

# Une approche lexicométrique du rap français

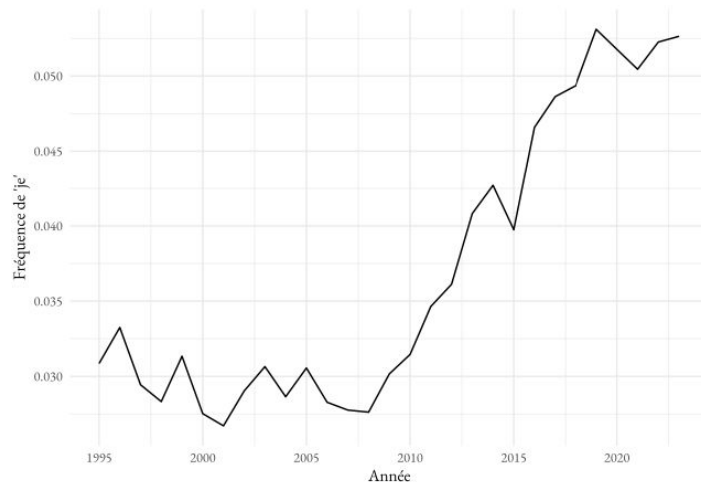
Juliette Hubert (Université Laval & Université Polytechnique Hauts-de-France)

featuring

Benoît de Courson (École Normale Supérieure-PSL & Leiden University)

# Introduction

- Au commencement était Gallicagram
- “Ma qualité de lyrics est quantifiable”  
(Médine, 2018)
- Juliette, une alliée de taille



## Juin :

- \* 01/06 : Ärsenik - *Quelques gouttes suffisent...*
- \* 03/06 : Ad'Hoc-1 - *À Bord Du N-333*
- 05/06 : Kheops - *Sad Hill* (réédition 1CD)
- \* 15/06 : Dee Nasty - *Le diamant est éternel*
- \* 15/06 : Saïan Supa Crew - *Saïan Supa Land* (EP)
- \* 19/06 : Basic (Vestat & Endo) - *98 jusqu'à l'infini...*
- \* 19/06 : Mellowman - *Au jour le jour*
- \* 22/06 : Le Péril Noir - *Le destin*
- \* 23/06 : I.m.s. - *La bible du soldat*

## Juillet :

- 01/07 : Kabal - *États d'Âmes*
- 21/07 : MC Solaar - *MC Solaar*

https://genius.com/Medine-bangerang-lyrics

Bangerang - Médine ↗

Lyrics

About

Q&A

Comments

Edit Lyrics

Cette année, j'fête mes dix ans au Wakanda

Ma qualité d'lyrics est quantifiable

Ne confonds pas Paul Bocuse et l'McDonald's

Tous tes rappeurs sont là par le piston

Moi, je ne mens pas, c'est la vérité qui s'trompe

J'leur déclare la guerre toutes les sept secondes

J'commande mes hommes et mes pizzas sur le même ton

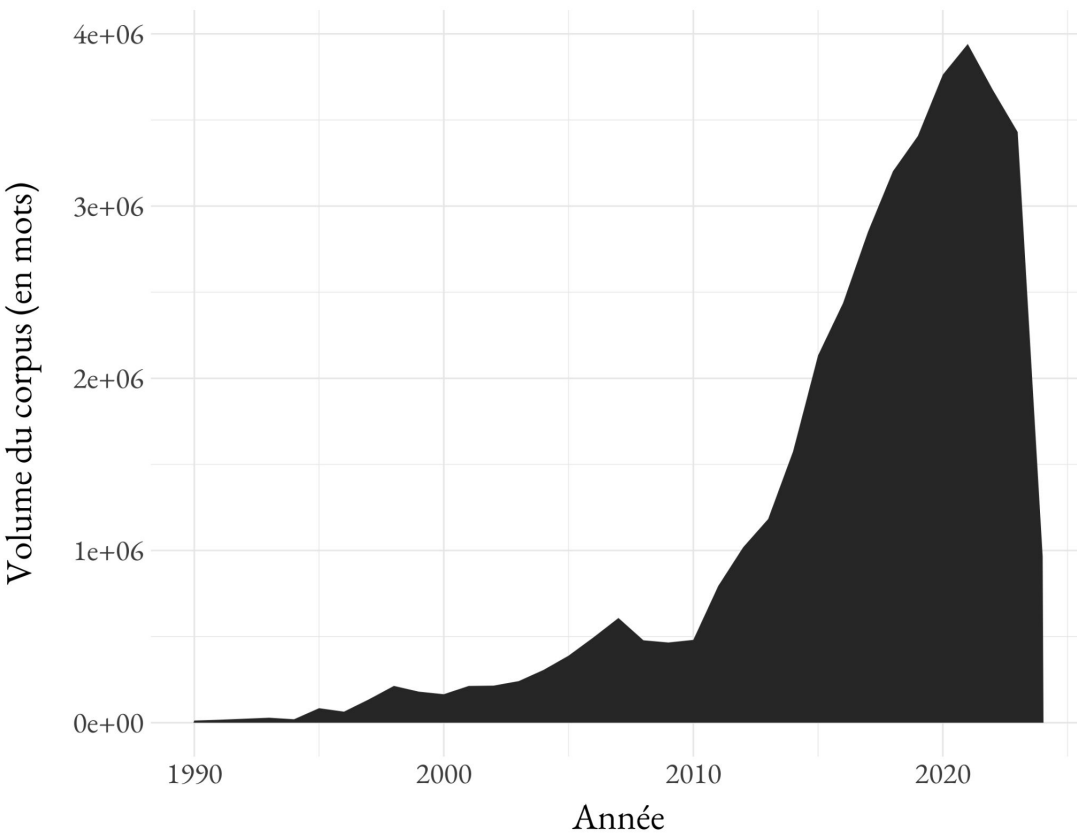
J'ai fini ma sèche, j'suis en prise de masse

J'ai fait de mes seize une musique de masse

Rap hardcore, rappeurs qui s'clashent

Même entre nous, y a des rapports de classe

# Le corpus



Remarques :

- Rap comme entité
- Corpus
- Textualité

# Plan

Que faire de ces données textuelles ? Un tutoriel à travers démarches possibles :

- Une batterie de mesures
- Une démarche *data-driven*
- Tester une hypothèse pré-existante

# I. Une approche aveugle : la “batterie de mesures”

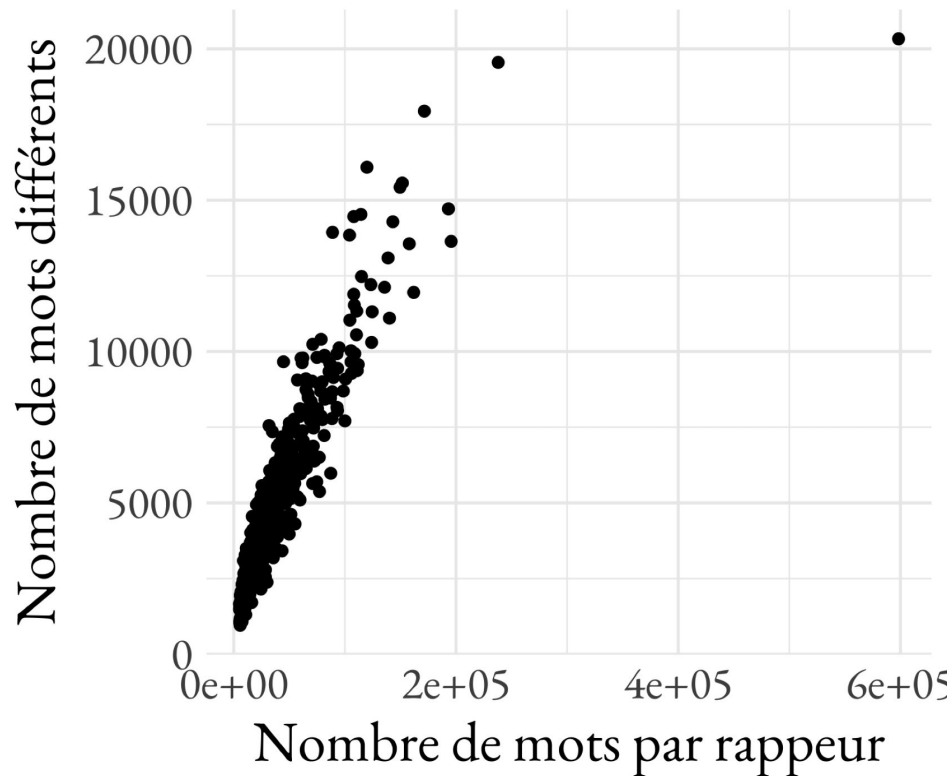
- Une ‘lecture distante’ peut appréhender un corpus de taille inhumaine
- On ne peut pas mesurer *la* qualité, mais on peut mesurer *des* qualités
- En particulier :
  - Les rappeurs sont plus ou moins « lyricistes », « obsédés textuels »
  - Les rappeurs sont plus ou moins violents verbalement
- Quel lien statistique entre les deux ? Quelle évolution ?

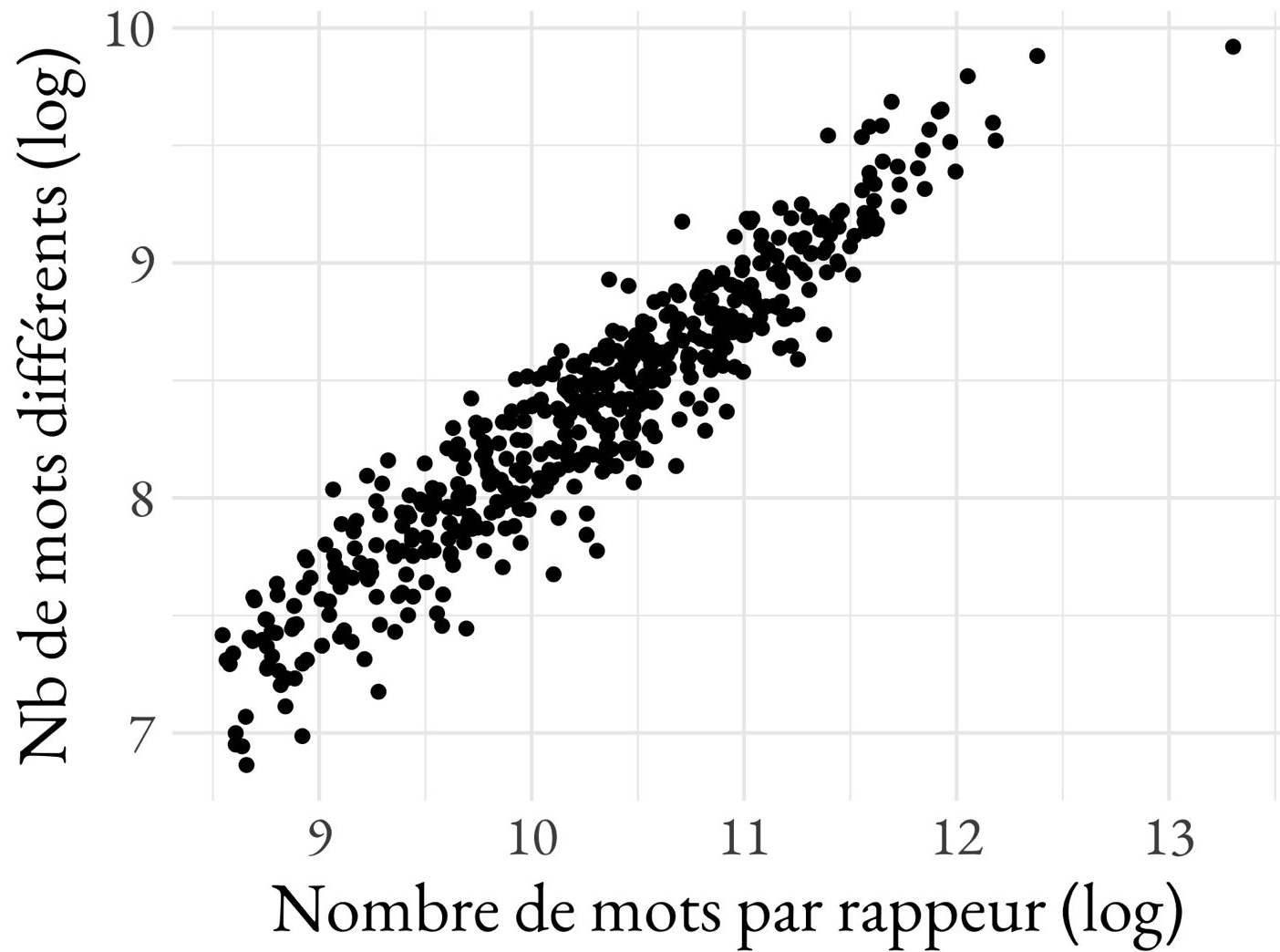
# Que mesurer ?

