre la négation d'une négation revient à émettre une affirmation.

Cependant, le style parlé ne manque pas d'expressivité. L'expressivité peut venir de l'usage d'accents emphatiques, comme illustré par l'accent de focalisation sur la syllabe initiale de *incompétence* à la figure 5. Cet accent initial atypique produit l'allongement de la syllabe [ɛ̃] et est marqué par un contour intonatif de focalisation (mouvement dynamique haut-bas). Les trois dernières syllabes du mot *incompétence* sont désaccentuées et réalisées sur un ton bas :

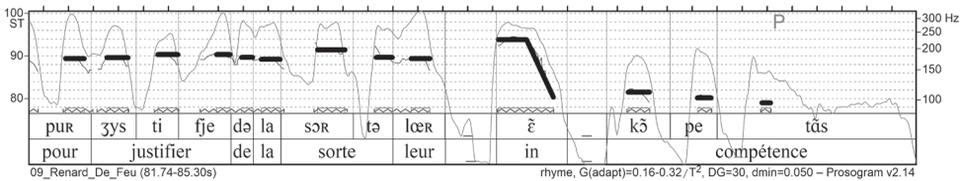
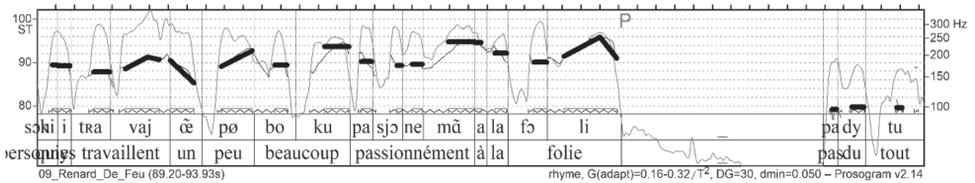


Figure 5 : Accent initial de focalisation sur le mot *incompétence*
[représentation visuelle par un Prosogramme, Mertens 2004]
[Renard de Feu]

L'expressivité du style parlé provient aussi d'une utilisation iconique de la prosodie, c'est-à-dire d'une imitation par la forme sonore (le signifiant) d'un aspect du contenu (le signifié). Dans la figure 6, on voit que le locuteur prononce l'énoncé *autant de personnes qui y travaillent un peu, beaucoup, passionnément, à la folie, pas du tout* avec un ample mouvement de crescendo jusqu'à à la folie (la cible mélodique de chaque syllabe finale est un peu plus haute que la précédente) suivi d'un rapide et brutal décroscendo sur *pas du tout* (chute de la mélodie et de l'intensité). La chute prosodique imite ainsi la chute de la productivité au travail.

Figure 6 : Mouvement mélodique de crescendo suivi d'un décroscendo
 (représentation visuelle par un Prosogramme, Mertens 2004)
 [Renard de Feu]



Il existe une variante de ce style parlé qui présente des caractéristiques prosodiques typiques d'une voix plus intime, proche de la confidence. Deux slams sont représentatifs de cette tendance, ceux de Josh et de Volauvent, où le registre mélodique est compressé, les accents initiaux peu nombreux et le débit d'articulation lent (2,8 à 3,7 syllabes par seconde). Dans ce style parlé intime, l'expressivité est marquée principalement par des pauses silencieuses emphatiques placées à des endroits non prévus par la syntaxe¹⁶ et donnant un relief particulier aux mots qui suivent. De cette manière, le silence marque « des pauses, des blancs, des syncopes, obligeant ainsi à des lignes de brisure entre les mots qui, en les soudant différemment et de façon inusitée, rompent l'unité sémantique » (Tyszler 2009 : 17).

(8) [Volauvent ; durée des pauses en secondes ; débit 2,1 syll/s]

L'amour (1,4) n'est jamais propre.

(9) [Volauvent ; durée des pauses en secondes ; débit 2,6 syll/s]

À bien y regarder (0.6) la vie (0.9) c'est exténuant (1.2), et une âme (0.5) c'est ténu (0.5), et un homme (0.2) ça doit tenir debout (1.1). Mais le plus souvent (0.3), ça boîte (0.8) et ça titube (1.2). C'est vertigineux (1.4).

