



**HAL**  
open science

# Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. Une analyse quantitative d'un panel d'utilisateurs de streaming

Jean-Samuel Beuscart, Sisley Maillard, Samuel Coavoux

## ► To cite this version:

Jean-Samuel Beuscart, Sisley Maillard, Samuel Coavoux. Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. Une analyse quantitative d'un panel d'utilisateurs de streaming. 2017. halshs-01639788v1

**HAL Id: halshs-01639788**

**<https://shs.hal.science/halshs-01639788v1>**

Preprint submitted on 20 Nov 2017 (v1), last revised 24 Apr 2019 (v2)

**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.

## **Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. Une analyse quantitative d'un panel d'utilisateurs de streaming**

Working paper. Version étendue et retravaillée d'une communication présentée au colloque Musiques et publics jeunes à l'ère de la musimorphose, Faculté de Musique / Université de Montréal, 8-10 Juin 2017.

### **Introduction**

Le streaming est devenu une composante importante de l'écoute musicale. Les données d'usage disponibles sont assez éparpillées et divergentes, mais les enquêtes suggèrent qu'environ la moitié des internautes, et les trois quarts des jeunes internautes, écoutent régulièrement de la musique en streaming (GfK, 2015 ; IFPI, 2016). En France, on compte 13 millions d'utilisateurs mensuels du streaming audio, dont 3,7 millions d'abonnés payants ; les usages gratuits comme payants ont été multipliés par deux entre 2014 et 2016 (SNEP, 2017). Le succès des plateformes spécialisées (Spotify, Deezer) ou généralistes (YouTube) se traduit par une part accrue des revenus du streaming dans le chiffre d'affaires de l'industrie musicale – même si de nombreux acteurs trouvent cette part encore insuffisante : ils représentent, sur le marché français, 78% des revenus numériques, et 32% de l'ensemble du marché. Ces indicateurs en forte hausse témoignent dans l'ensemble de l'installation du streaming comme régime de consommation de la musique.

L'environnement sociotechnique du streaming met l'auditeur dans une position spécifique, à maints égards différente de celle de l'auditeur de disque ou de musique téléchargée. Il met à sa disposition un catalogue virtuellement illimité, ayant une prétention à l'exhaustivité : Deezer comme Spotify offrent environ 30 millions de titres immédiatement accessibles, Youtube plus encore. Pour guider l'auditrice dans ce catalogue, les services proposent une variété d'outils algorithmiques suggérant ou choisissant des titres : depuis l'outil de filtrage collaboratif (« les gens qui ont aimé ce titre que vous avez écouté ont aussi aimé celui-ci »), de recommandation thématique (« les autres titres de cet artiste, de ce sous-genre du hip-hop que vous avez écouté ») jusqu'aux radios thématiques (« radio inspirée de cet artiste ») ou aux playlists complètement « inspirée de vos goûts » (le « flow » de Deezer, les « découvertes de la semaine » de Spotify), en passant par les nombreuses « ambiances » mises en playlists pour s'adapter aux humeurs de l'auditrice. Ces algorithmes sont en constante évolution ; Spotify annonce le « zero button music player », un outil de sélection musicale fondé non seulement sur les goûts, mais aussi sur l'activité de l'utilisateur, telle que mesurée par les capteurs du smartphone<sup>1</sup>.

Les plateformes de streaming se situent, en matière d'agentivité, à mi-chemin entre deux techniques d'écoute musicale : la collection de disque et la radio. Elles construisent une figure d'auditeur potentiellement plus assistée et dirigée que l'écoute de disque. Alors que, pour reprendre l'expression d'Antoine Hennion, « on choisit deux fois un disque », lors de son achat et lors de son écoute sur la platine<sup>2</sup>, il est possible qu'un morceau ou album écouté en streaming n'ait jamais été « choisi » - c'est-à-dire consciemment et intentionnellement sélectionné dans un univers de possibles - mais entièrement suggéré ou mis en playlist par le service de streaming en fonction de logiques plus

---

<sup>1</sup> Sur le blog de Paul Lamere, fondateur de The Echo Nest et directeur du développement chez Spotify, <https://musicmachinery.com/2014/01/14/the-zero-button-music-player-2/>. Consulté le 10/11/2017.

<sup>2</sup> A. Hennion, S. Maisonneuve, E. Gomart, Figures de l'amateur. Formes, objets, pratiques de l'amour de la musique aujourd'hui, La Documentation Française, p.233.

ou moins transparentes. Par rapport à la radio, l'auditeur peut se trouver dans une position similaire de choisir un flux auquel il délègue le choix ; mais ce flux est plus individualisé que celui, de masse, de la radio. Alors que cette dernière se fonde sur la répétition d'un petit nombre de morceaux choisis pour correspondre aux goûts attribués à un public cible décrit par des indicateurs socio-démographiques grossiers, notamment l'âge, le flux de streaming puise dans un large catalogue et adapte sa proposition au profil individuel. Ce flux peut-être interrompu et réorienté à tout moment par l'utilisateur. Le choix n'est cependant pas entièrement idiosyncratique : les algorithmes des services de streaming s'appuient le plus souvent sur les comportements des utilisateurs au profil similaire pour sélectionner des propositions.

Tout comme le format MP3 contient une figure d'un utilisateur accumulateur compulsif (Sterne, 2008), les plateformes et leurs algorithmes dessinent une figure particulière de l'amateur de musique : il s'agit d'un auditeur défini par ses traces d'activité et pris en charge par des machines, guidé au sein de l'océan musical par des robots infaillibles. L'exercice du jugement de goût est suspendu, ou réduit à des acquiescements et refus (passer le morceau proposé). Pour reprendre les termes de Karpik (2007), le consommateur est ici passif (il délègue le choix) et hétéronome (il endosse les goûts porté par le dispositif). Cette figure est à l'opposé de la figure classique de l'amateur de musique, activement engagé dans le choix des musiques permettant la félicité, et attentif à la définition et au maintien de ses goûts personnels (ibid, p. 144). Cette figure de l'auditeur donne de fortes prises à la critique, puis qu'elle dépouille l'amateur du contrôle de ses goûts, de son rôle moteur dans l'interaction avec les œuvres, de son activité parfois virtuose dans l'entretien de ses passions.

