

Séminaire de l'EHESS « Approches pluridisciplinaires du rap »

Mercredi 9 avril 2025 : Charting Musical Boundaries: A Ground-Up Mapping of Rap Music in France Through Actual Consumption Data, par Myriam BOUALAMI et Camille ROTH

Compte-rendu de Thomas Bop

A l'occasion de la séance présentée par Myriam Boualami et Camille Roth, il nous a été partagé les conclusions provisoires de leur article intitulé : *Charting Musical Boundaries : A Ground-Up Mapping of Rap Music in France Through Actual Consumption Data*. Ce travail de recherche qui nous a été présenté est le fruit de trois laboratoires du CNRS qui se sont regroupés à l'occasion du projet ANR RECORDS (pRatiques dEs publiCs des platefORmes De Streaming), réalisé en partenariat avec Deezer. A cette occasion, l'entreprise spécialisée dans le streaming musical a partagé les données de ses utilisateur.ice.s afin de permettre l'étude de l'impact de l'algorithme dans leurs pratiques musicales, ainsi que la production d'une cartographie détaillée du rap et de ses frontières.

Myriam Boualami est doctorante en géographie à l'Université Paris 1, affiliée au Centre Marc Bloch à Berlin. Sa thèse, menée dans le cadre du projet RECORDS, étudie les effets spatiaux de la numérisation de l'industrie musicale, en particulier via les plateformes de streaming. Elle analyse les comportements d'écoute géolocalisés pour cartographier les dynamiques de consommation du rap en France. Camille Roth est directeur d'études à l'EHESS, spécialisé en sciences sociales computationnelles. Il dirige le Centre d'Analyse et de Mathématiques Sociales (CAMS) à Paris et a fondé l'équipe de sciences sociales computationnelles au Centre Marc Bloch. Ses travaux portent sur les réseaux socio-sémantiques, les dynamiques de l'espace public numérique et la modélisation des systèmes sociaux complexes.

Le cadre théorique mobilisé à l'occasion de cette étude par Myriam Boualami et Camille Roth pour étudier les formes de consommation du rap s'appuie sur les analyses bourdieusiennes, leur permettant, dans une double perspective, de tracer les contours du rap et d'en saisir les formes possibles de hiérarchisation sociale.

De même que les outils méthodologiques utilisés ont été empruntés au cadre analytique bourdieusien, notamment avec l'usage de L'Analyse en Composante Principales (ACP) afin d'objectiver les structures de relations et mettre en lumière les « distances » et les « proximités » entre les individus et les modalités sélectionnées – principalement le genre et la génération.

Pour construire leurs analyses, les chercheur.euse.s ont d'abord filtré les utilisateur.ices français.e.s écoutant plus de rap que la moyenne, soit environ 1000 personnes sur un échantillon de 4000. Ce groupe est composé majoritairement de jeunes, sans qu'il y ait de surreprésentation marquée ni d'hommes ni de femmes. Concernant les artistes étudié.e.s, iels se sont appuyé.e.s sur la classification de Deezer en sélectionnant une partie des rappeur.euse.s rangé.es dans la catégorie « Rap », ce qui correspond à 810 artistes. A partir des données de consommation des utilisateur.ice.s, iels ont ensuite eu accès à des millions de morceaux, qu'iels ont trié.e.s en plusieurs groupes, afin de mieux comprendre comment le rap est écouté et consommé.

L'accès aux données permis par Deezer a ouvert une nouvelle voie pour l'étude des pratiques culturelles et du genre rap : à partir de données individuelles – non déclaratives –, de modèles mathématiques et de données informationnelles, iels ont pu reconstituer des catégories de rap, non plus fondées sur un modèle de production industriel *Top-down*, mais selon une logique scientifique *Bottom-up*. Cette démarche présente comme intérêt de permettre de sortir des catégories produites par l'industrie qui invisibilise les différentes pratiques de consommation, pour redonner du sens dans les catégories de la musique rap sur la base du discours artistique des

rappeur.euse.s et des auditeur.ice.s.

Pour éviter la suspicion de coïncidence, iels ont cherché à quantifier dans quelle mesure un.e utilisateur.ice écoute un.e artiste par rapport à sa popularité générale, afin d'écartier toute analyse biaisée qui pourrait suggérer que les préférences observées sont dues au hasard. À partir de ce calcul, il devient alors possible d'identifier des préférences marquées ou des formes d'indifférence dans les pratiques d'écoute.