- Des mesures de complexité (diversité lexicale, longueur des mots)
- La fréquence de lexiques (insultes, verlan, argot, ... )
- L'analyse de sentiments (LIWC)
- Appliquer des détecteurs de haine/sexisme

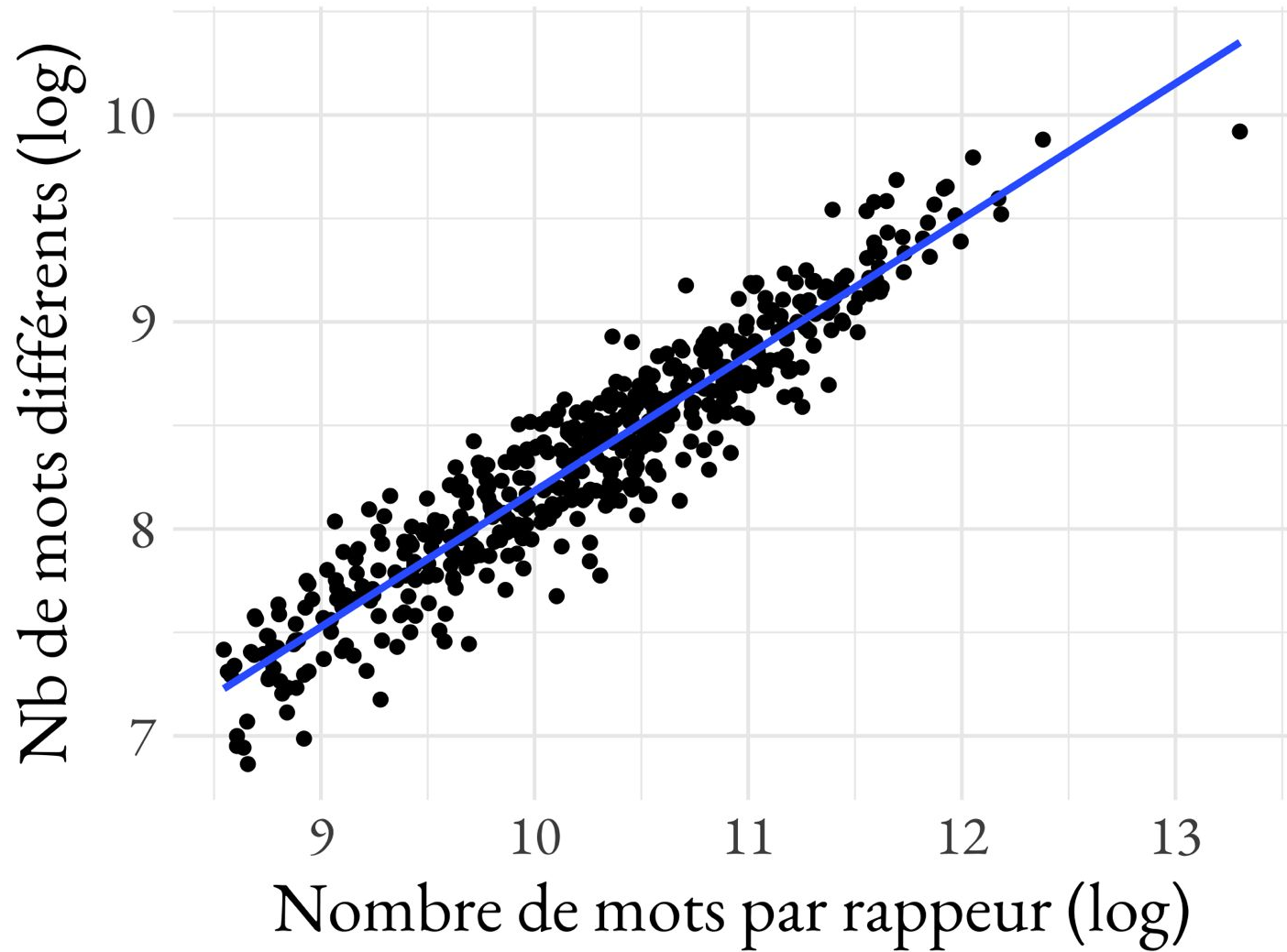
# Comment mesurer l'étendue du vocabulaire ?

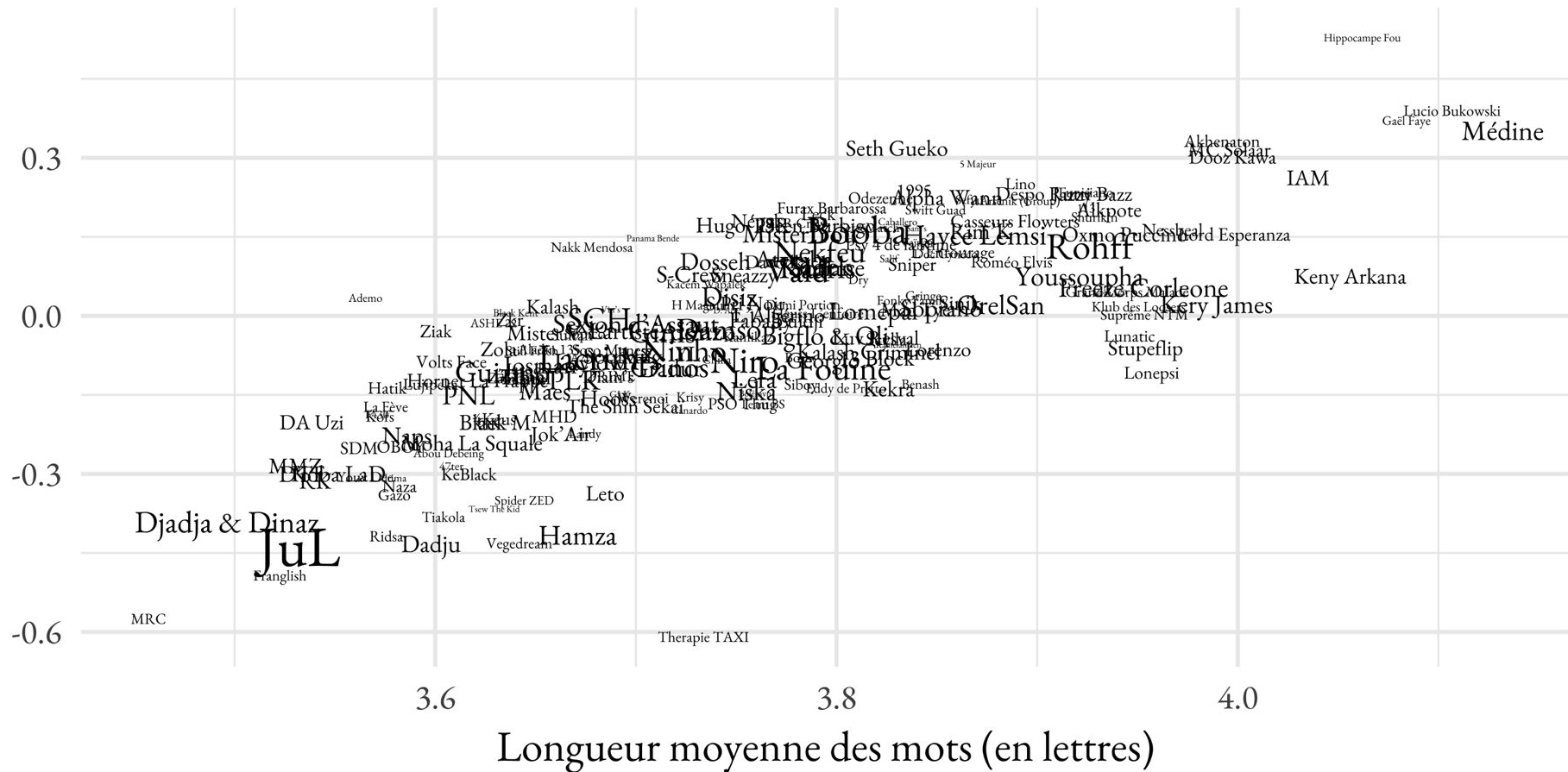
- On peut compter le nombre de mots différents
- ...mais JuL a publié près de 1000 morceaux, Tiakola seulement 24
- Et si on divisait le nombre de mots différents par le nombre de mots ?
- Le problème s'inverse