C'est à la réalité de cette figure que s'intéresse notre recherche : dans quelle mesure la figure dessinée par les interfaces se réalise-t-elle, se traduit-elle dans les comportements réels des auditeurs ? Les utilisateurs suivent-ils les recommandations des algorithmes ? Dans quels contextes, avec quelle régularité, quels effets ? Comment ces recommandations sont-elles perçues, comprises ? Pour répondre à ces questions, nous nous appuyons sur un corpus de traces d'activité d'un échantillon de 4000 utilisateurs de Deezer, observés pendant une période de cinq mois, au cours de laquelle ils ont réalisé 17 millions d'écoutes<sup>3</sup>. Pour environ la moitié d'entre elles, il est possible de caractériser le contexte de l'interface dans lequel elles ont été réalisées, notamment dans quelle mesure elles ont été induites par les algorithmes. A partir de ces données, nous suggérons que la figure de l'auditeur algorithmisé est encore (très) minoritaire dans les usages ; l'essentiel des écoutes sont issues d'un choix explicite et conscient de l'auditeur, et les modes d'exploration actifs et autonomes dominant. L'enjeu est de préciser dans quels contextes, et par quels types d'auditeurs, les algorithmes de recommandation sont suivis : existe-t-il plutôt des auditeurs particulièrement passifs, ou des moments de passivité partagés par tous les auditeurs ?

Cette réflexion s'inscrit dans le cadre d'une réflexion plus globale sur l'agentivité des algorithmes, sur la capacité des individus à s'appuyer sur les algorithmes sans pour autant leur abandonner le contrôle de leur expérience. Il s'agit de comprendre comment, dans un monde d'agentivité de plus en plus distribuée entre l'amateur et ses prothèses numériques, se construisent des situations d'autonomie et d'hétéronomie.

## **1 – La recommandation algorithmique, assistance ou contrôle ?**

Dans la littérature, les algorithmes de recommandation tels que ceux mis en place par les plateformes de streaming sont traités de deux façons principales : les travaux en économie et en

---

<sup>3</sup> Une première présentation de ces résultats figure dans la thèse de Sisley Maillard (Maillard, 2015).

marketing s'efforcent de mesurer leur efficacité, tandis que les travaux issus des *médias studies* mettent l'accent critique sur les potentialités de contrôle portées par les dispositifs.

### **1 / La mesure de l'efficacité des recommandations**

Une première tradition d'analyse s'intéresse, depuis les travaux de l'école de Columbia, aux formes d'influence et de recommandation autour de la consommation des produits. Depuis les travaux de Katz et Lazarsfeld (1955), de nombreuses recherches explorent, dans le cas des biens culturels notamment, le rôle respectif de l'entourage, des critiques, de la publicité, etc., sur la décision d'acheter un disque ou d'aller voir un film (pour une synthèse voir Beuscart, Mellet, 2012). Si le terme parfois utilisé d'« influence » peut laisser penser le contraire, ces travaux discutent peu la légitimité des différentes formes de sollicitation, et sont avant tout centrés sur la question de l'efficacité relative des différentes formes de « recommandation » qui sont faites au consommateur : il s'agit de montrer par exemple que la critique n'a d'effet que sur l'audience des films art et essai, tandis que la publicité accroît significativement des audiences des blockbusters (Prag et Casavant, 1994).

Dans cette perspective, les différentes formes de guidage numérique sont perçues comme de nouveaux outils de recommandation. Du fait de la disponibilité des données, de nombreux travaux s'intéressent ainsi au « bouche à oreille numérique » (eWoM) : ils étudient l'influence des avis de consommateurs sur les ventes de livres ou de disques (Chevalier et Mayzlin, 2006), l'impact des classements sur l'audience des morceaux de musique (Salganik et al., 2006), les effets des mentions sur les médias sociaux sur les ventes (Bastard et al., 2012), le lien entre le nombre de tweets à propos d'un film et les entrées au box-office (Asur et Huberman, 2011). En discussion plus ou moins explicite avec l'hypothèse de la longue traîne, la numérisation des pratiques culturelles est vue comme un double mouvement : d'un côté l'élargissement de l'offre, tendant vers l'illimité, de l'autre l'enrichissement des prises offertes aux consommateurs pour se faire un avis et naviguer cette abondance.

Les recherches sur la recommandation algorithmique proprement dite sont plus rares, du fait de la difficulté d'obtenir des données permettant d'évaluer leur efficacité. L'essentiel des travaux portent sur le filtrage collaboratif d'Amazon, qui établit des proximités statistiques entre les clients selon leurs paniers d'achat, et proposent un retour sous la forme « les clients qui ont aimé ce que vous venez d'acheter ont aussi aimé... ». L'un des points de discussion est ici de savoir comment mesurer l'efficacité de ces systèmes. Tous les travaux constatent ainsi qu'une partie non-négligeable des achats sur les sites de e-commerce se font suite à des liens de recommandations algorithmiques (jusqu'à 15% sur Amazon) ; mais il est plus difficile de mesurer quels sont, parmi ces achats, ceux qui n'auraient pas été conclus en l'absence d'algorithme (Sharma, Hofman, Watts, 2016).

### **2 / Critique de la raison algorithmique**

Un second ensemble de travaux prend pour point de départ le rôle croissant des algorithmes dans nos sociétés. Issus de la sociologie, de la communication, de la théorie des médias, ils documentent la prolifération des agents socio-informatiques dans différents domaines de la vie sociale (finance, culture, médias, consommation, travail). Une grande partie adopte une perspective critique quant au pouvoir pris par ces algorithmes. Ils soulignent l'absence de transparence, d'*accountability* de ces systèmes : ils agissent comme des boîtes noires, selon des principes qui ne sont ni clairs, ni connus, ni discutés, ni discutables (Pasquale, 2015 ; Gillespie, 2016). Ils insistent en outre sur leur effectivité : les algorithmes transforment le monde, ils représentent un pouvoir mal perçu sur les personnes et leurs environnements (Lasch, 2007 ; Galloway, 2012).

Les travaux plus empiriques soulignent néanmoins le caractère pluriel et difficile à saisir des algorithmes. S'ils ont un pouvoir, le lieu de ce pouvoir est souvent fuyant. Comme le rappellent Seyfert et Roberge (2016), il n'existe pas en général d'interprétation univoque de l'algorithme et de

ses effets. Ils doivent toujours être dépliés, ils ne peuvent qu’être interprétés ; ceci d’autant plus qu’ils sont le plus souvent imbriqués avec d’autres, interconnectés, pris dans des contraintes techniques locales. Leurs effets sont dépendants des contextes, évoluent avec les usages qui en sont faits et en fonction des données de leur environnement. Leur performativité répond moins au modèle de la règle à laquelle on obéit qu’à celui d’une agentivité distribuée entre les lignes de codes et les différentes catégories d’utilisateurs (Beuscart et Mellet, 2016). Seyfert et Roberge proposent de parler de « cultures algorithmiques » pour désigner ces ensembles de performativités et de significations qui émergent des usages des algorithmes (Seyfert et Roberge, 2016).