Ce faisant, à partir de l'hypothèse suivante : *Puisque tou.te.s les auditeur.ice.s ont accès au même catalogue musical, les préférences comme les indifférences deviennent observables, ce qui permet d'émettre l'hypothèse que les distinctions entre préférences individuelles reflètent des positionnements sociaux différenciés*, les recherches de Myriam Boualami et Camille Roth se sont données pour défi de répondre aux trois questions suivantes :

- a. *Comment identifier des classifications musicales collectives à partir de web-données de consommations individualisées, à l'échelle intra-genre ?*
- b. *Peut-on donner du sens (esthétique, social...) à ces constructions mathématiques ground up ?*
- c. *Sont-elles liées aux caractéristiques démographiques des audiences ?*

En l'état actuel de la thèse, il nous a été présenté deux grands points dont l'opérationnalisation des savoirs apparaît comme sérieusement stable et concluant :

Premier point : quatre grands thèmes opposant les artistes apparaissent à partir du traitement et de l'analyse des données actuelles. La première opposition est marquée entre les artistes « *savants* » (underground, niche...) et les artistes « *populaires* » (mainstream, tendance...). Une tendance dans les pratiques culturelles qui apparaît est alors que ceux qui écoutent des « petits » rappeur.euse.s n'écoutent pas de « gros » rappeur.euse.s. La seconde opposition se fait en termes de génération : l'audience du rap se scinde en deux générations, celle du rap des années 80 face à celle des années 2000. La troisième opposition se joue sur le thème que Myriam Boualami et Camille Roth nomment l'« *éthique* » – à savoir les thèmes abordés, tels que l'amour et les récits gangster, les deux catégories s'opposant par leurs registres symboliques. Enfin, quatrième et dernière opposition : celle des nationalités. Une ligne de démarcation semble exister entre les artistes francophones et internationaux, suggérant une forme de segmentation linguistique et géographique. Apparaissent alors, à travers ces thèmes, des catégories pouvant être mobilisées par la suite pour poursuivre et approfondir leurs analyses : « *rap savant* », « *rap populaire* », « *rap des années 80* » (old school), « *rap des années 2000* » (nouvelles générations), « *rap francophone* » et « *rap international* » ainsi que des thèmes comme l'amour ou l'univers gangster...

Deuxième point : une présentation détaillée des frontières du rap, fondée sur la construction de sept clusters, s'appuyant sur une nomenclature précise des catégories et des esthétiques. Cette cartographie s'accompagne d'une analyse des vocabulaires récurrents, des caractéristiques propres à chaque cluster, ainsi que des principales figures associées à chaque forme de rap. Le premier cluster est le cluster « *américain* » : on y retrouve les mots « *love* », « *gril* » et « *die* ». Il se caractérise par la nouvelle génération, une esthétique « *gangster* » et une inscription plus marquée sur la scène internationale. Le deuxième cluster est celui des artistes dits « *conscients* » : les morceaux mobilisent un vocabulaire savant, souvent associé à une ancienne génération de rappeur.euse.s et d'auditeur.ice.s. Il s'agit d'une écoute de niche, principalement francophone. On y retrouve des figures comme IAM ou Tandem. Le troisième cluster, nommé « *éthéré* » : rassemble des artistes dont les morceaux sont traversés par des mots comme « *ciel* », « *piège* », « *fleurs* », « *anges* » ou « *paradis* ». Il s'agit d'un rap savant, à l'esthétique de niche, porté notamment par Damso ou NES. Le quatrième cluster est celui dit « *gloomy* » : les morceaux sont ponctués de mots comme « *adieu* », «

pluie », « souvenir » ou « cendre ». Il s'inscrit dans une esthétique sentimentale, mêlant une écoute de niche, une nouvelle génération et un registre plutôt savant. Le cinquième cluster correspond à une catégorie « *pop* » : il se distingue par l'absence de thématique propre, les artistes reprenant des vocabulaires empruntés aux autres catégories. Il s'agit d'un rap très écouté, grand public, avec des figures comme JUL ou GIMS. Le sixième cluster est le cluster « *rough* » : il mobilise un lexique dur et explicite, autour du crime organisé, de la « purge » ou des « billets ». On y retrouve des artistes de la nouvelle génération, porteurs d'une esthétique gangster, non-savante, avec une écoute plutôt de niche – comme Fresh la Douille. Enfin, le septième cluster, dit « *soft-heart* » : se distingue par des mots comme « *Habibi* », « *loca* », « *chérie* » ou « *piqué* ». C'est un rap à tonalité amoureuse, porté par la nouvelle génération, francophone, non-savante et écouté en marge du mainstream, avec par exemple l'artiste Melina.

Cette catégorisation fait éclater de nouvelles couleurs, permettant de recatégoriser les formes de rap sur le cercle chromatique de ses pratiques, offrant ainsi, à travers une approche *ground up*, une vision plus nuancée et authentique des dynamiques de production et de consommation culturelles rap – lequel ne saurait être réduit à une entité monolithique. Une fois ces nouvelles catégories établies, on peut les croiser avec les caractéristiques démographiques de leur audience. A partir d'un modèle de régression, Myriam Boualami et Camille Roth analysent comment l'écoute de tel ou tel artiste varie selon les groupes démographiques par âge genre des auditeur.ice.s.