(10) [Josh ; durée des pauses en secondes ; débit 3,8 syll/s]

Sur moi j'ai jamais parié (0.7). Je dirais jamais que personne n'arrive à ma cheville car (0.3) pour lutter je me suis jamais foulé.

16. Habituellement, la durée d'une pause silencieuse est proportionnelle au degré de frontière syntaxique : les pauses les plus longues sont situées en fin de phrase et renforcent le découpage syntaxique (Candea 2000 ; Grosman, Simon & Degand 2018).

Ce style parlé intime partage les caractéristiques de la « voix de charme » décrite par Léon (1993 : 77) : atténuation de l'intensité, vocalisation par adoucissement des consonnes, souffle, ralentissement du tempo.

5. 2. Le phonostyle poétique

Le style poétique représente 16 % du temps de parole dans le corpus. La dimension poétique de ce style est produite par un système de rimes ou de vers numériquement réguliers dans le texte. Les scansions prosodiques sont absentes et le rythme est produit par les régularités du signifiant phonique. Nous avons considéré comme rimée toute suite d'au moins deux séquences (syntagmes ou phrases) terminées des phonèmes identiques, comme les rimes en [ʃɛk] ou [sɛk] dans (11) ou les différentes assonances dans (12).

(11) [Josh]

On se demande pourquoi on court à l'**échec**.
En deux-mille-quatorze, t'es personne si t'as pas de **chèque**.
Pour oublier, il y en a qui boivent cul **sec**.

(12) [II-Y]

Transgresser les **normes**, les on-**dit**.
Oublier sa **donne**, son nom **maudit**.
Effacer des **tablettes**, les **règles** et ses **lingettes**.
Absoudre la **propreté** dans un **nitrate** de **soude** **aseptisé**.

Dans certains textes, l'élaboration verbale inclut l'utilisation de paronymes ou d'homophones, comme dans (13). Dans cet exemple, l'iconicité prosodique intervient avec une accélération du débit d'articulation sur la fin de la séquence mimant le contenu verbal ça va si vite.

(13) [Volauvent, homophonie /sa.va.si.vit/]

C'est que c'est vaste aussi, et que ça vacille vite, ça va si vite.

Dans le style poétique tel qu'il est défini ici, le signifiant verbal prime sur le paraverbal dans la construction du rythme. Cette dimension sonore du slam a été analysée par Vorger qui précise : « cette musique des lettres est souvent première dans le slam, l'enjeu étant de faire sourdre une mélodie à l'intérieur d'un texte » (2011 : 161).

Indépendamment de la présence de rimes, un texte peut être perçu comme métré quand les énoncés successifs comptent un nombre régulier

de syllabes prononcées. L'exemple (14) s'entend comme une suite de quatre octosyllabes car chaque énoncé compte huit syllabes (excepté le mot *enfin* qui s'intercale avant le dernier vers). C'est aussi un rythme octosyllabique qui est créé dans (15) et qui est renforcé par la présence de rimes internes et en fin de vers.

(14) [NK]

Et j'me souviens d'ses premiers mots...
Elle m'a choisi, c'est stupéfiant
et moi j'la dévorais des yeux...
Enfin,
d'un œil rougi, sous stupéfiants.

(15) [Mineige]

Tu cherch(e)s son nom dans le vin sombr(e).
Tu compt(e)s le nombre de tes ombr(e)s.
Mon pauvre loup, tu es relou.

La diction de ces vers n'est pas régulière au regard des règles de la métrique classique (Jenny 2003) et les syllabes finales de vers ne sont pas isochroniques. Cependant, une régularité du texte est perçue par la répétition métrique.