### **3 / Le projet des plateformes de streaming**

Le projet des principales plateformes de streaming offre en effet un rôle important aux algorithmes de recommandation. Au fur et à mesure de leur évolution, les services tels que Deezer, Spotify, LastFM développent les outils de suggestion de titres, les diversifient et leur accordent une place croissante dans le design du site. Alors que la page d’accueil de Deezer ou de Spotify présentait initialement les albums écoutés et archivés par l’utilisateur, à la manière d’une discothèque personnelle, elle organise aujourd’hui une variété de suggestions d’artistes, d’albums et de playlists inspirés des écoutes passées des utilisateurs ; en outre, elles mettent en avant diverses modalités de flux musicaux (radio de l’artiste, découvertes de la semaine, ambiances, « flow », etc.) qui prennent complètement en charge le choix de la musique à diffuser. Sur YouTube, chaque écoute de musique est automatiquement suivie d’une autre sélectionnée par l’algorithme de recommandation, qui élabore la playlist de façon ininterrompue en l’absence d’action de l’utilisateur. Le centre de recherche de Spotify évoque le projet d’un « zero button music player », soit d’un service qui choisisse la musique en l’absence de toute action de l’utilisateur, à partir de la connaissance de ses habitudes, ses écoutes passées et des informations issues de son smartphone (Melançon, 2016). Se dessine ainsi une figure de l’utilisateur pris en charge par l’algorithme, déchargé de la tâche de choisir la musique du moment. Morris (2015), dans une étude du service de profilage musical the Echo Nest racheté par Spotify en 2014, décrit ainsi la figure d’un utilisateur dessiné de façon très précise à travers ses goûts, classé au sein de micro-segments de marché, et nourri par des flux très individualisés. Il note néanmoins le maintien d’une figure complémentaire de la curation humaine, via les playlists partagées ; la culture algorithmique qui émerge sur ces sites reste – pour le moment du moins – hybride. Symétriquement, des études qualitatives se penchent sur l’expérience musicale des services de streaming, du moins sur les discours que ses utilisateurs portent sur leurs pratiques et les figures qu’ils dessinent. Weiss Hanrahan (2016) suggère, à partir de focus group, que le modèle de l’écoute en streaming est celui du flux satisfaisant, de l’évitement de la mauvaise surprise, proche d’une radio très individualisée et capable de s’adapter au moment, mais à forte distance du modèle classique de l’amateur faisant l’expérience esthétique d’une œuvre ; autrement dit, les utilisateurs interrogés viennent habiter, au moins dans leurs discours en focus group, la figure proposée par les plateformes. Dans le même registre, Hagen (2016) évoque une pluralité des métaphores de l’usage du streaming, qui toutes s’éloignent de la figure de l’auditeur choisissant de prêter attention à une œuvre.

L’écoute de musique apparaît ainsi comme un cas emblématique d’un questionnement plus large sur la performativité des algorithmes ; leur omniprésence sur les plateformes de streaming dessine une figure de l’auditeur calculé et pris en charge, en opposition à la figure de l’amateur explorant les œuvres. Au-delà de cette figure extrême, brandie tant par les développeurs des plateformes que par les critiques les plus virulents, il est très probable que la culture algorithmique qui émerge sur des usages est plurielle et composite. Nos données permettent d’en esquisser des grandes lignes.

### **2 / Un panel d’utilisateurs d’une plateforme de streaming**

## 1 / Données et méthode

Nos données décrivent la consommation d'un échantillon représentatif d'utilisateurs d'une plateforme de streaming musical, suivis pendant cinq mois. L'échantillon est constitué d'environ 4 000 utilisateurs inscrits en France, sélectionnés aléatoirement parmi ceux qui ont utilisé la plateforme au cours de la première semaine d'avril 2014. Leurs écoutes ont été enregistrées du 1er avril au 31 août 2014, aboutissant à une base d'un peu plus de 20 millions d'observations. Les données collectées nous renseignent à la fois sur le profil des utilisateurs (âge, sexe, lieu de résidence, date d'inscription, abonnement, titres et artistes favoris), sur leur consommation (date, durée, supports d'écoute), et sur les caractéristiques des contenus écoutés (artiste, titre, date de sortie). En particulier, la variable « contexte d'écoute » nous permet d'appréhender par quel type de dispositif, de recommandation ou de collection, un consommateur a sélectionné un titre à écouter.

La plateforme propose un catalogue de 35 millions de titres, et deux offres illimitées, sur le modèle économique du « freemium » : l'une est gratuite et financée par la publicité, l'autre est payante, à un tarif de 9.99 euros par mois. La formule payante permet à l'utilisateur d'écouter sa musique en haute qualité, sans publicité, et surtout en mode hors connexion. Une offre intermédiaire à 4,99 euros par mois, aujourd'hui disparue, n'offrait pas de mode hors connexion. Toutes les offres sont disponibles sur l'ordinateur et les terminaux mobiles, comme les smartphones et les tablettes (via des applications pour iOS, Android et Windows). Ces radios sont (i) une radio personnalisée en fonction des goûts de l'utilisateur et enrichie de recommandations jouées aléatoirement, (ii) des radios éditoriales et thématiques, (iii) des radios algorithmiques basés sur les similarités entre artiste, et enfin, (iv) des radios construites à partir des *playlists* de l'utilisateur.

Après retraitement des valeurs aberrantes, notre échantillon compte précisément 3 813 utilisateurs inscrits en France. Il s'agit d'une population plutôt masculine, jeune et urbaine. On compte seulement 37% de femmes au sein de l'échantillon. L'âge médian des utilisateurs est de 28 ans et 80% de l'échantillon a entre 17 et 46 ans. Notre échantillon est très urbain. Parmi les 75% d'utilisateurs à avoir renseigné leur lieu de résidence, 14% déclare résider à Paris (qui rassemble un peu plus de 3% de la population française) et 73% vivent dans un grand pôle urbain (contre 59% de la population française). Il s'agit dans l'ensemble d'utilisateurs plutôt expérimentés : la majorité des utilisateurs (53%) ont rejoint le service depuis plus de deux ans, et 18% depuis plus de cinq ans. Seul un quart des utilisateurs se sont inscrit assez récemment, il y a moins d'un an.