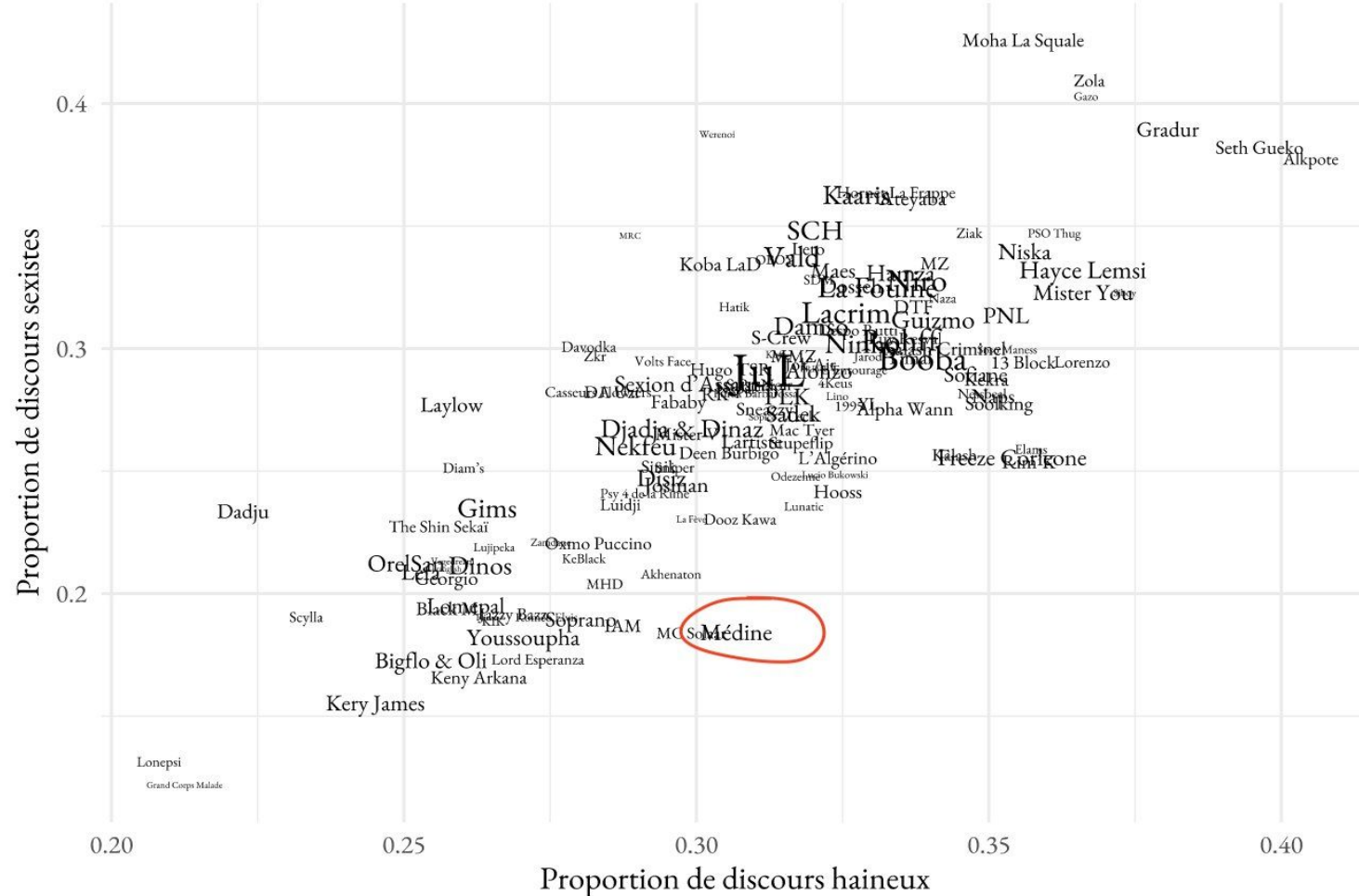
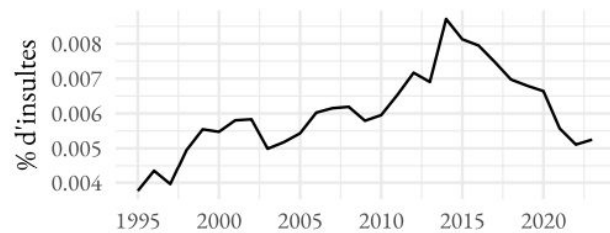
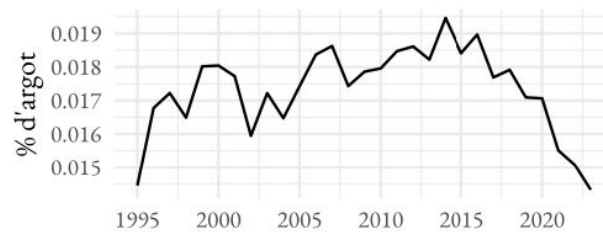
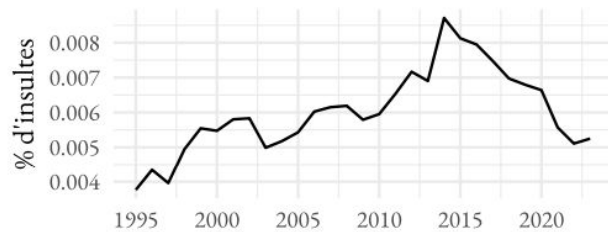


FIGURE 12 – Proportion de discours sexistes et haineux pour les rappers les plus vus

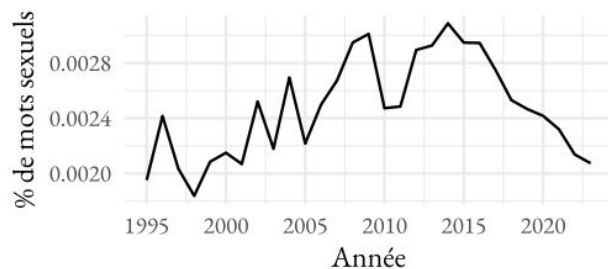
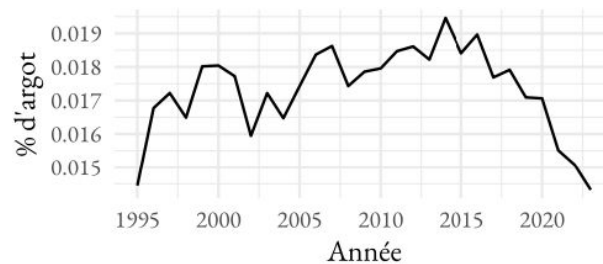
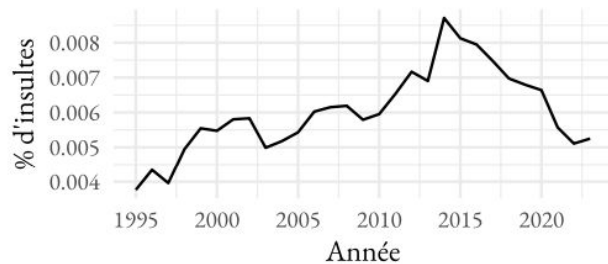
# Une évolution en deux temps



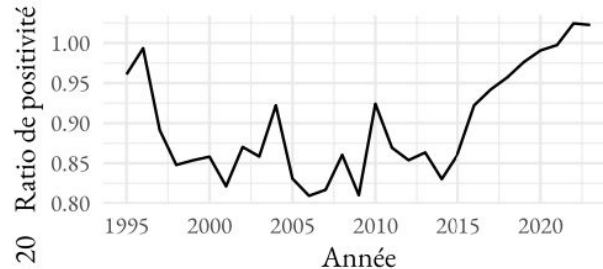
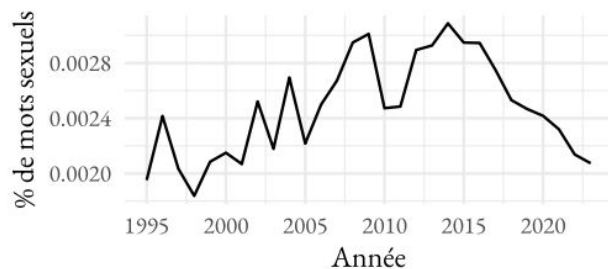
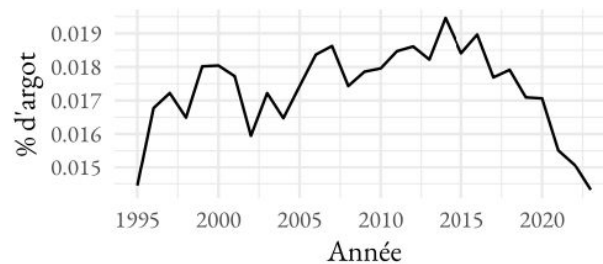
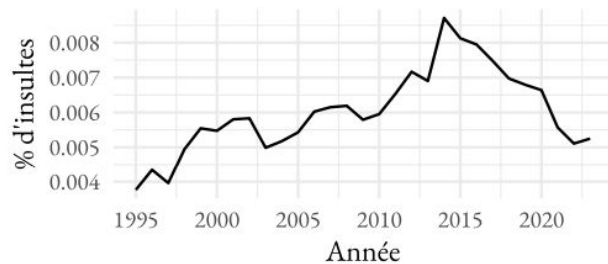
# Une évolution en deux temps



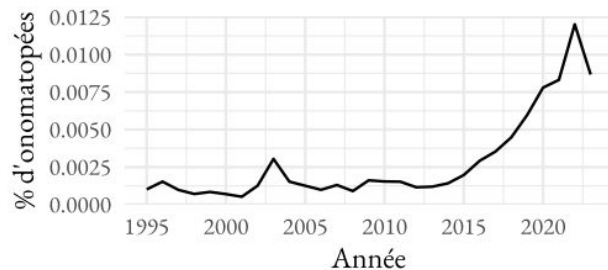
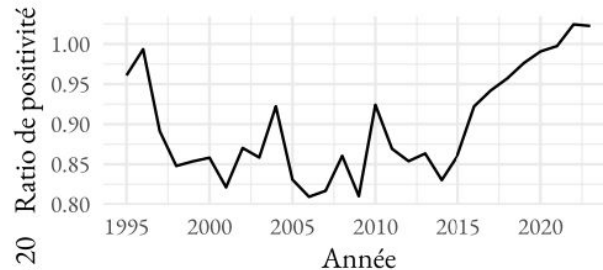
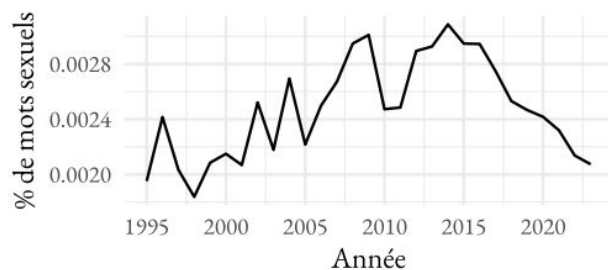
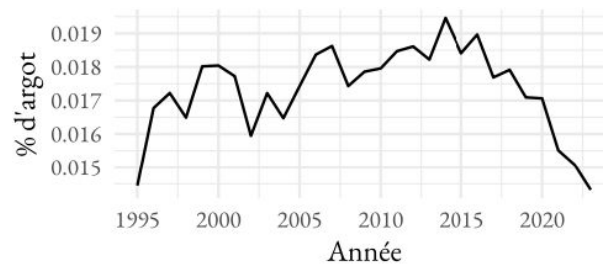
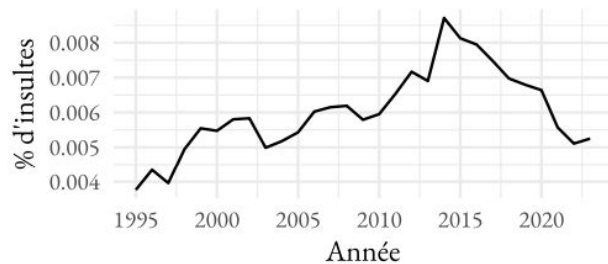
# Une évolution en deux temps



# Une évolution en deux temps

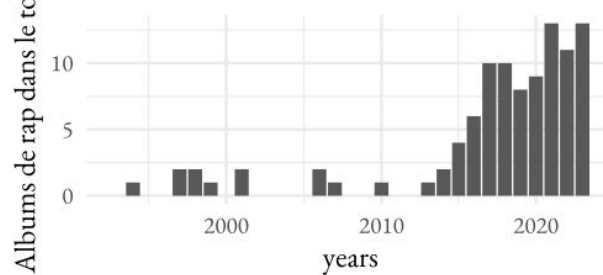
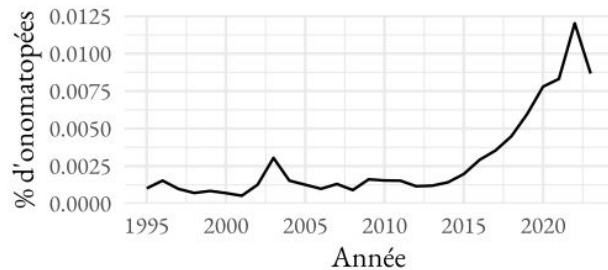
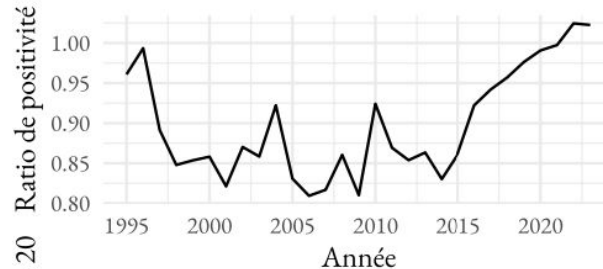
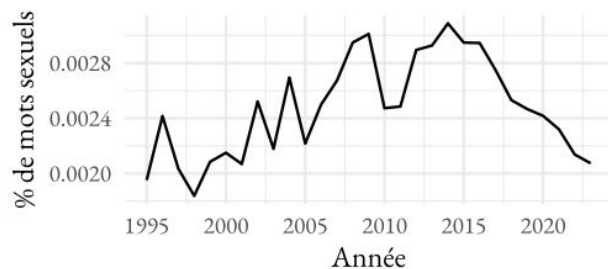
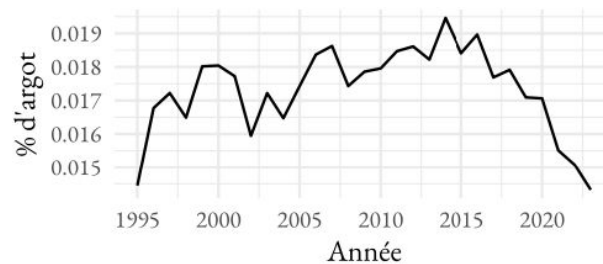
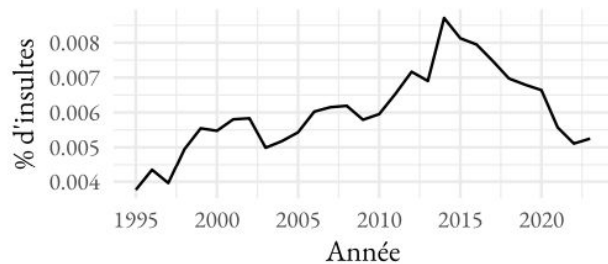