5. 3. Le phonostyle accentué

Le style accentué représente 19 % du temps de parole dans le corpus : ce style caractérise des passages très accentués (densité accentuelle élevée) mais dépourvus d'isochronie ou d'isométrie. Le plus souvent, on y observe une proportion importante d'accents initiaux. Les accents initiaux ne font pas partie de la base du système accentuel du français et sont moins fréquents que les accents finaux de groupes (Di Cristo 1999). On les a dénommés accents d'intensification, accents expressifs, accents d'insistance, ce qui met en avant leur caractère saillant. Ces accents sont rendus plus frappants et leur saillance est augmentée par la répétition de consonnes initiales de mots (allitérations). C'est le cas dans (16) : les allitérations des consonnes [t] et [s], ainsi que la répétition des voyelles [i] et [y], renforcent les syllabes accentuées et rendent ce passage plus percutant. Dans le symbolisme sonore, qui associe certains timbres vocaliques à des impressions auditives, les sons [i] et [t] sont considérés comme durs ou anguleux (Léon 1993 : 50-52), ce qui renforce l'aspect incisif de ce passage. En outre, ce slam se caractérise par un débit plutôt rapide (4,1 syllabes par seconde), un taux de pause bas (13 %) et une densité accentuelle élevée (31 %) produisant un rythme haché. Tous ces

paramètres prosodiques sont aussi ceux qu'on associe habituellement à la voix de colère (Léon 1993 : 117).

(16) [Skash ; les syllabes accentuées sont précédées du symbole ']

C'est pour tous mes 'têtus 'taciturnes qui diluent leur 'lassitude au fond d'un verre de 'cigüe, jusqu'à ne plus savoir où ils se 'situent. Et ce qui me 'tue c'est que c'est peut-être le 'fait qu'on 'titube qui nous 'consti'tue, je sais plus j'ai qu'une 'seule 'certi'tude, une 'seule chose 'dont je suis 'sûr, c'est que c'est 'pas parce qu'on se dit 'tu qu'on se dit 'tout.

Le style accentué peut aussi produire une impression plus sautillante ou légère. C'est le cas dans l'exemple (17), qui est caractérisé par un débit est plus lent (3,4 syllabes par seconde), une densité accentuelle moindre (24%), un taux de pause plus élevé (25%) et une voix mélodieuse (contours intonatifs amples et variés).

(17) [Mineige]

Tu prends 'goût au 'lait 'des u'sines, au 'miel de Mélu'sine, au 'ciel de ton 'lit, aux 'cachets d'aspi'rine.

Pour conclure, le caractère très rythmé et irrégulier du phonostyle accentué peut produire des effets interprétatifs différents en fonction des autres paramètres prosodiques, depuis une allure énergique ou révoltée jusqu'à une allure plus dansante et plus douce, avec un côté insistant des accents initiaux.

5. 4. Le phonostyle élaboré

Le style élaboré représente 44 % du temps de parole dans le corpus. Quantitativement, il est le style le plus fréquent dans ce corpus de slam. Ce style est façonné par la combinaisons d'une dimension rimée ou versifiée dans le texte avec des scansions rythmiques régulières dans la performance orale. Ce style permet de nombreuses variantes selon le débit d'articulation, le tempo des scansions et leur densité accentuelle. Le tempo d'une scansion est fonction de la durée entre les accents successifs et s'exprime en nombre de battements par minute (bpm). Les tempos observés dans notre corpus varient entre 20 bpm et 360 bpm.

Un tempo est perçu comme lent lorsque les syllabes accentuées successives sont espacées de deux secondes ou plus. Ce tempo sera perçu différemment selon le nombre de syllabes faibles intercalées entre deux syllabes accentuées. Dans (18), chaque mot monosyllabique est accentué et séparé du suivant par une longue pause, de manière à créer un tempo parfaitement

régulier de 30 bpm. L'effet interprétatif est celui d'une décomposition du langage, comme formulé dans la phrase qui suit cette scansion : *Tu peux décomposer chaque mot mais tu ne décomposeras jamais notre vie.*

(18) [Gaëtan Sortet]	Intervalle moy.	Tempo
Je nous veux heureux et saints et saufs (1,56)		
Je (1,33) /		
nous (1,40) /		
vie (1,57) /		
eux (1,58) /		
heure (1,59) /		
eux (1,56) /		
sains (1,27) /		
saufs. /	1,9 s	30 bpm

Dans l'exemple (19), le tempo est aussi lent (25 bpm), mais l'effet est très différent. D'une part, ce tempo lent entre en contraste avec le passage qui précède, réalisé sur un tempo près de quatre fois plus rapide (92 bpm). Ce brusque ralentissement incarne une dimension iconique, puisque qu'on passe d'une course effrénée (tempo rapide) à la résolution du morceau (tempo lent) qui donne la réponse à cette course : *ne pas mourir*. Les syllabes accentuées de cette scansion lente sont séparées par un grand nombre de syllabes non accentuées, ce qui produit un effet apaisé, très différent des monosyllabes dans l'exemple précédent.