Le mobile et l'ordinateur sont les supports privilégiés d'écoute (la tablette et le navigateur web, les deux autres options possibles, sont marginales). En moyenne, sur l'ensemble de l'échantillon, un utilisateur se sert du support mobile pour deux tiers de ses écoutes (63%) et utilise l'ordinateur pour le tiers restant. Le type d'offre souscrit influence évidemment les supports utilisés. Parmi nos utilisateurs, 58% ont souscrit à l'offre premium à 9,99 euros, qui permet l'écoute illimitée sur mobile et le mode écoute hors connexion, 27% ont souscrit à l'offre premium à 4,99 euros, sans publicité, et 15% à l'offre gratuite incluant de la publicité. Les abonnés aux offres premium, sans publicité, utilisent plus fréquemment, parfois même presque exclusivement, le mobile pour écouter de la musique, contrairement aux abonnés gratuits, qui utilisent plutôt l'ordinateur. Néanmoins, même pour les abonnés payant qui en bénéficient, l'option hors connexion, pour écouter la musique hors ligne, est relativement peu utilisée par les abonnés. Les écoutes qu'ils ont réalisées hors ligne représentent seulement 11% de leurs écoutes, principalement sur le mobile.

Offre	Support	Fréquence moyenne d'utilisation	Utilisation du mode hors-connexion
Gratuite	téléphone	9%	-
	ordinateur	84%	-
Payante	téléphone	73%	15%

	ordinateur	25%	1%
--	------------	-----	----

Table 1: Utilisation des supports en fonction du type d'abonnement

## 2 / les volumes consommés

Au cours de la période d'observation de cinq mois, on enregistre plus de 17 millions d'écoutes par les usagers de l'échantillon. Cependant, ce chiffre doit être pondéré par la durée réelle d'écoute des titres sur la plateforme : seule un peu plus de la moitié des écoutes (56%) vont jusqu'à la fin du titre, et un tiers des écoutes ne durent pas plus de 30 secondes. Ces écoutes ont porté sur près d'un million de titres différents. Parmi les titres écoutés au moins une fois plus de 30 secondes (81%), un titre est écouté en moyenne 16 fois, soit 4 écoutes pour 1000 utilisateurs. La distribution du nombre d'écoute est sans surprise très asymétrique : 39% des titres ont été écoutés une seule fois par un seul utilisateur, tandis que 5% l'ont été plus de 50 fois. Comme attendu, le « top » des titres affiche des audiences considérables : les cinq titres les plus écoutés au cours de la période de cinq mois ont cumulé chacun plus de 25 000 écoutes et chacun a été écouté par presque la moitié des individus de l'échantillon. Cela correspond donc à une moyenne de 13 écoutes par usager ayant écouté le titre au moins une fois. Les titres concernés sont *Stolen Dance* (Milky Chance), *Sur ma route* (Black M), *Rather Be* (Clean Bandit), *Waves* (Mr. Probz) et *Liar Liar* (Cris Cab feat. Pharell Williams).

L'intensité d'usage est assez élevée : en moyenne, les utilisateurs se sont connectés à la plateforme 86 jours au cours des cinq mois (soit un jour sur deux), pour écouter près de 50 titres chaque jour de connexion. Ces moyennes cachent bien sûr des disparités très importantes, comme le résume le tableau ci-dessous : les 10% les plus actifs ont réalisé plus de 10 000 écoutes au cours de la période, tandis que les 10% les moins actifs en ont réalisé moins de 300.

Variables	moy	med	sd	Min	Max	.01	.10	.25	.50	.75	.90	.99
Nombre d'écoutes totales	4617	300	5064	101	4121	133	422	1180	3006	6202	1079	2421
Nombre de jours d'utilisation de la plateforme	86	93	46	1	154	5	17	46	93	127	145	153
Nombre moyen d'écoutes lors des jours d'utilisation	47	36	36	3	412	9	17	25	36	57	88	187
Consommation quotidienne de musique (durée d'écoute)	1h 50m	1h 23m	1h 39m	0h 1m	19h 5m	0h 1m	0h 30m	0h 51m	1h 23m	2h 13m	3h 39m	8h 28m

Table 2: Intensité de la consommation musicale

Comment caractériser la diversité consommée sur les plateformes de streaming ? La promesse est que, lorsque l'offre est vaste, gratuite ou forfaitaire, et les outils de recommandation plus perfectionnés, la consommation musicale devrait être plus diversifiée sur les plateformes en streaming que sur les marchés existants du disque et du téléchargement. Notons d'abord que la diversité est difficile à mesurer et à résumer, tant il s'agit d'un concept multidimensionnel (Stirling 2007). Dans ce chapitre, nous ne nous intéressons qu'à une de ces dimensions, la variété, mesurée par le nombre d'artistes différents écoutés. Pour notre échantillon d'environ 4 000 utilisateurs, les écoutes portent sur 925 555 titres différents, soit 3% des titres présents dans le catalogue. Elles correspondent également à un peu plus de 3% des artistes et 6% des albums proposés sur Deezer (114 182 artistes et 226 939 albums différents). Ces chiffres portent cependant sur un échantillon. Une extrapolation grossière, à partir de ces résultats, laisse penser que la quasi-totalité du catalogue serait consommé à un moment ou à un autre par deux millions d'abonnés. Cependant, nous ne disposons pas de moyen de savoir combien de personnes sont réellement actives sur la plateforme - notre mode d'échantillonnage surreprésente mécaniquement les usagers les plus actifs- et d'estimer ainsi la diversité globale de l'écoute.

La comparaison de cette variété avec d'autres médias est délicate. Les conditions d'écoute ne sont pas les mêmes : par rapport au disque ou au téléchargement, l'écoute d'un morceau en streaming n'implique pas de dépense supplémentaire ; par rapport à la radio, le rôle du consommateur dans le choix des morceaux est – jusqu'à preuve du contraire – plus actif, et l'offre est bien moins contrainte. Par ailleurs, les données issues des études de la diversité de l'industrie musicale (Bourreau et al., 2011) concernent des flux (le nombre d'albums vendus par semaine), et non des stocks (le nombre d'albums possédés et écoutés), et sous-représentent donc la consommation effective. On constate cependant que le nombre de titres écoutés sur Deezer par seulement 4000 utilisateurs au cours de cinq mois représente environ 9 fois plus de titres que la diffusion de l'ensemble des radios françaises au cours d'une année (d'après l'Observatoire de la musique également)<sup>4</sup>.

La variété de la consommation varie fortement selon les individus. Le tableau ci-dessous décrit le nombre de titres et d'artistes différents écoutés par l'échantillon. Un utilisateur écoute en moyenne 1 149 titres différents ; 5% en ont écouté le triple, alors qu'un quart de notre échantillon n'en a pas écouté plus de 359.