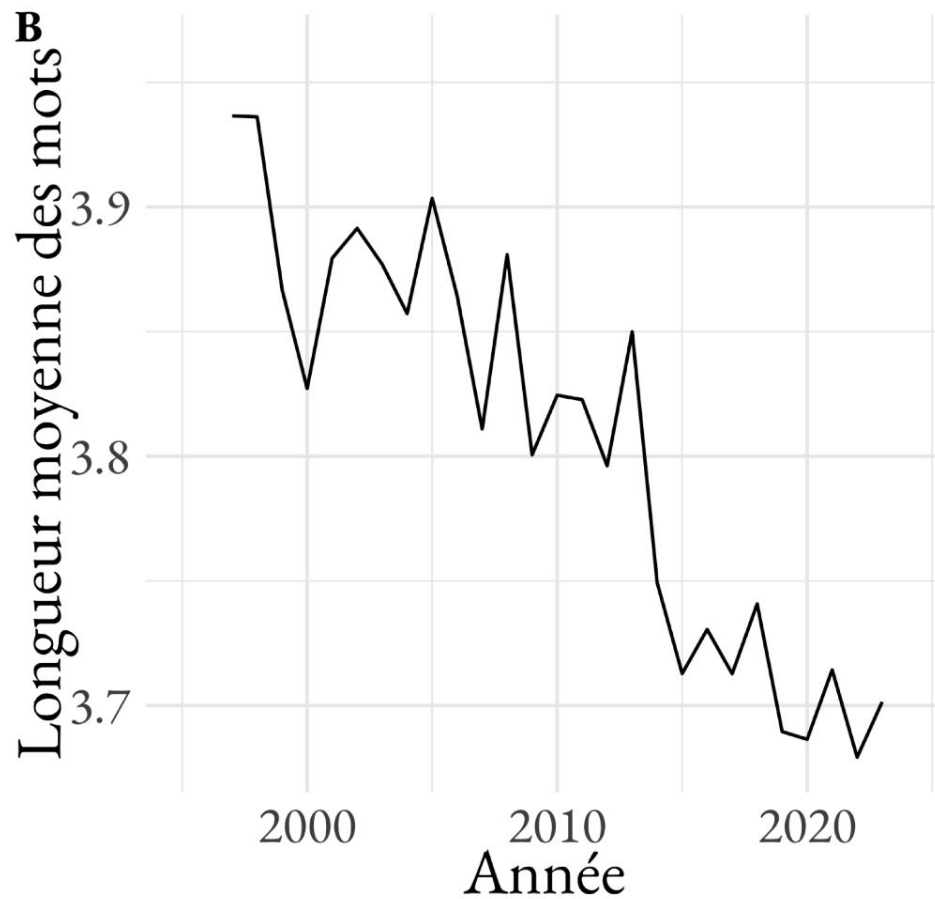
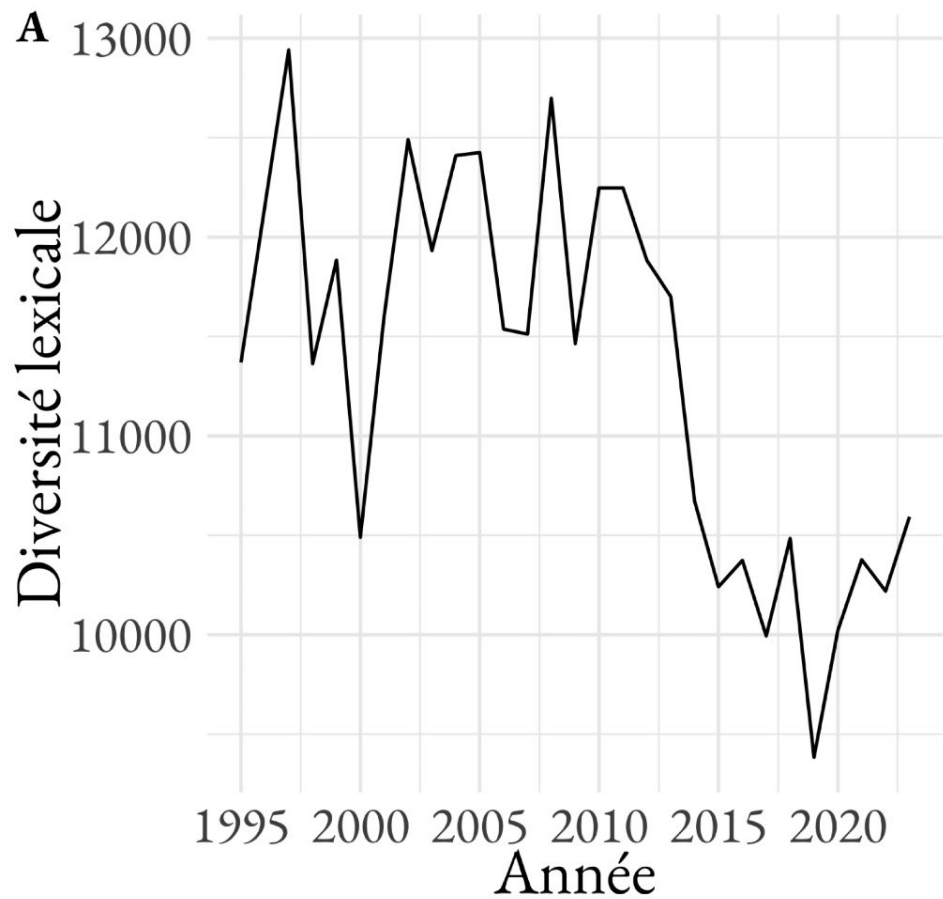
# Une évolution en deux temps

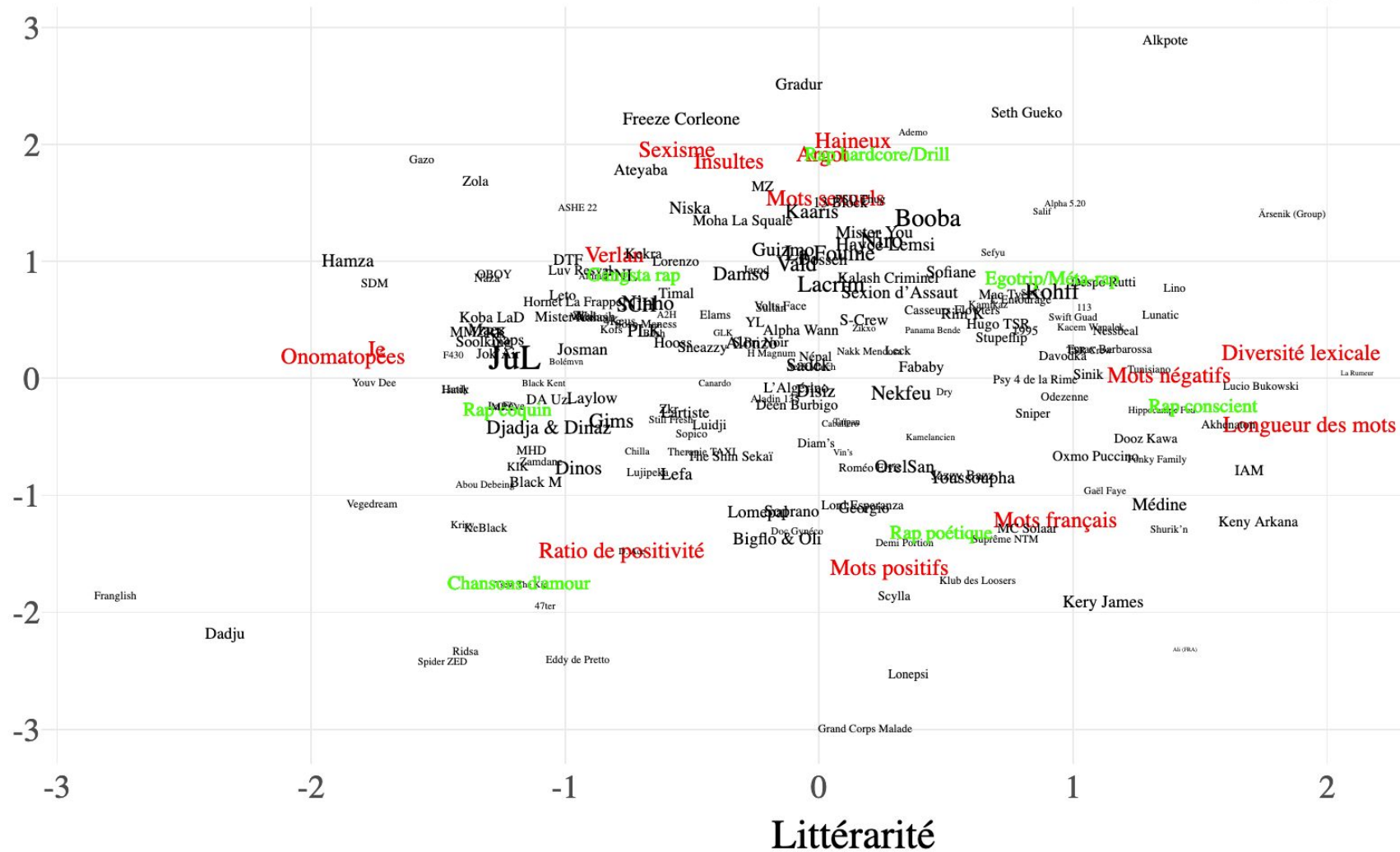




# Une évolution en deux temps







Où en est-on ? A-t-on objectivé le rap français ?