(19) [Petit Poussin]	Intervalle moy.	Tempo
On 'court /		
encore, on 'court /		
sans 'savoir /		
vraiment 'pour- /		
'quoi /		
On se 'trimbale /		
presque à 'poil /		
mais notre 'cœur /		
porte la 'bur- /		
'qa /	0,65 s	92 bpm
Et même perdus dans le dé'dale /		
on s'arrêtera pas de cou'rir /		
En fait le slam ça sert à que 'dalle /		
à part à ne pas mou'rir /	2,5 s	25 bpm

Le ralentissement du tempo est une technique fréquemment utilisée pour annoncer la fin d'un slam : on la retrouve dans l'exemple (20), extrait d'un slam retraçant une rencontre amoureuse d'une nuit, et où le tempo lent final (20 bpm) vient signaler en même temps la conclusion

de la nuit et du morceau. Ce ralentissement fonctionne comme un indice de contextualisation au sens de Gumperz (1982).

(20) [NK]	Intervalle moy.	Tempo
C'était juste un autre 'soir / et deux âmes perdues au fond d'un 'bar / Juste une autre nuit qui dé'cline / les idées troublées par le 'gin /	2,8 s	20 bpm

Un tempo très rapide, où les syllabes accentuées se succèdent à intervalles très brefs (à partir d'un quart de seconde), peut produire un rythme staccato (détaché, haché) quand le slameur accentue chaque syllabe et parle d'un ton plat, peu modulé. Dans l'exemple (21) représenté à la figure 7, les syllabes accentuées se succèdent à intervalles de 0,25 s, ce qui produit un tempo extrêmement rapide de 240 bpm. Dans ce slam, caractérisé par un rythme effréné, l'impression générale est celle de la peur et de la fuite. Le débit d'articulation est le plus rapide du corpus (4,4 syllabes par seconde en moyenne¹⁷) et le taux de pause le plus réduit (à peine 7% du temps total). La mécanique rythmique est renforcée par un ton très monocorde.

(21) [Philippon ; chaque syllabe est accentuée]
'Ce 'que 'je 'cro 'vais 'ré'el 'est 'il'lus'oïre. 'Ce 'qui 's'est 'pas'sé 'n'est 'pas 'ce 'que 'j'a'vais 'cru 'voir.

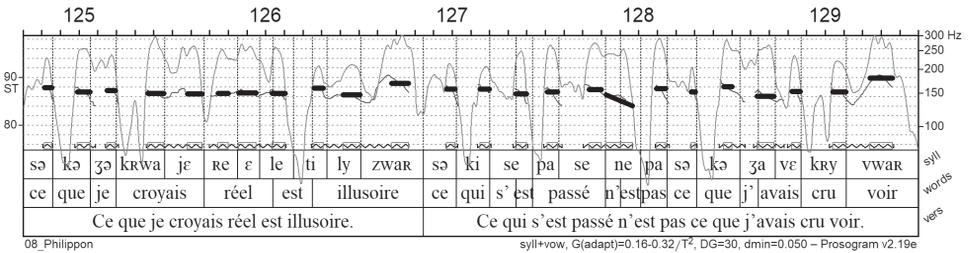


Figure 7 : Rythme staccato et débit rapide (représentation visuelle par un Prosogramme, Mertens 2004) [Philippon]

17. On pourrait s'étonner que ce débit corresponde à ce que Duez et Nishinuma (1987) qualifient de débit « lent » en parole lue. Cependant, lorsque la parole est scandée, c'est aussi l'écart temporel entre syllabes accentuées (le tempo) qui va donner une impression de lenteur ou de vitesse, et pas uniquement la durée des syllabes.

La vitesse sera d'autant plus saillante qu'elle contraste avec un débit plus lent. Dans (22), on passe d'un débit très lent (3,6 syllabes par seconde) à un débit extrêmement rapide (7,04 syllabes par seconde) : ce contraste et le tempo rapide de la scansion (113 bpm) signalent le début d'une frénésie d'activités qui vont peupler la nuit narrée dans le slam.