Variable	mo y	me d	sd	min	ma x	.01	.10	.25	.50	.75	.90	.99
Nombre de titres différents écoutés	114	789	121	9	130	46	157	359	789	150	251	578
Nombre d'artistes différents écoutés	9		1		42					2	9	2
	402	278	412	1	381	12	57	129	278	527	877	202
					1							8

Table 3: Diversité de la consommation musicale

Une étude simple des corrélations linéaires montre qu'il existe, avec la diversité consommée : (i) une corrélation positive et significative avec l'âge des utilisateurs, le temps écoulé depuis l'inscription sur la plateforme, l'intensité de la consommation musicale, la fréquence d'écoute des radios et le nombre de dispositifs de recommandation utilisés, et (ii) une corrélation négative et significative avec la fréquence d'écoute des nouveautés et des artistes stars. Autrement dit, la consommation la plus diversifiée est le fait d'utilisateurs plus intensifs, plus anciens dans leur usage de la plateforme, plus enclins à utiliser les outils de recommandation ; il s'agit d'utilisateurs plus âgés. Ils s'opposent tendanciellement à des utilisateurs plus jeunes, écoutant des artistes connus sur leur mobile. Ces résultats préliminaires devront être affinés.

### 3 / Stars vs Longue Traîne

Sur la plateforme, l'audience des différents artistes est très hétérogène : la moitié des titres et des artistes ne sont écoutés que par un seul individu de l'échantillon, alors que le top 1% des titres et des artistes est écouté par plusieurs centaines d'entre eux ; l'artiste le plus écouté l'a été par les trois quarts de l'échantillon. 10% des artistes concentrent 90% des écoutes, ce qui dessine une courbe de longue traîne tout à fait classique. Pour approfondir le phénomène, nous divisons notre population d'artistes en trois groupes :

- les stars : les 0.1% les plus écoutés, soit 116 artistes écoutés chacun par au moins 25% de l'échantillon ;
- la mid-tail, des 0,1% aux 10% d'artistes les plus écoutés, divisée en deux sous-catégories :
  - Haute: les 1% d'artistes les plus écoutés, moins les stars, soit 1 038 artistes écoutés chacun par au moins 5% de l'échantillon;

<sup>4</sup> Observatoire de la musique, « Indicateurs de la diversité musicale dans le paysage radiophonique. Rapport 2015 », disponible sur <https://philharmoniedeparis.fr/fr/ressources-numeriques/observatoire-de-la-musique/etudes/la-diversite-musicale-dans-le-paysage>, consulté le 10/11/17.

- Basse: des 1% au 10% d’artistes les plus écoutés, soit 10 363 artistes écoutés par au moins 18 usagers différents, soit un peu moins de 0,5% de l’échantillon;
- la long-tail : les 90% d’artistes les moins écoutés, qui ont pu n’être écoutés qu’une seule fois par un seul usager.

Sur la base de cette catégorisation, il est possible d’analyser le panier moyen du consommateur. Le premier constat est que, au sein de l’échantillon, presque tous les individus consomment des artistes *stars* ; en outre, ces artistes constituent une part importante de leur consommation. Le panier « moyen » d’un consommateur surreprésente les artistes *stars* (17% contre 0,1% du catalogue) et de popularité moyenne (31% pour 0,9% du catalogue pour la higher-mid-tail, 35% pour 9% du catalogue pour la lower-mid-tail). Les artistes de niche ne représentent, eux, que 17% de l’ensemble des artistes écoutés par un utilisateur. Ces écarts sont encore plus importants si l’on considère la fréquence d’écoute plutôt que le panier. Les artistes *stars* représentent alors 29% des écoutes des auditeurs ; un utilisateur sur quatre n’écoute quasiment que des artistes *stars* ou de la higher-mid-tail – plus de 80% des écoutes – alors qu’il est rare que des utilisateurs en écoutent peu – seuls 4% des usagers consacrent moins de 20% des écoutes aux 1% d’artistes les plus écoutés. Au contraire, il est plus fréquent de ne consommer pas ou peu d’artistes de niche (53% des utilisateurs consacrent moins de 5% de leurs écoutes aux 90% d’artistes les moins écoutés), mais il est rare d’en écouter dans plus de la moitié de ses écoutes (2% des utilisateurs).

Autrement dit, tout le monde écoute les stars (et c’est d’ailleurs ce qui les définit). La différence entre les utilisateurs intensifs et les autres réside dans le fait que les stars dominent plus le panier des utilisateurs dont la consommation musicale est faible (inférieure au premier quartile) alors que ceux dont la consommation est forte (supérieure au troisième quartile) font une part un peu plus importante aux artistes de la lower-mid-tail. Ces écarts, cependant, concernent surtout le panier et pas la fréquence moyenne d’écoute : les utilisateurs intensifs ont des pratiques d’exploration ponctuellement plus diversifiée, mais réécoutent les mêmes artistes que les autres.

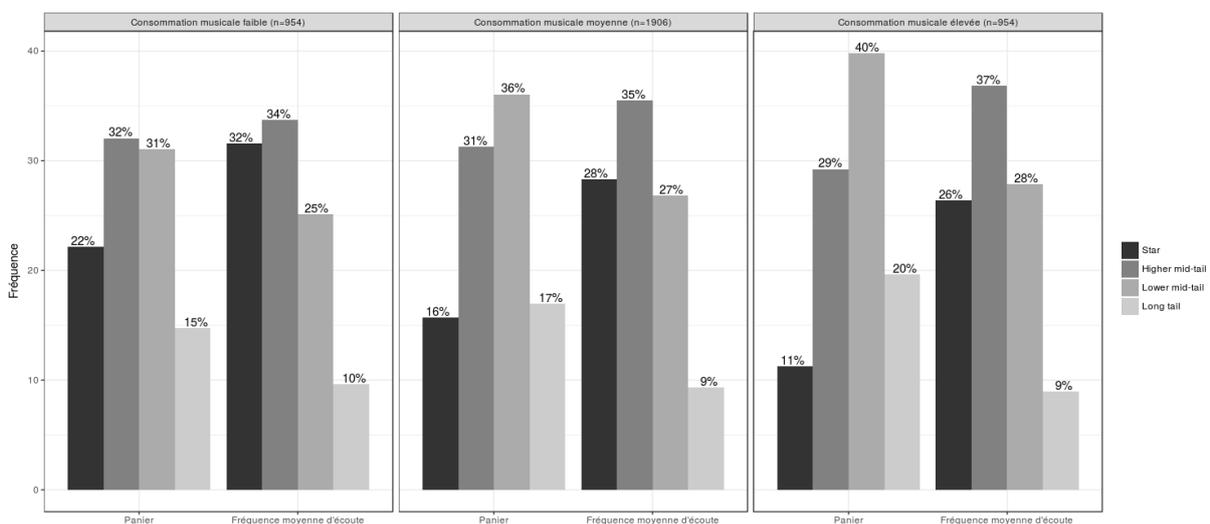


Figure 1 : Consommation et fréquence d’écoute d’artistes stars, mid-tail et tail en fonction de l’intensité d’usage