## II - Une approche réellement data-driven ?

- Comment mesurer sans juger ?
- Le *topic modeling* : rassembler des textes sémantiquement similaires
- Une méthode “non-supervisée”

# Le *topic modeling*

Mots caractérisant le topic	%	Nom (selon l'auteur)
beuh, sacoche, binks, billets, poto, monnaie, Audi, moula, zone, condés, euros, pétard, cli, mets, frappe	19.96	Koba LaD
rêves, ciel, espoir, bonheur, vie, larmes, lumière, souvenirs, jour, destin, âme, vide, monde, silence, soleil	16.37	Dooz Kawa
peuple, pays, système, justice, liberté, jeunes, violence, médias, familles, politique, État, droits, racisme, peuples, pauvres	15.47	Keny Arkana
rap, mic, flow, MC, rimes, beat, style, hip, micro, rime, rappeurs, phases, lyrics, textes, rappeur	14.91	Ärsenik
baby, soir, might, chérie, yeah, bébé, nuit, miss, night, soirée, copines, danse, fille, mama, belle	13.43	JuL
amour, cœur, yeux, might, larmes, sentiments, fois, nuit, fille, souvenirs, temps, vie, soir, mal, relation	10.35	Dadju
négro, négros, ekip, bitch, LDO, Fuck, nigga, pétasse, Dakar, chatte, game, bite, secte, flow, gang	8.92	Freeze Corleone

# Le *topic modeling*

Mots caractérisant le topic	%	Nom (selon l'auteur)	Artiste représentatif
beuh, sacoché, binks, billets, poté, monnaie, Audi, moulé, zone, condés, euros, pétard, cli, mets, frappe	19.96	Koba LaD	Gangsta rap
rêves, ciel, espoir, bonheur, vie, larmes, lumière, souvenirs, jour, destin, âme, vide, monde, silence, soleil	16.37	Dooz Kawa	Rap poétique
peuple, pays, système, justice, liberté, jeunes, violence, médias, familles, politique, État, droits, racisme, peuples, pauvres	15.47	Keny Arkana	Rap conscient
rap, mic, flow, MC, rimes, beat, style, hip, micro, rime, rappeurs, phases, lyrics, textes, rappeur	14.91	Ärsenik	Egotrip/Méta-rap
baby, soir, might, chérie, yeah, bébé, nuit, miss, night, soirée, copines, danse, fille, mama, belle	13.43	JuL	Rap coquin
amour, cœur, yeux, might, larmes, sentiments, fois, nuit, fille, souvenirs, temps, vie, soir, mal, relation	10.35	Dadju	Chansons d'amour
négro, négros, ekip, bitch, LDO, Fuck, nigga, pétasse, Dakar, chatte, game, bite, secte, flow, gang	8.92	Freeze Corleone	Rap hardcore/Drill

Part dans le corpus

1.00  
0.75  
0.50  
0.25  
0.00

2000

2010

2020

Année

Topic

- Chansons d'amour
- Egotrip/Méta-rap
- Gangsta rap
- Rap conscient
- Rap coquin
- Rap hardcore/Drill
- Rap poétique





# Le *topic modeling*

Mots caractérisant le topic	%	Nom (selon l'auteur)	Artiste représentatif
beuh, sacoché, binks, billets, poté, monnaie, Audi, moulé, zone, condés, euros, pétard, cli, mets, frappe	19.96	Koba LaD	Gangsta rap
rêves, ciel, espoir, bonheur, vie, larmes, lumière, souvenirs, jour, destin, âme, vide, monde, silence, soleil	16.37	Dooz Kawa	Rap poétique
peuple, pays, système, justice, liberté, jeunes, violence, médias, familles, politique, État, droits, racisme, peuples, pauvres	15.47	Keny Arkana	Rap conscient
rap, mic, flow, MC, rimes, beat, style, hip, micro, rime, rappeurs, phases, lyrics, textes, rappeur	14.91	Ärsenik	Egotrip/Méta-rap
baby, soir, might, chérie, yeah, bébé, nuit, miss, night, soirée, copines, danse, fille, mama, belle	13.43	JuL	Rap coquin
amour, cœur, yeux, might, larmes, sentiments, fois, nuit, fille, souvenirs, temps, vie, soir, mal, relation	10.35	Dadju	Chansons d'amour
négro, négros, ekip, bitch, LDO, Fuck, nigga, pétasse, Dakar, chatte, game, bite, secte, flow, gang	8.92	Freeze Corleone	Rap hardcore/Drill

# “Blanche”, Maes ft. Booba

<Couplet=2>

Je n'connais pas le banc à part celui de la G.A.V  
Drapeau est noir et blanc, je reste devant, je préfère être à l'attaque

Si j'n'ai plus d'limite, je m'arrête où ?  
J'pends mon linge sale devant personne, je rachète tout

Je ne demande qu'à t'aimer, te montrer mon Makélélé  
Le Versace, le D&G, elle est trop bonne, wallaye, c'est B

Je vais l'emmener à l'étranger, peut-être sur place, j'vais l'échanger  
Faire le hlel ou l'étrangler, j'suis qu'en léger, que la CB

J'drague à la Momo Henni, beaucoup de cash, je la séduis  
On m'tire dessus, j'entends des bruits, même en réserve, je la détruis

J'viens du 9.2, la quality, tu connais d'jà l'état d'esprit  
Elle m'a d'mandé combien ça coûte comme si j'avais regardé l'prix

<refrain>

<outro>

# “Alphonse Karr”, Freeze Corleone

[Couplet 2]

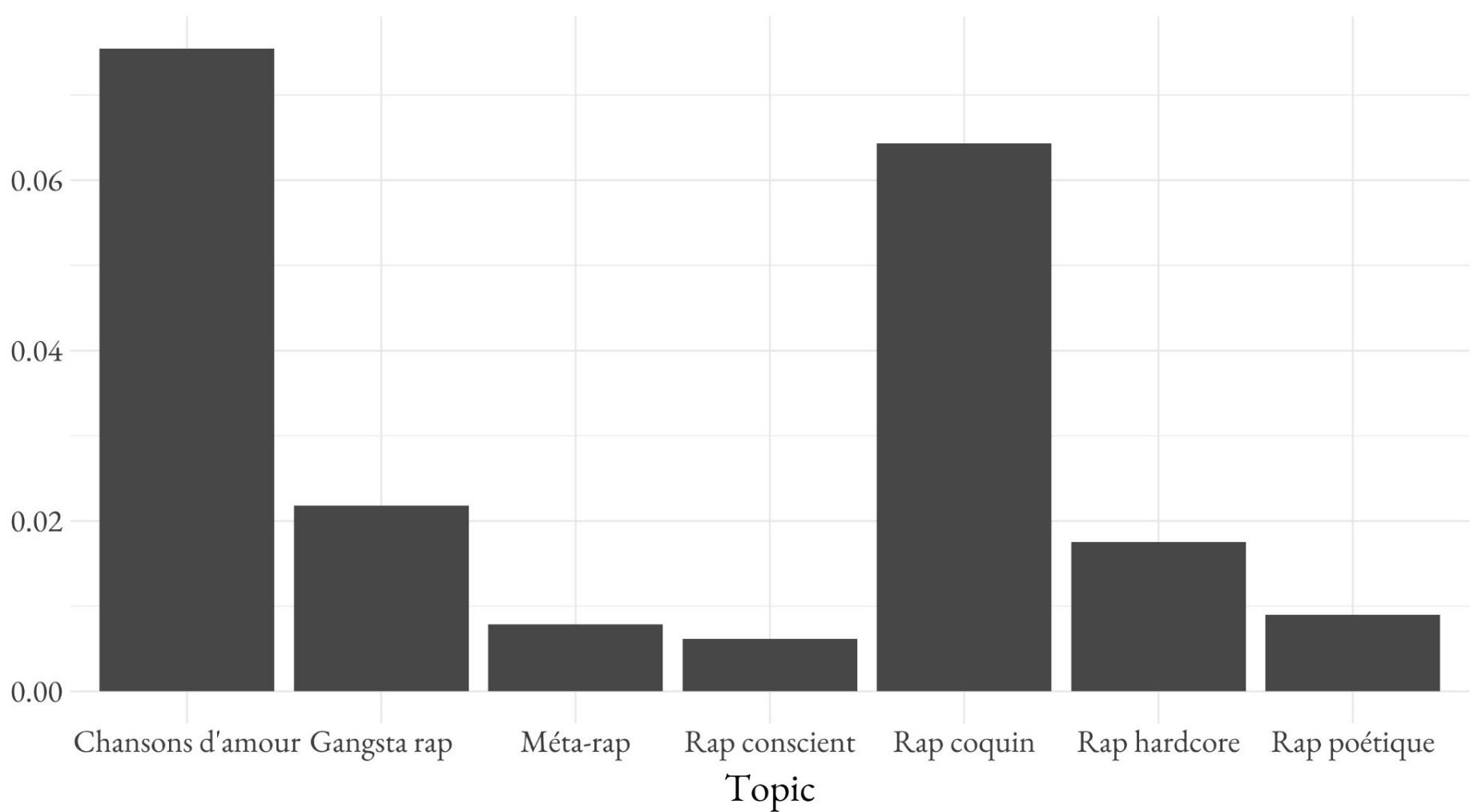
Alphonse Karr, j'arrive mal fonscar

J'cons mon paki, j'arrive chinois comme Jackie

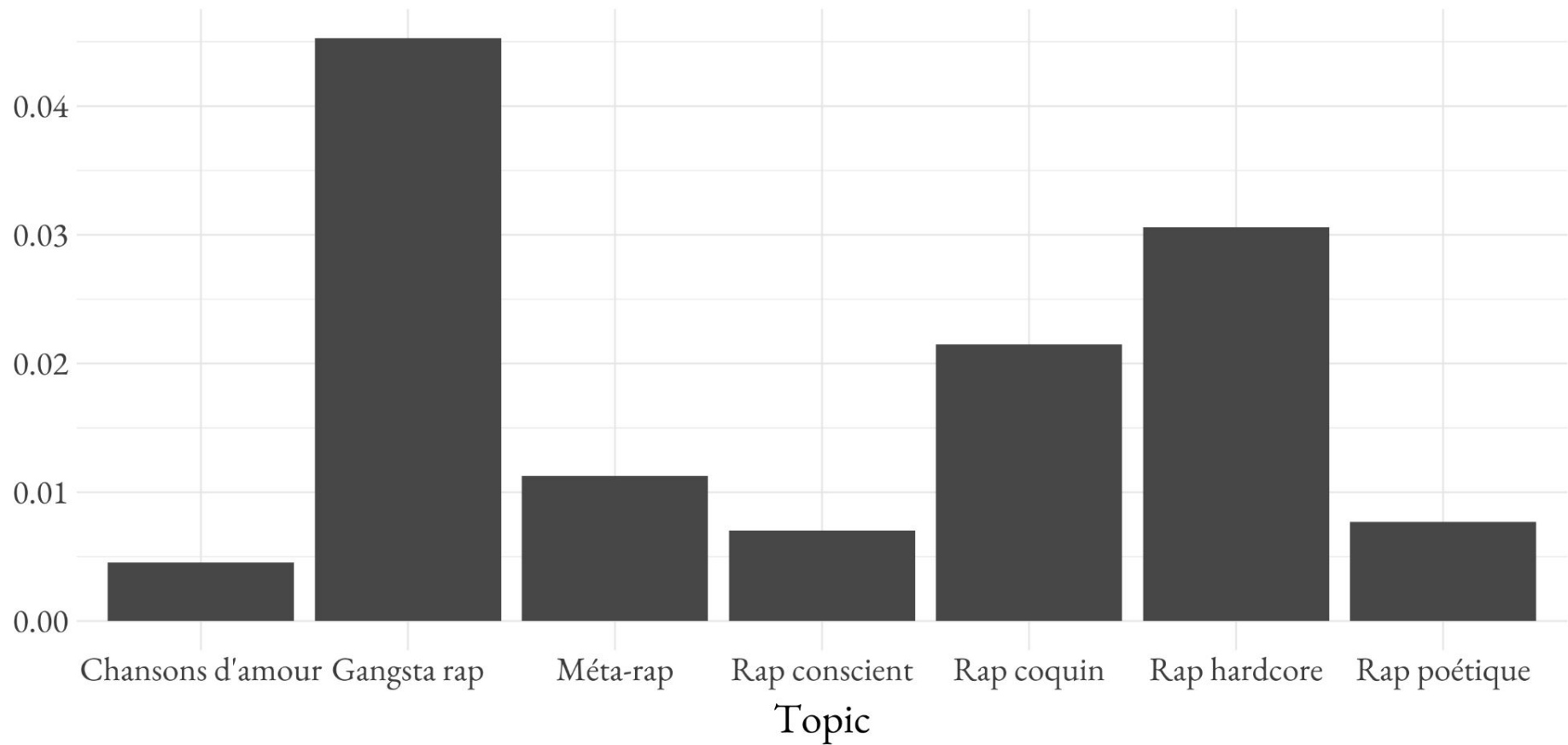
S/o Kaki, au micro nous sommes comme vos grands-paters