(22) [NK]		Intervalle moy.	Tempo
Quand je suis entré, moi je cherchais rien (3,60 syll/sec)			
j'voulais 'rien	/		
qu'on me foute la 'paix	/		
j'voulais pla'ner	/		
juste décol'ler	/		
mater le 'cirque	/		
des gens bour'rés	/		
boire une gor'gée	/		
recommen'cer	/		
et oubl'ier (7,04 syll/sec)	/	0,53 s	113 bpm

En outre, plus une scansion est régulière (isochronique et isométrique), plus elle sera perçue comme saillante, surtout si cette régularité est renforcée par des répétitions lexicales (23).

(23) [Skash]	Nb de syl.	Intervalle moy.	Tempo
Je lâche une petite dernière rime			
pour tous mes anc'iens	/ 5		
pour tous mes gen'tils	/ 5		
pour tous mes gens 'bien	/ 5		
pour tous mes gens 'tristes	/ 5		
pour tous mes gens 'ivres	/ 5		
pour toutes les gen'cives	/ 5		
à qui mes pha'langes	/ 5		
ont fait la 'bise	/ 4	0,84 s	72 bpm

La scansion de l'exemple (24) est intéressante tant au plan verbal, par la richesse des allitérations et des paronomases, qu'au plan prosodique, où se construit un rythme syncopé : les syllabes accentuées alternent à intervalles réguliers de 0,6 et 1,1 seconde, pendant toute la durée du morceau qui s'apparente à un jazz, comme le dit le refrain : *Son dada de dandy, c'est d'aider les dames en dandinant son doux dard à tâter du swing.*

(24) [l'Ami Terrien]		Intervalle moy.	Tempo
Des 'dindons	/		
'd'Indochine dode-	/	0,6 s	
'linent du	/	1,1 s	
'dos	/	0,7 s	
'dodus dodos	/	1,1 s	
'd'un zoo d'un	/	0,7 s	
'zozo du	/	1,0 s	
'Siam	/	0,7 s	70 bpm

Pour conclure, on décrit une scansion complexe qui combine des syllabes accentuées initiales et finales de mots. Dans chaque vers transcrit sous (25), on trouve deux battements rythmiques isochrones qui forment ce que Fónagy appelle des arcs accentuels : un arc accentuel apparaît dans un groupe de deux mots lexicaux lorsque la première syllabe du premier mot et la dernière du second sont accentuées. À la scansion, se superpose la mise en évidence de groupes de mots : *secrets de famille, cadavres dans les placards, je-t'aime-pour la vie*, etc. Il y a donc un double tempo, celui de la scansion organisée sur les accents finaux et celui de la scansion imbriquée avec les accents initiaux.

(25) [Petit Poussin]		Nb. syl.	Tempo
combien de 'secrets de fa'mille	/	8	
de 'cadavres dans les pla'cards	/	8	
combien de 'je-t'aime-pour-la-'vie	/	8	
et de tu 'comprendras plus 'tard	/	8	
t'oc'cupe	/	2	40 bpm

Quelque 150 scansions rythmiques accompagnées de régularités du signifiant verbal ont été détectées et annotées dans le corpus : elles forment le phonostyle élaboré. Leur signification varie selon leur tempo, leur métrique et les rimes ou allitérations concomitantes. Elles contribuent à produire des séquences de pulsations communes au slameur et à l'auditoire, invitant à une synchronisation interactionnelle basée sur des régularités sensibles et audibles.

6. Combinaisons des styles vocaux et flow

L'analyse des quatre phonostyles a montré l'étendue des techniques d'écriture et de mise en voix qui constituent les ressources des slameurs. Chaque slam est construit à partir de combinaisons spécifiques de ces phonostyles, avec souvent un style dominant.