### 3 / Les usages du guidage

Dans quelle mesure les auditeurs se laissent-ils guider dans leur écoute, et comment ? Sur quelles formes de recommandations s’appuient-ils ? L’originalité de nos données est de pouvoir observer l’usage des différents dispositifs offerts par la plateforme à travers une variable de *contexte*, qui

représente la page à partir de laquelle le titre a été écouté. Si l'écoute d'un titre se fait sur la page d'une playlist éditoriale, il est raisonnable de penser que l'utilisateur suit les recommandations de cette playlist fournie par la plateforme ; si au contraire l'écoute a lieu au sein d'une page d'album mis en favori par l'utilisateur, on peut estimer qu'il s'agit d'une écoute volontaire d'un artiste déjà connu de l'auditeur, de sa discothèque virtuelle personnelle, à la manière d'un disque posé sur la platine. Pour pouvoir interpréter ainsi les différentes modalités du contexte, nous avons créé différentes catégories qui agrègent les dispositifs proches ou les modalités redondantes : recommandations algorithmiques personnalisées, classements, recommandations sociales, sélections éditoriales, outils contextuels, et radios.

Une limite à l'utilisation de la variable *contexte* est néanmoins que la moitié des écoutes (52%) ne peut pas être reliée à un dispositif particulier : que celui-ci ne soit pas identifié par l'API à cause d'un problème technique (« inconnu ») ou qu'il soit caché derrière la navigation de l'utilisateur dont nous n'observons que le résultat final (« indéterminé »), c'est-à-dire l'écoute d'un titre à partir de la page d'un album ou d'une playlist sans que nous puissions déterminer par quel chemin l'utilisateur est parvenu sur cette page (par une recommandation de la plateforme, une exploration, ou une navigation « au hasard »). Nous travaillerons donc, dans la suite, sur les 50% d'écoutes qualifiées, pour lesquelles la variable est renseignée.

### 1 / *Écoutes autonomes et hétéronomes*

Le premier résultat est que, au sein des écoutes qualifiées (8,8 millions), l'écoute « de stock » - par laquelle l'auditeur va chercher quelque chose qu'il connaît et a stocké - est nettement dominante : elle représente 5,1 millions d'écoutes, soit 59% des écoutes qualifiées. Il s'agit pour l'essentiel de l'écoute d'artistes placés en favoris par l'utilisateur ; plus marginalement, de playlist personnelles et d'albums archivés. Symétriquement, les écoutes « guidées » d'une manière ou d'une autre représentent 42% des écoutes qualifiées, soit 3,7 millions. Notons qu'elles ne concernent pas pour autant nécessairement des artistes inconnus de l'utilisateur : les recommandations qui partent des goûts établis peuvent y ramener.

Pour autant, l'ensemble des écoutes guidées ne correspondent pas à la figure d'un consommateur hétéronome pris en charge par les algorithmes. Elles se répartissent de la façon suivante.

Dispositifs	Nb. écoutes	Fréquence	Fréquence hors-stock	Nb. Usagers distincts	écoutes moyennes par usager
<b>Recommandations personnalisées</b>					
Playlist	541 612	6,1%	14,9%	2 442	222
Album	264 872	3%	7,3%	2 691	98
Titre	49 904	0,6%	1,4%	2 655	19
Radio éditoriale	656 651	7,4%	18%	1 970	333
Radio algorithmique	24 694	0,3%	0,7%	1 192	21
<b>Recommandations sociales</b>	1 154	0,01%	0,02%	190	6
<b>Recommandations éditoriales</b>	9 141	0,1%	0,3%	466	20

<b>Classements</b>	270 641	3,1%	7,4%	1 693	160
<b>Nouvelles sorties</b>	16 670	0,2%	0,5%	440	38
<b>Radios</b>					
Radio éditoriale	128 443	1,5%	3,5%	1 469	87
Radio "flow"	107 952	1,2%	3%	763	142
Radio algorithmique	97 194	1,1%	2,7%	1 031	94
<b>Outils contextuels</b>					
Discographie de l'artiste	502 611	5,7%	13,8%	1 908	263
Top de l'artiste	536 030	6,1%	14,7%	3 526	152
<b>Moteur de recherche</b>	418 133	4,7%	11,5%	3 030	138
<b>Player (par défaut)</b>	16 647	0,2%	0,5%	1 228	14
<b>Stock</b>	5 172 732	58,7%	N/A	3 535	1 463

Table 4: Distribution des écoutes qualifiées entre les différents dispositifs

Au sein des dispositifs de guidage, une partie renforcent plutôt la figure d'un utilisateur autonome, explorant activement et consciemment le catalogue : c'est le cas du moteur de recherche et des outils contextuels tels que la discographie des artistes (qui permet de naviguer l'ensemble de l'œuvre d'un artiste) ou le top des artistes (qui propose les titres les plus écoutés de cet artiste). Pour ces trois dispositifs, il est raisonnable de penser que l'utilisateur les mobilise volontairement dans l'intention d'explorer un univers musical. Ces dispositifs de l'exploration autonome comptent pour 1,5 millions d'écoutes, soit 17% des écoutes qualifiées. Ce sont donc, avec les écoutes de stock, 76% des écoutes qui relèvent de la figure de l'auditeur autonome.

Restent deux formes d'écoutes guidées où l'agentivité des dispositifs est plus marquée. L'une relève de la recommandation proprement dite, de la suggestion : il s'agit des propositions algorithmiquement construites de titres, d'artistes, de playlists, qui jouent un rôle dans 8% des écoutes. Selon l'interprétation que l'on fait du clic de l'auditeur sur la recommandation, ces dispositifs peuvent aussi bien être lus comme des formes de recommandation pertinentes suivies volontairement par un consommateur averti, que comme des outils de manipulation d'un consommateur mal informé. L'autre regroupe les écoutes de flux, radios d'un artiste et radios algorithmiques : dans cette configuration, l'essentiel de l'agentivité revient à la plateforme, qui sélectionne les titres et organise leur enchaînement, sans aucune intervention de l'utilisateur. L'écoute algorithmique représente ainsi un peu moins d'un quart des écoutes, dont une moitié d'écoute de flux.

## 2 – Guidage et découverte

Les différentes formes d'exploration et de guidage permettent-elles de renforcer la diversité de la consommation (ce qui est une de leurs promesses), ou conduisent-elles au contraire à renforcer l'uniformité de l'écoute musicale ? Nos données ne permettent pas de répondre clairement à cette question : notamment, la durée d'observation de cinq mois n'est pas suffisante pour identifier avec certitude des découvertes, et nous n'avons aucune information sur les autres formes d'écoute et d'information musicale mobilisées par les utilisateurs en dehors de la plateforme. A titre exploratoire, nous avons cependant repéré, au sein des écoutes guidées, celles qui relevaient d'une

« première écoute » pour l'utilisateur au sein de notre période d'observation, en excluant le premier mois. On obtient ainsi un ensemble de 639 319 écoutes, que l'on peut considérer sinon comme des découvertes, en tous cas comme des morceaux amenés ou ramenés à la conscience de l'utilisateur via une modalité d'exploration sur le site. Nous pouvons voir si les morceaux ainsi (re)-découverts via les dispositifs sont plutôt le fait d'artistes stars, mid-tail ou longue traîne.