667 ekip des comme nous t'en trouves pas d'autres dans la terre

Fréquence de bébé



Fréquence de 'beuh'



### III - Une approche confirmatoire : mesurer les ruptures

- Thématique
- Syntaxique
- Énonciative

# Blanche, Maes ft. Booba

<couplet=1>

<locuteur=Maes>

Très loin d'être parfait mais t'inquiète pas, ça ira  
Au ciel, on va aller, / j'veais laisser mes biens en bas

Toutes mes années d'galère où j'me disais ça ira  
Fais un pas de travers, / une bécane, tu t'en iras

J'v'esqui le parquet, / c'est soit Bercy soit du  
ferme  
J'suis dans le carré, / loin d'Bois-d'Arcy, j'ai mon  
fer

J'barode dans la ville, / sept cents chevaux / Ferrari  
Tu connais l'tarif, / fais pas l'grossiste pour un kil'

J'bicrave à pas d'heure, / j'vends la moola dans le  
bendo  
Douze heures à douze heures, / j'fais partir trois  
kilos d'aya

Ta go' veut un marmot, / moi, j'voulais juste la ken  
Te-shi m'rend paro, / j'me rappelle plus d'la veille

<refrain>

J'visser la night, elle va s'en remettre, si tu savais  
Blanche ou caramel, j'ai son remède

J'visser la night, elle va s'en remettre, si tu savais  
Blanche ou caramel, j'ai son remède

<locuteur=Booba>

Nous restâmes sur le bateau, jamais il ne chavira  
Je devais voir Anne-So', au final, j'ai vu Samira

Passe-moi à travers, jamais tu n'me graviras  
Tu n'as pas de valeur, jamais on n'se ralliera

<couplet=2>

Je n'connais pas le banc à part celui de la G.A.V  
Drapeau est noir et blanc, je reste devant, je préfère être  
à l'attaque

Si j'n'ai plus d'limite, je m'arrête où ?  
J'pends mon linge sale devant personne, je rachète tout

Je ne demande qu'à t'aimer, te montrer mon Makélélé  
Le Versace, le D&G, elle est trop bonne, wallaye, c'est B

Je vais l'emmener à l'étranger, peut-être sur place, j'veais l'  
échanger  
Faire le hlél ou l'étrangler, j'suis qu'en léger, que la CB

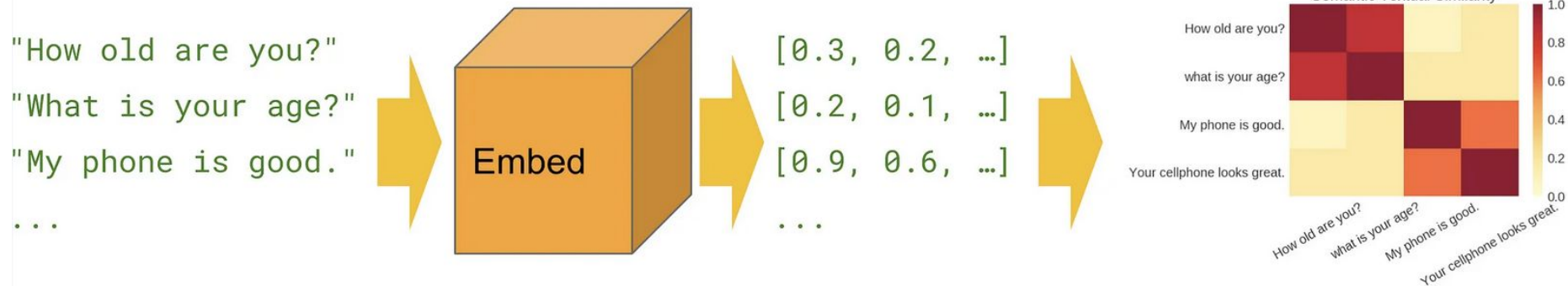
J'drague à la Momo Henni, beaucoup de cash, je la séduis  
On m'tire dessus, j'entends des bruits, même en réserve, je la  
détruis

J'viens du 9.2, la quality, tu connais d'jà l'état d'esprit  
Elle m'a d'mandé combien ça coûte comme si j'avais regardé  
l'prix

<refrain>

<outro>

# Les “plongements sémantiques”



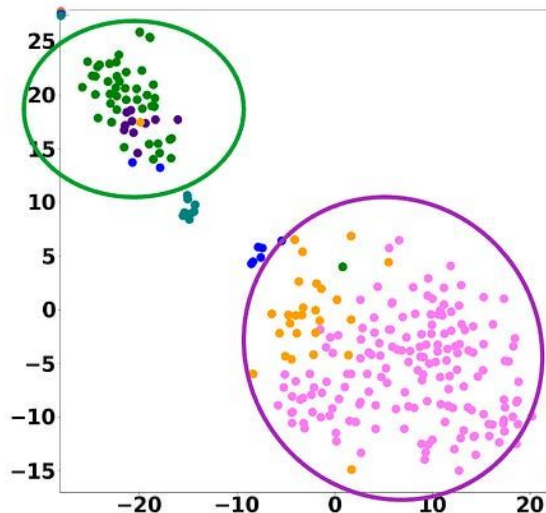


# Les “plongements sémantiques”



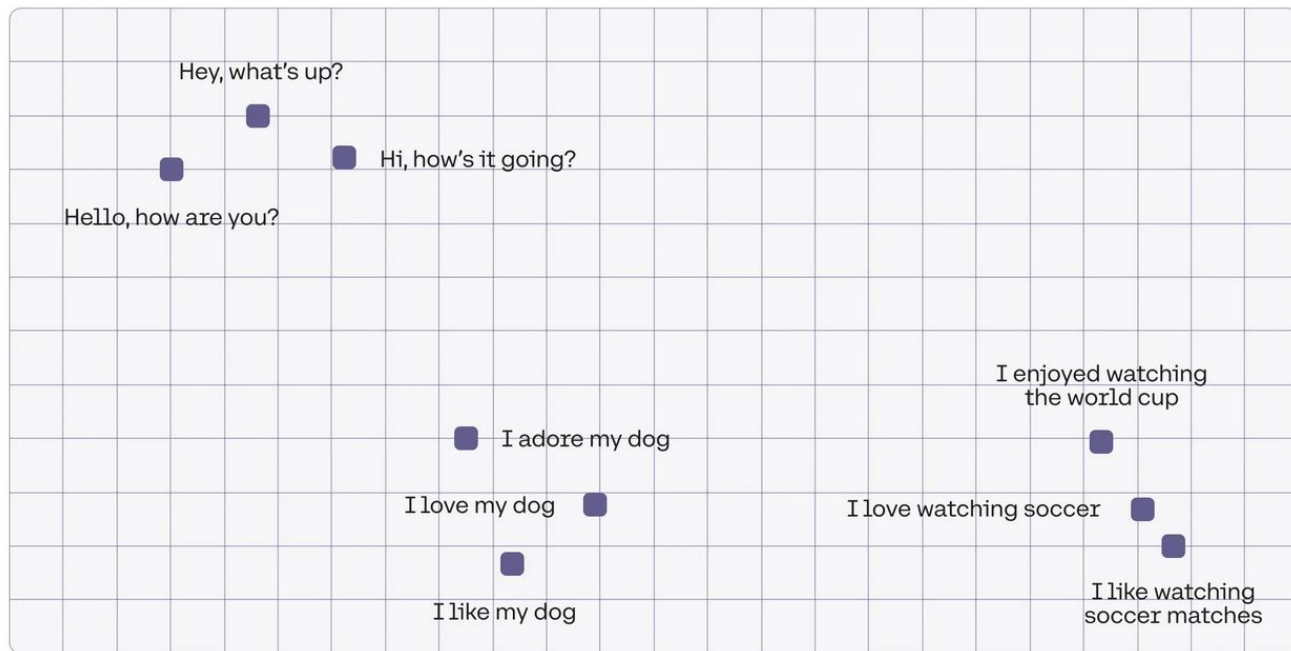
# Les “plongements sémantiques”

The word 'bank'  
when used in  
water sense



The word 'bank'  
when used in  
financial sense

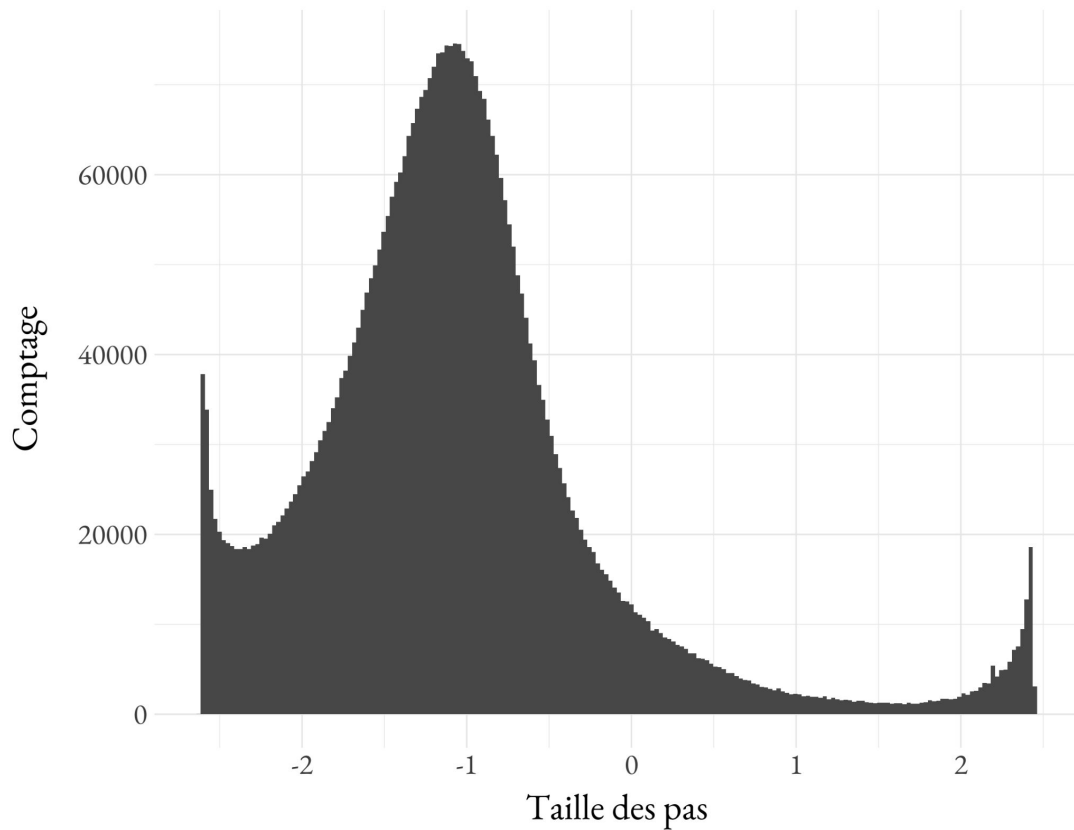
# Les “plongements sémantiques”