Voici un exposé des tendances qui se dessinent : concernant le cluster « *américain* », aucune composition démographique particulière n'est retenue, aucune classe ne présente de préférence. Concernant les clusters « *éthéré* », « *gloomy* » et « *soft-heart* », les femmes âgées de 12 à 24 ans présentent une nette préférence, là où les hommes de 35 à 44 ans y sont sous-représentés, soulignant une certaine indifférence. Le cluster « *conscients* » est, quant à lui, marqué par une forte préférence chez les 25-44 ans, tous genres confondus. Les femmes adultes les plus âgées ne manifestent pas de préférences marquées : ce sont les moins spécialisées dans le rap. Les hommes les plus âgés, tout comme les 25-34 ans, partagent une affinité pour le rap « *conscious* ». Les jeunes femmes rejoignent les femmes adultes dans leurs goûts, traduisant une préférence féminine marquée pour les clusters à tonalité « *émotionnelle* » ou « *savante* », tandis que les préférences des plus jeunes hommes se rapprochent davantage de celles des femmes, toutes tranches d'âge confondues, que de celles des hommes plus âgés.

Une tendance dont il fait bon cœur de se réjouir, tant ce problème a caractérisé et continue de marquer le rap, réside dans la féminisation de l'audience et la présence de plus en plus importante des femmes dans cet univers. Là, où, auparavant, l'étiquette de « connaisseur » ou de « puriste » du rap était l'apanage exclusif des hommes – tel que cela a pu être démontré dans les travaux de Karim Hammou¹ ou encore Paulo Higgins² – il apparaît que les jeunes femmes, au même titre que les hommes, se spécialisent. Ce constat donne alors à espérer l'avènement d'une remise en question des différences genrées historiquement présentes dans le rap.

D'un point de vue personnel, cet exposé m'a permis d'ouvrir mon regard sur la recherche appliquée. Ayant développé une certaine préférence pour la recherche fondamentale, en dépit de la recherche appliquée qui, avec un certain recul critique, m'apparaît comme pouvant servir des entreprises intellectuelles et des intérêts économiques avec lesquels je doute que la production de savoir doive chercher à se concilier, il m'est apparu un exemple de recherche appliquée tout à fait enrichissant et, me semble-t-il, travaillant autant au compte de la recherche que du rap.

¹Hammou, K. (2024, septembre 30). « C'est ça l'égalité » ? Les hommes amateurs de rap et les rappeuses [Billet]. *Sur un son rap*.

²Higgins, P. (2020). Femmes et Queers : Des publics subalternes et cachés du rap français ? *Volume!*, 172(2), 185-193.

Cette recherche m'apparaît comme particulièrement brillante sur le plan scientifique. Dans le cadre de mon mémoire consacré au rap, je me suis intéressé à la littérature issue de différentes disciplines afin de mieux saisir ce que la recherche en dit et comment elle le présente. La majorité de mes lectures reposaient sur des approches qualitatives, et lorsque ce n'était pas le cas – ce qui restait très rare – la collecte et l'analyse de données à grande échelle se révélaient limitées, tant le rap constitue un objet vaste et complexe à appréhender. Je pense particulièrement aux travaux de Benoît de Courson, notamment son article intitulé *Une exploration numérique du rap français depuis les années 1990*³(2024), dans lequel il a étudié les données textuelles d'un grand nombre de morceaux de rap en les soumettant à un traitement lexicométrique. Pour opérer, il a constitué un large corpus de chansons à partir du site *Genius*. Bien qu'il soit parvenu à de nombreuses conclusions, comme la présentation de l'évolution du rap français depuis 1990 jusqu'aux années 2023, son corpus présente certaines limites, liées à la qualité inégale des transcriptions et à la sélection des morceaux disponibles sur la plateforme *Genius*. Par conséquent, les travaux actuels de Myriam Boualami et Camille Roth, en collaboration avec Deezer, permettent l'ouverture à de grandes bases de données offrant à la science de pouvoir travailler à partir d'échantillons importants, et de fait, de produire des conclusions scientifiques fiables, méthodiquement établies et inédites, tant l'accès à ces ressources en dehors d'un tel cadre collaboratif ne peut quasiment être opéré.

D'un point de vue épistémologique, la manière de faire science des deux chercheur.euse.s est intéressante car l'étude du rap ne se fait pas depuis une position extérieure, prétendant le comprendre de loin, mais au contraire, elle repose sur une approche par le bas, c'est-à-dire qu'elle comprend le rap à partir de ses acteur.ice.s. Ainsi, la production des savoirs se fait dès lors à partir d'une position non pas surplombante, où l'objectivité des chercheur.euse.s serait la condition de la bonne conduite de cette étude, mais à l'inverse, elle part du sensible, des subjectivités de tout.e.s celleux qui constituent le rap, pour en rendre visibles les logiques à l'extérieur en partant de l'intérieur. Cette nouvelle catégorisation apporte aussi au rap et à ses acteur.ice.s une classification digne, qui ne répond pas en premier lieu à des intérêts marketing et industriels, mais à une catégorisation fidèle à la réalité de la musique rap.

En conclusion, cette recherche propose une approche novatrice et passionnante du rap, qui redéfinit ses frontières et dévoile des dynamiques de consommation inédites, invitant ainsi à une exploration plus profonde et authentique d'un genre musical souvent réduit à des clichés.

³Courson, B. de. (2024). *LRFAF : Une exploration numérique du rap français depuis les années 1990*. OSF