Dispositifs	Effectif	Fréquence	Long tail	Lower mid-tail	Higher mid-tail	Star	Total
<b>Recommandations personnalisées</b>							
Playlist	190 730	29,8%	29,9%	36,2%	23,7%	10,3%	100 %
Album	40 842	6,4%	33,6%	46,4%	<u>15,2%</u>	<u>4,8%</u>	100 %
Titre	18 086	2,8%	33,9%	39,1%	18,8%	8,2%	100 %
Radio éditoriale	53 301	8,3%	30,9%	42,8%	19,9%	6,4%	100 %
Radio algorithmique	15 358	2,4%	<b>75,6%</b>	<u>21,5%</u>	<b>2,6%</b>	<b>0,4%</b>	100 %
<b>Recommandations sociales</b>	400	0,1%	37,0%	<u>27,0%</u>	19,0%	17,0%	100 %
<b>Recommandations éditoriales</b>	3 707	0,6%	15,2%	40,6%	17,5%	<b>26,7%</b>	100 %
<b>Classements</b>	39 409	6,2%	<u>2,1%</u>	<b>8,2%</b>	<u>14,3%</u>	<b>75,3%</b>	100 %
<b>Nouvelles sorties</b>	6 110	1,0%	<b>42,1%</b>	37,1%	<u>16,2%</u>	<u>4,5%</u>	100 %
<b>Radios</b>							
Radio éditoriale	41 502	6,5%	32,3%	53,1%	<b>11,4%</b>	<u>3,1%</u>	100 %
Radio "flow"	11 619	1,8%	38,8%	38,2%	16,8%	6,2%	100 %
Radio algorithmique	30 636	6,2%	<b>68,7%</b>	<u>24,3%</u>	<b>5,6%</b>	<b>1,4%</b>	100 %
<b>Outils contextuels</b>							
Discographie de l'artiste	20 469	3,2%	<b>50,6%</b>	33,3%	<u>12,3%</u>	<u>3,8%</u>	100 %
Top de l'artiste	78 874	12,3%	23,6%	<u>29,3%</u>	30,6%	16,5%	100 %
<b>Moteur de recherche</b>	80 163	12,5%	<b>42,1%</b>	<u>30,3%</u>	20,6%	7,0%	100 %
<b>Player (par défaut)</b>	8 113	1,3%	<b>0,1%</b>	<b>14,0%</b>	<b>54,1%</b>	<b>31,8%</b>	100 %
<b>Ensemble premières écoutes</b>			21,1%	42,5%	26,1%	10,4%	100 %
<b>Ensemble écoutes qualifiées</b>			9,4%	22,8%	34,6%	27,2%	100 %

Tableau 5 – Dispositifs d'exploration et longue traîne

Le premier constat est que, en dehors des dispositifs de classements du type « les plus écoutés », la plupart des systèmes de recommandation orientent tendanciellement les auditeurs des titres de la mid-tail et de la longue traîne ; les premières écoutes en général sont moins concentrées sur les artistes les plus écoutés (seul 10,4% de ces écoutes concernent des artistes stars, contre 27,2% du total des écoutes qualifiées).

Mis à part les classements, seules les recommandations éditoriales de la plateforme et la chanson présentée à l'utilisateur au lancement du service orientent plus que la moyenne vers les artistes star ou de la higher mid-tail. L'orientation vers la longue traîne, cependant, varie beaucoup d'un dispositif à l'autre. Elle est la plus forte avec les recommandations algorithmiques, puis dans un second temps avec l'exploration autonome de la discographie des artistes et des nouvelles sorties. Contrairement à ce que l'on pourrait penser, dans les données agrégées, les dispositifs d'exploration autonome ne sont pas les plus « pointus » musicalement, même s'ils proposent une ouverture plus large que les recommandations éditoriales et les classements.

La durabilité de ces “découvertes” est néanmoins très variable. On peut mesurer cette durabilité au nombre de réécoutes des artistes ainsi découverts. Or, ce sont les artistes découverts via leur page qui sont le plus susceptibles d'être réécoutes (18 et 15 réécoutes en moyenne sur la période, contre 8 pour l'ensemble des dispositifs), puis ceux découverts via les recommandations éditoriales de Deezer (14 réécoutes en moyenne) ou via des recommandations du réseau de l'utilisateur (13 réécoutes en moyenne). En somme, les formes d'exploration autonome<sup>5</sup> et les recommandations humaines, par opposition aux recommandations algorithmiques, semblent produire des relations plus durables avec les musiques découvertes. On note enfin que les recommandations humaines sont légèrement plus durables que l'exploration autonome pour les stars, mais le sont nettement moins pour les artistes de la mid-tail, où l'exploration est plus importante.

#### **4 / Discussion : des profils ou des contextes ?**

Notre étude, encore exploratoire – les analyses de la base de données ne sont pas terminées – permet d'établir un certain nombre de résultats. Par rapport à l'inquiétude sur l'émergence d'un auditeur passif et hétéronome, nous montrons que, dans les usages réels, les écoutes autonomes dominent : elles représentent les trois quart des écoutes. Autrement dit, si les plateformes de streaming dessinent en effet la figure d'un amateur complètement pris en charge par les machines, cette figure est loin d'être habitée par les utilisateurs réels, qui choisissent la majorité de leurs écoutes, et usent préférentiellement de dispositifs autonomes d'exploration tels que les moteurs de recherche et les discographies d'artistes.

Le motif de l'écoute guidée représente un quart des usages mesurés, répartis équitablement entre écoute de flux et suivi ponctuel de recommandations algorithmiques. Dans tous ces cas, il y a effectivement délégation du choix à un dispositif technique : cette délégation est proche de celle du choix d'une station de radio dans le cas du flux, plus ponctuelle dans le cas d'un suivi de lien de recommandation. Ces outils tiennent leurs promesses d'exploration de la richesse des catalogues et d'enrichissement de la diversité consommée des utilisateurs, en les guidant préférentiellement vers des titres de longue traîne ou de notoriété intermédiaire.