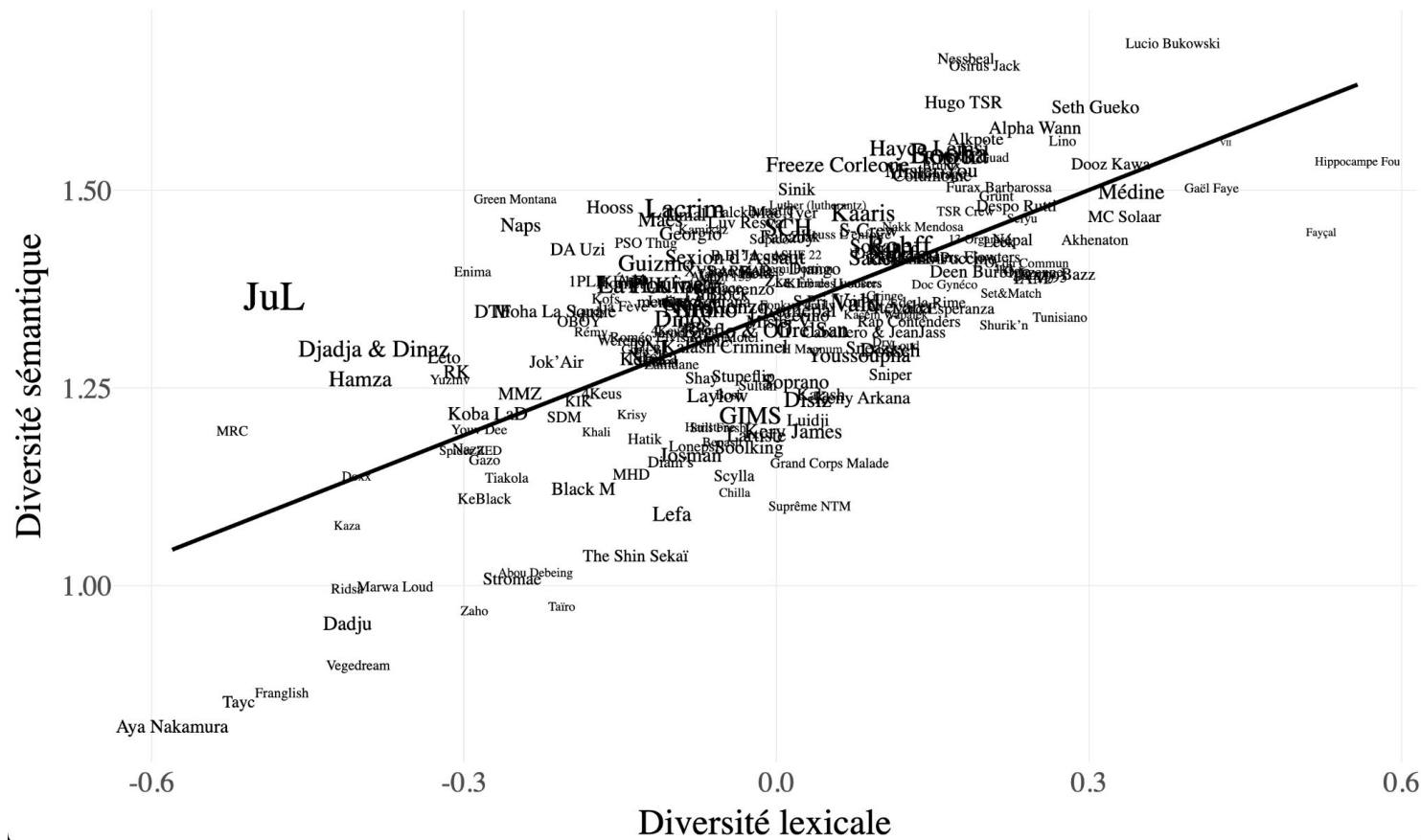
-> On peut mesurer une “distance sémantique” entre deux lignes !

-> Vegedream a la plus faible “taille de pas” moyenne, Lucio Bukowski la plus grande

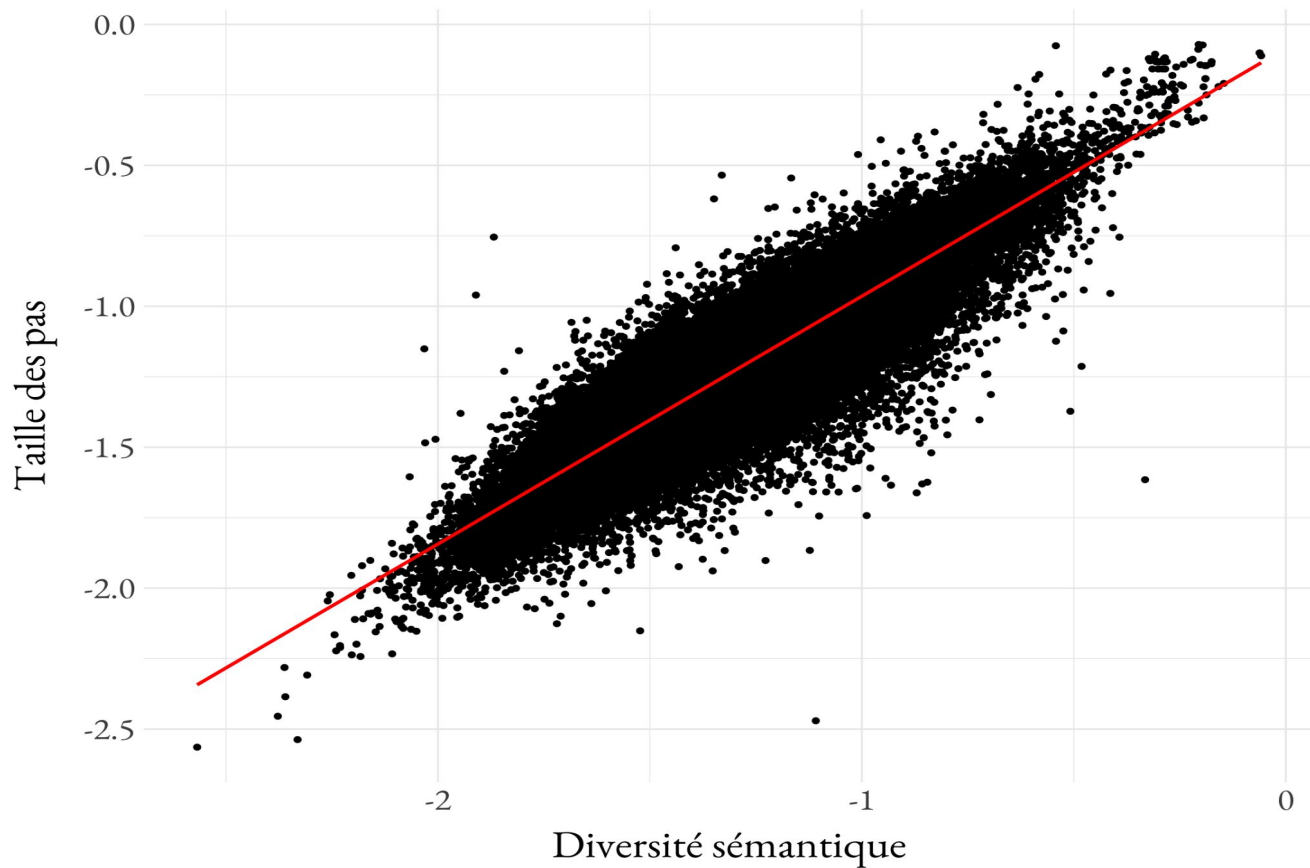
# Problème 1 : les refrains



## Problème 2 : la diversité sémantique



La solution : les “puzzles de mots et de pensée”



## Les chansons les plus cohérentes de Médine

artist	title
Médine	Enfant du destin (Petit Cheval)
Médine	Arabospiritual
Médine	Enfant du destin (Kounta Kinté)
Médine	Sister Act
Médine	Besoin de résolution
Médine	Du Panjshir à Harlem
Médine	Combat de femme
Médine	Enfant du destin (Daoud)
Médine	Biopic