Tableau 6 : Proportion de temps de parole occupé par chaque phonostyle.
Le style quantitativement dominant est souligné (en gras)

Slameurs	Phonostyles				Total
	parlé	accentué	poétique	élaboré	
16_Volauvent	48 %	20 %	5 %	27 %	100 %
11_7 ^e Œil	60 %	26 %	0 %	13 %	100 %
09_Renard de Feu	47 %	23 %	6 %	25 %	100 %
20_Josh	16 %	5 %	62 %	18 %	100 %
21_PiY	23 %	25 %	19 %	33 %	100 %
15_Mineige	8 %	43 %	13 %	35 %	100 %
03_NK	10 %	0 %	38 %	52 %	100 %
04_Skash	4 %	35 %	1 %	60 %	100 %
01_Petit Poussin	10 %	4 %	23 %	64 %	100 %
13_Gaëtan Sortet	10 %	15 %	0 %	75 %	100 %
08_Philippon	0 %	11 %	2 %	87 %	100 %
02_Ami Terrien	0 %	0 %	0 %	100 %	100 %
Total général	21 %	19 %	16 %	44 %	100 %

Un premier groupe, composé des slams de Volauvent, 7^e Œil et Renard de Feu, se caractérise par un usage dominant du style parlé et un usage moindre des styles accentué et élaboré. Dans ce groupe, Volauvent se distingue par un débit très lent, des pauses silencieuses longues et des paronomases. Un second groupe, composé des slams de Josh, PiY et Mineige, combine l'ensemble des phonostyles dans des proportions variables. Josh se distingue par une utilisation massive du

style poétique tandis que les styles sont plus panachés dans les deux autres slams. Le troisième groupe rassemble les 6 slams restants : ils se caractérisent par une utilisation majoritaire du style élaboré. Chez Skash, le style accentué est aussi très présent (35 %) ; chez Petit Poussin et NK, le style élaboré se combine avec le style poétique. Chez Gaëtan Sortet, Philippon et Ami Terrien, le style élaboré est quasiment uniformément utilisé.

Des combinaisons particulières de phonostyles forment le flow particulier de chaque slam. Cette description ne constitue pas un profil définitif de chaque slameur, car ils peuvent tout à la fois se réinventer en déclamant leur slam autrement la fois suivante ou utiliser d'autres techniques dans d'autres slams. L'analyse a cependant montré l'étendue des techniques possibles et l'importance de la combinaison de paramètres verbaux et paraverbaux dans la création des rythmes.

7. Conclusion

Dans cette étude, nous avons analysé un corpus de 12 enregistrements afin d'identifier et de définir quatre phonostyles du slam à partir de paramètres verbaux et prosodiques. Au plan méthodologique, nous avons objectivé le rythme à l'aide d'outils d'analyse prosodique disponibles librement pour la communauté scientifique. Au plan théorique, nous avons défini le rythme comme une combinaison de paramètres verbaux et paraverbaux et développé plusieurs pistes pour rendre compte des effets interprétatifs propres à différents rythmes.

L'analyse a permis de d'identifier quatre phonostyles – parlé, poétique, accentué et élaboré. Le style élaboré est le plus fréquent et peut-être le plus typique du slam. Certains slameurs ont catégorisé les slams sous les étiquettes suivantes : « rap », « théâtre » ou « poésie »¹⁸. Ces catégories émiques semblent correspondre aux phonostyles mis au jour par notre analyse : le style identifié comme rap correspondrait aux phonostyles accentué et élaboré, où une scansion rythmique domine dans la performance (Marc-Martínez 2010 : 191), que cette scansion soit irrégulière (style accentué) ou régulière (style élaboré) ; le style « théâtre » correspondrait au style parlé, qui peut s'apparenter à une parole publique ou plus intime selon les paramètres secondaires ; enfin, le style « poésie » correspondrait au style « poétique », ancré principalement sur

18. Cette catégorisation a été évoquée par plusieurs slameurs rencontrés par Aline Colau (UCLouvain) dans le cadre d'entretiens menés pour sa recherche doctorale. Nous la remercions pour cette information qu'elle nous a communiquée.

les jeux sonores (allitérations, rimes, assonances) et exploitant davantage les pauses et les silences que les rythmes scandés.