Il reste qu'il est difficile de qualifier systématiquement d'hétéronomes et de passives ces écoutes où l'agentivité algorithmique est plus marquée. Le suivi de lien de recommandation (« vous avez aimé tel artiste, vous devriez aimer tel autre ») peut aussi bien être vu comme une pratique de novice (comblant son incompetence par la délégation à l'algorithme) que comme un démarche d'amateur éclairé (utilisant ponctuellement un outil informatique pour compléter sa palette exploratoire), comme le montre une récente étude des gros utilisateurs de recommandation algorithmique qui met

---

<sup>5</sup> Il faut toutefois préciser que les dispositifs page de l'artiste sont d'interprétation encore plus difficile que les autres dispositifs, dans la mesure où l'utilisateur a pu arriver sur cette page de nombreuses manières différentes.

en évidence leur réflexivité (Karakayali 2017). Même dans le cas des flux, ceux-ci sont choisis, peuvent être interrompus, transformés par l'évaluation active des morceaux. Autrement dit, si le constat de l'usage de ces dispositifs rend crédible l'émergence d'une figure de l'écoute hétéronome, elle demande à être vérifiée et affinée par d'autres méthodes.

En particulier, une des questions qui restent ouvertes est de savoir si le recours à l'écoute guidée relève plutôt des dispositions ou des situations. Existe-t-il des individus plus ou moins autonomes dans leur écoutes musicales, plus ou susceptibles de se laisser guider ? Ou, au contraire, tous les individus ont-ils ponctuellement recours au guidage en fonction du contexte attentionnel dans lequel ils se trouvent ? Cette question pourra faire l'objet de traitements complémentaires à partir de nos données. Pour l'instant, la recherche d'associations entre les propriétés des individus et le recours au guidage algorithmique n'est guère convaincante ; la seule relation un peu tangible établit que les systèmes de guidage sont plutôt plus utilisés – de façon non exclusive - par les usagers intensifs.

Les travaux qualitatifs sont alors cruciaux pour comprendre dans quels contextes, avec quel conscience, quelle compréhension, quelle forme d'attention, quel objectif, les utilisateurs acceptent de s'en remettre aux algorithmes. Le travail exploratoire de Louis Melançon (2016) auprès de dix utilisateurs permet ainsi d'esquisser à la fois une variété des profils, entre les utilisateurs refusant de voir les recommandations (« je sais déjà ce que je veux écouter »), ceux qui les soupçonnent de relever de la publicité, et ceux qui acceptent certaines des suggestions. Il souligne aussi l'importance des contextes dans le recours aux recommandations : certaines sont acceptables dans un moment d'exploration, d'autres dans un contexte d'écoute distraite, pour être immédiatement débranchée si l'attention se refocalise. Ces travaux empiriques à prolonger, tout comme notre enquête quantitative, nous éloignent dans tous les cas de la figure d'un utilisateur gouverné par les algorithmes, nous immunisent contre la croyance au premier degré dans les discours techno-déterministes tenus par les entrepreneurs des plateformes comme par leurs critiques ; ils nous orientent au contraire vers la compréhension des cultures algorithmiques émergentes, au sein desquelles les auditeurs, selon leurs compétences et selon les contextes, vont refuser ou brancher ponctuellement les algorithmes, sans pour autant abandonner le contrôle de leur expérience musicale.

## Références

- Beuscart, Jean-Samuel, Kevin Mellet, Département des Etudes de la Prospective et des Statistiques, and Ministère de la Culture et de la Communication. *Promouvoir les oeuvres culturelles - Usages et efficacité de la publicité dans les filières culturelles*. Paris: La Documentation française, 2012.
- Beuscart, Jean-Samuel, et Kevin Mellet. « Shaping Consumers' online voices : algorithmic apparatus or evaluation culture ? », in Seyfert, Robert, et Roberge, Jonathan (dir.), *Algorithmic Cultures*, Routledge, 2016.
- Bourreau, Marc, François Moreau, et Pierre Senellart. «La diversité culturelle dans l'industrie de la musique enregistrée en France (2003-2008)». *Cultures études*, no. 5, 2011.
- Chevalier, Judith A., and Dina Mayzlin. “The Effect of Word of Mouth on Sales: Online Book Reviews.” *Journal of Marketing Research* 43, no. 3 (2006): 345–354.
- Galloway, Alexander R. *The Interface Effect*. Cambridge, UK ; Malden, MA: Polity Press, 2012.
- Godes, David, and Dina Mayzlin. “Using Online Conversations to Study Word-of-Mouth Communication.” *Marketing Science*, November 1, 2004. doi:10.1287/mksc.1040.0071.
- Hagen, Anja Nylund. “The Metaphors We Stream by: Making Sense of Music Streaming.” *First Monday* 21, no.3 (March 3, 2016).

- Hennion, Antoine, Maisonneuve, Sophie, Gomart, Emilie, *Figures de l'amateur. Formes, objets, pratiques de l'amour de la musique aujourd'hui*, Paris, La Documentation Française, 200
- Karakayali, Nedim, Burc Kostem et Idil Galip, "Recommendation systems as technologies of the self. Algorithmic control and the formation of music taste", *Theory, Culture and Society* (2017).
- Karpik, Lucien. "L'économie de La Qualité." *Revue Française de Sociologie* 30, no. 2 (April 1989).
- Katz, Elihu, Paul Lazarsfeld, Eric Maigret, and Daniel Céfai. *Influence personnelle : Ce que les gens font des médias*. Paris; Paris: Armand Colin, 2008.
- Lash, Scott. "Power after Hegemony: Cultural Studies in Mutation?" *Theory, Culture & Society* 24, no. 3 (May 2007): 55–78.
- Maillard, Sisley. "Information et Consommation des biens culturels à l'ère numérique", Thèse d'économie, Telecom-ParisTech, 2015.
- Melançon, Louis. « La Nubémorphose des pratiques de consommation musicale : le cas des plateformes de diffusion musicale en continu », rapport de recherche, 2015.
- Morris, Jeremy. "Sounds in the Cloud: Cloud Computing and the Digital Music Commodity." *First Monday* 16, no. 5 (April 10, 2011).
- Pasquale, Frank. *The Black Box Society – The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press, 2015.
- Seyfert, Robert, and Jonathan Roberge. *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*. London ; New York: Routledge, 2016.
- Sharma, Amit, Jake M. Hofman, and Duncan J. Watts. "Estimating the Causal Impact of Recommendation Systems from Observational Data," 453–70. ACM Press, 2015.
- Sterne, Jonathan. "The mp3 as Cultural Artifact." *New Media & Society* 8, no. 5 (October 1, 2006): 825–42.
- Weiss Hanrahan, Nancy. "La découverte musicale en régime numérique", in Le Guern (dir.), *Où va la musique ?*, Presse des Mines, 2016.