Ce faisant, nous avons pu montrer que le rythme, reconnu comme caractéristique définitoire et essentielle du slam, se construit de manière complexe en recourant à deux sources indissociables : des éléments verbaux (jeux sur le signifiant sonore) et paraverbaux (modification du débit, de l'accentuation, du tempo). La structure métrique et sonore d'un texte de slam sert de préparation à la performance orale et emprunte « un ensemble de procédés que la poésie écrite a hérité de ses lointaines origines orales, qu'elle a codifiés et a persisté à cultiver longtemps après s'être préférentiellement diffusée par l'écrit, jusqu'au XIX^e siècle, et dont elle n'a eu de cesse de s'affranchir depuis plus d'un siècle : les allitérations, les assonances, les homophonies, les paronymies, les vers fixes, la rime » (Cabot 2017). D'autre part, la maîtrise de la prosodie permet au slameur de varier le débit d'articulation, de créer des scansions et d'utiliser les silences d'une manière emphatique. Ces procédés permettent de créer une parole expressive, vivante, partagée. L'ancrage dans la substance phonétique est fondamental : la matérialité verbale et prosodique créent un sens propre, supplémentaire au sens du texte. Ce sens relève de l'emphase et de l'expressivité, de l'iconicité prosodique ou de l'expression des attitudes et des émotions. En cela, nous avons montré comment l'analyse acoustique peut être au service de la phonostylistique définie par Léon comme un « travail sur la substance et la forme de l'*expression* [manifestant] des sens nouveaux par rapport au discours ordinaire » (1993 : 44).

Références bibliographiques

- Auchlin A. (2013), « Prosodic iconicity and experiential blending », dans Hancil S. & Hirst D. (dir.), *Prosody and Iconicity*, Amsterdam, Benjamins, p. 1-32.
- Auer P., Couper-Kuhlen E. & Müller F. (1999), *Language in Time. The Rhythm and Tempo of Spoken Interaction*, Cambridge, Cambridge University Press.
- Boersma P. & Weenink D. (2018), « Praat: doing phonetics by computer [Computer program] ». En ligne : <<http://www.praat.org/>>.

- Boutet J. (2014), « L'interprétation des faits phonétiques: Dialogue posthume entre John J. Gumperz et Yvan Fonagy », *Langage & Société* 150 (4), p. 71-84.
- Cabot J. (2017), « La scène ouverte de slam : dispositif, situation, politique », dans Cabot J. (dir.), *Performances poétiques*, Lormont, Éditions Nouvelles Cécile Default, p. 79-100.
- Candea M. (2000), *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané : étude sur un corpus de récits en classe de français*, thèse de l'université de la Sorbonne nouvelle-Paris 3.
- Candea M. & Trimaille C. (2015), « Introduction. Phonétique, sociolinguistique, sociophonétique : Histoires parallèles et croisements », *Langage & Société* 151 (1), p. 7-25.
- Couper-Kuhlen E. (1993), *English Speech Rhythm: Form and function in everyday verbal interaction*, Amsterdam, Benjamins.
- Degand L. & Simon A. C. (2009), « Mapping prosody and syntax as discourse strategies: How basic discourse units vary across genres », dans Wichmann A., Barth-Weingarten D. & Dehé N. (dir.), *Where Prosody Meets Pragmatics: Research at the Interface*, Bingley, Emerald, p. 79-105.
- Dessons G. & Meschonnic H. (1998), *Traité du rythme. Des vers et des proses*, Paris, Dunod.
- Di Cristo A. (1999), « Vers une modélisation de l'accentuation du français : première partie », *Journal of French Language Studies* 9 (2), p. 143-179.
- Duez D. & Nishinuma Y. (1987), « Vitesse d'élocution et durée des syllabes et de leurs constituants en français parlé », *TIPA* 11, p. 157-180.
- Fónagy I. (1980), « L'accent en français : accent probabilitaire », *Studia Phonetica* 15, p. 123-233.
- Fónagy I. (1983), *La vive voix. Essai de psycho-phonétique*, Paris, Payot.
- Fónagy I. (1999), « Why iconicity », dans Nänny M. & Fischer O. (dir.), *Iconicity in Language and Literature*, Amsterdam/Philadelphia, Benjamins, p. 3-36.
- Fraisse P. (1974), *Psychologie du rythme*, Paris, PUF.

- Goldman J.-P. (2011), « EasyAlign: an automatic phonetic alignment tool under Praat », dans *Proceedings of InterSpeech*, Florence.
- Goldman J.-P. & Simon A. C. (2020), « ProsoBox, a praat plugin for analysing prosody », dans *Proceedings Speech Prosody*, Tokyo.
- Grosjean F. & Deschamps A. (1972), « Analyse des variables temporelles du français spontané », *Phonetica* 26, p. 129-156.
- Grosman I., Simon A. C. & Degand L. (2018), « Variation de la durée des pauses silencieuses : Impact de la syntaxe, du style de parole et des disfluences », *Langage* 211 (3), p. 13-40.
- Gumperz J. J. (1982), *Discourse Strategies*, Cambridge, Cambridge University Press.
- Jenny L. (2003), « Versification », dans *Méthodes et problèmes. Cours d'initiation aux méthodes et problèmes de littérature française moderne*, département de français moderne, Université de Genève. En ligne : <www.unige.ch/lettres/framo/enseignements/methodes/index.html>.
- Kraxenberger M., Menninghaus W., Roth A., & Scharinger M. (2018), « Prosody-based sound-emotion associations in poetry », *Frontiers in Psychology* 9.
- Lacheret A., Simon A. C., Goldman J.-P. & Avanzi M. (2013), « Prominence perception and accent detection in French: from phonetic processing to grammatical analysis », *Language Sciences* 39, p. 95-106.
- La Zone Slam [site web]. En ligne : <www.lazone.be/slam/>, consulté le 05/07/2019.
- Léon P. (1993), *Précis de phonostylistique. Parole et expressivité*, Paris, Nathan.
- Léon P. (2009), « Nouveau regard sur la phonostylistique », *La Linguistique* 45 (1), p. 159-169.
- Mahiou I. (2015), *Flow et versification : Poétique et scansion, oralité et écriture en rap français*, mémoire de master (M1), dirigé par Marty Ph., université Paul Valéry Montpellier III.
- Marc-Martínez I. (2010), « Voix signifiantes : Le cas du rap français », *Thélème. Revista Complutense de Estudios Franceses* 25, p. 183-195.
- Mertens P. (2004), « Un outil pour la transcription de la prosodie dans les corpus oraux », *Traitement Automatique des Langues* 45 (2), p. 109-130.

- Prsir T. & Simon A. C. (2013), « Iconic interpretation of rhythm in speech », dans Hancil S. & Hirst D. (dir.), *Prosody and Iconicity*, Amsterdam/Philadelphia, Benjamins, p. 161–180.
- Raket S. & Mezmizi M. (2017), *Slam. Poésies et voix de Liège*, Liège, Éditions de la province de Liège.
- Reynolds S. & Vorger C. (2014), « Les pérégrinations du mot slam. Du slam américain au slam français », *Cahiers de lexicologie* 104, p. 219-231.
- Simon A. C. (2004), *La structuration prosodique du discours en français*, Bern, Peter Lang.
- Simon A.-C. & Auchlin A. (2001), « Multi-modal, multi-focal : Les “hors-phase” de la prosodie », dans Cavé C., Guaitella I. & Santi S. (dir.), *Oralité et gestualité. Interaction et comportements multimodaux dans la communication*, Paris, L'Harmattan, p. 629-633.
- Simon A. C. & Grobet A. (2005), « Interprétation des scansion rythmiques en français », dans *Actes du colloque Interface Discours Prosodie (IDP 2005)*, université d'Aix-en-Provence.
- Simon A.C., Avanzi M., Goldman J.-P. (2008), « La détection des proéminences syllabiques. Un aller-retour entre l'annotation manuelle et le traitement automatique », dans *Actes du 1^{er} Congrès mondial de linguistique française*, p. 1685-1698.
- Tyszler C. (2009), « Entre rap et slam : un souffle nouveau dans la langue ? », *Journal français de psychiatrie* 34, p. 16–18.
- Vorger C. (2011), *Poétique du slam: de la scène à l'école. Néologie, néostyle et créativité lexicale*, thèse de doctorat en sciences du langage – linguistique, sous la direction de Grossman F. & Abry-Deffayet D., université de Grenoble.
- Vorger C. (2016), *Slam, une poétique : De Grand Corps malade à Boutchou*, Paris, Les Belles Lettres.
- Vorger C. & Abry D. (2009), « Du rap au slam, le flow ne se tarit pas », *Synergie Espagne* 4, p. 63-76.

Article reçu en octobre 2019. Révision acceptée en avril 2020.