# Le virage de la “trap consciente”





Diversité sémantique

1.50

1.25

1.00

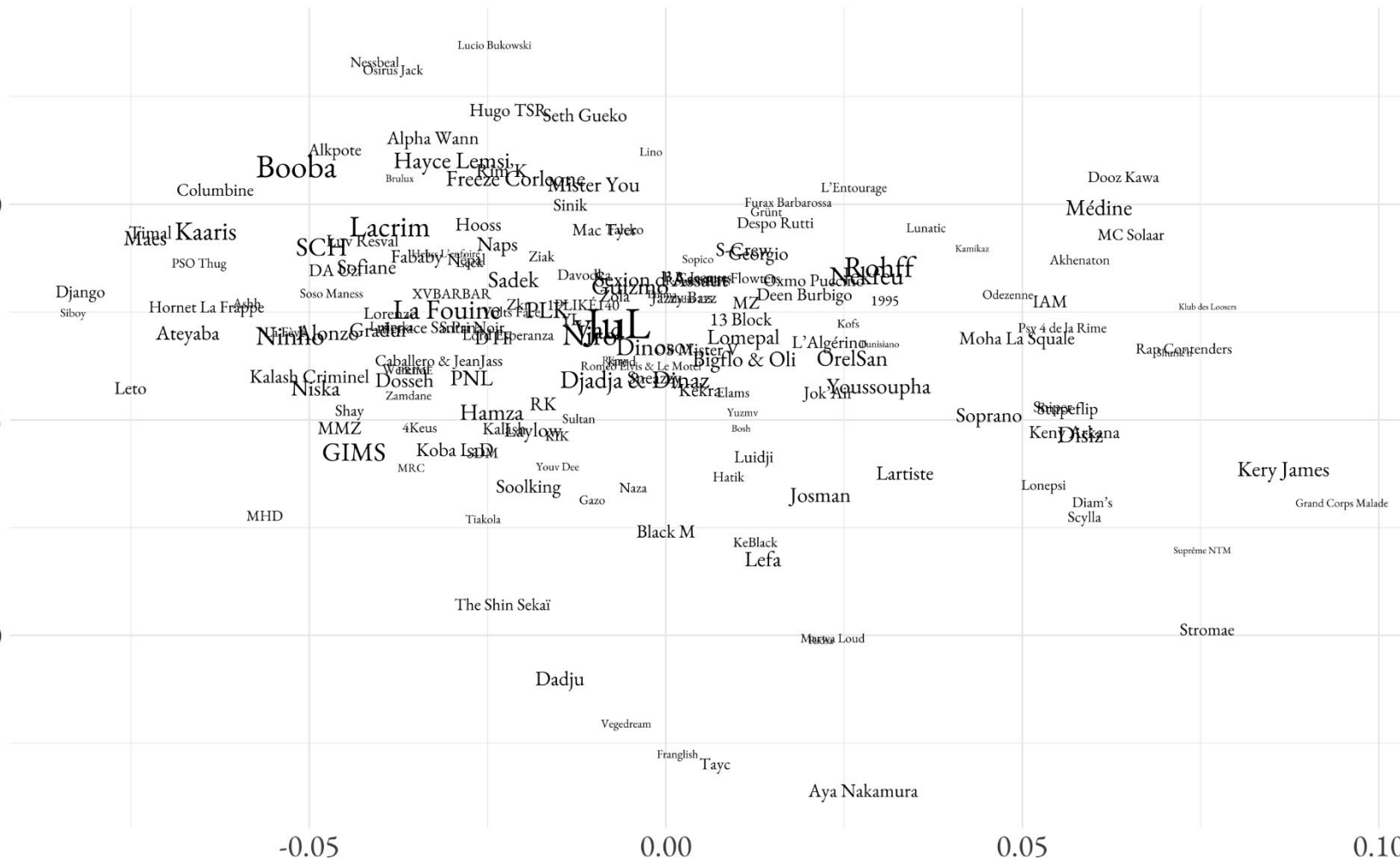
-0.05

0.00

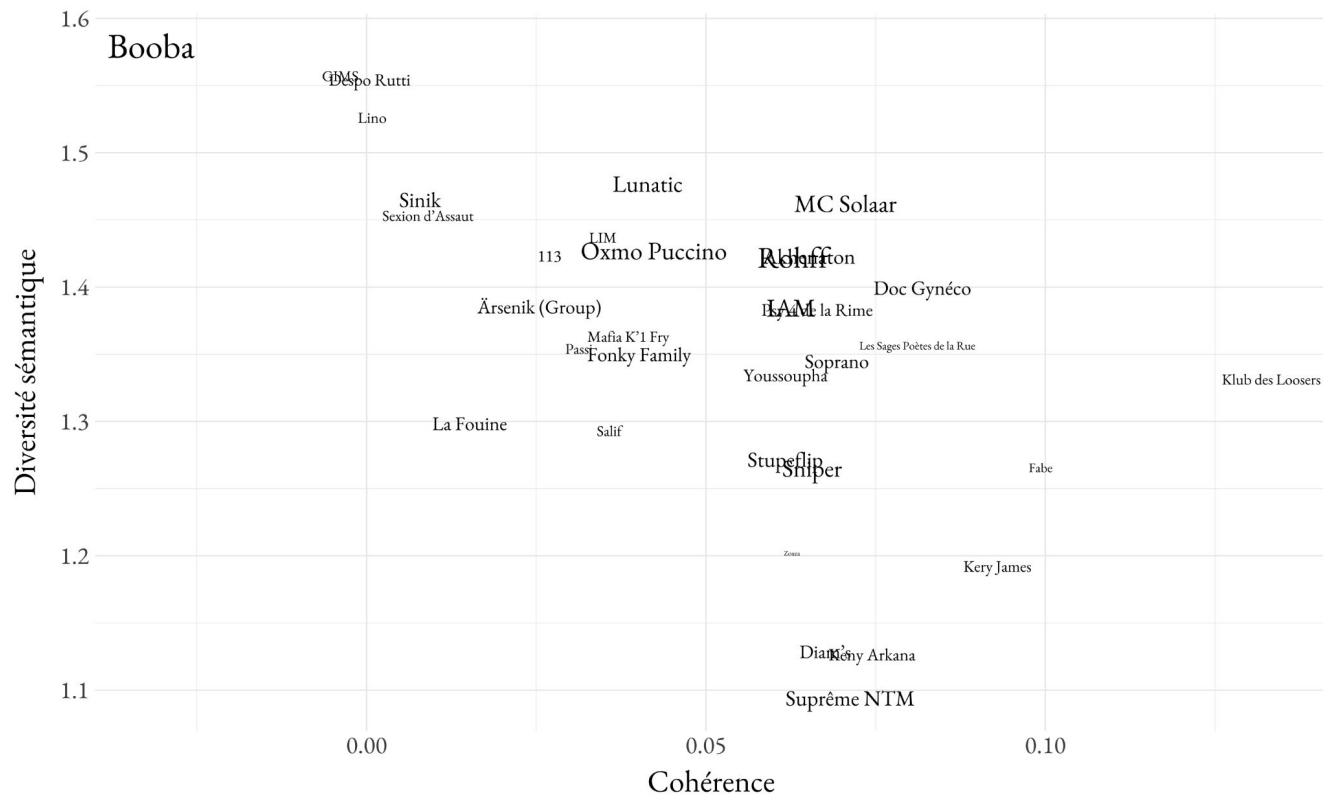
0.05

0.10

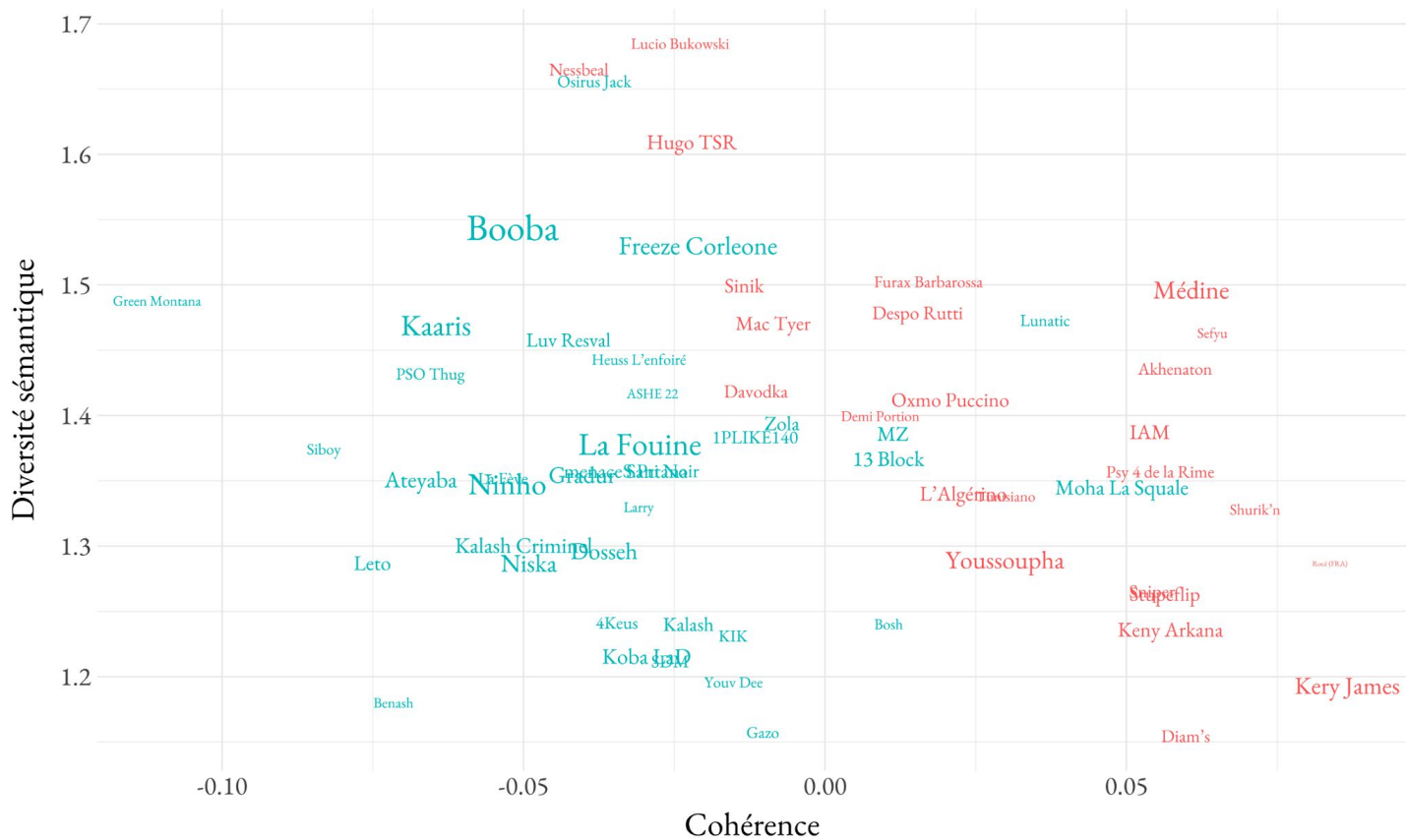
Cohérence



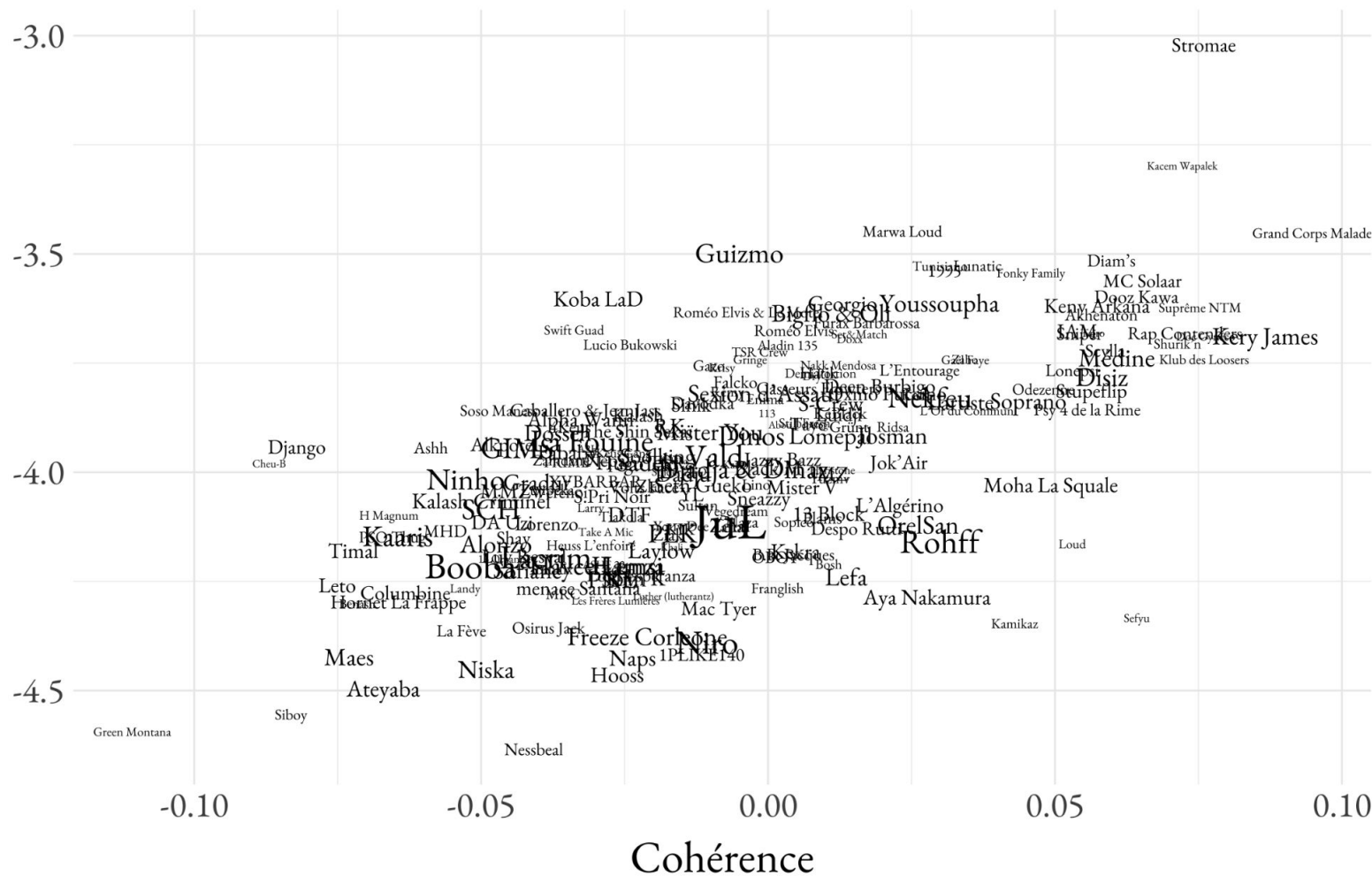
# Booba, pionnier de l'incohérence



# Une dimension structurante



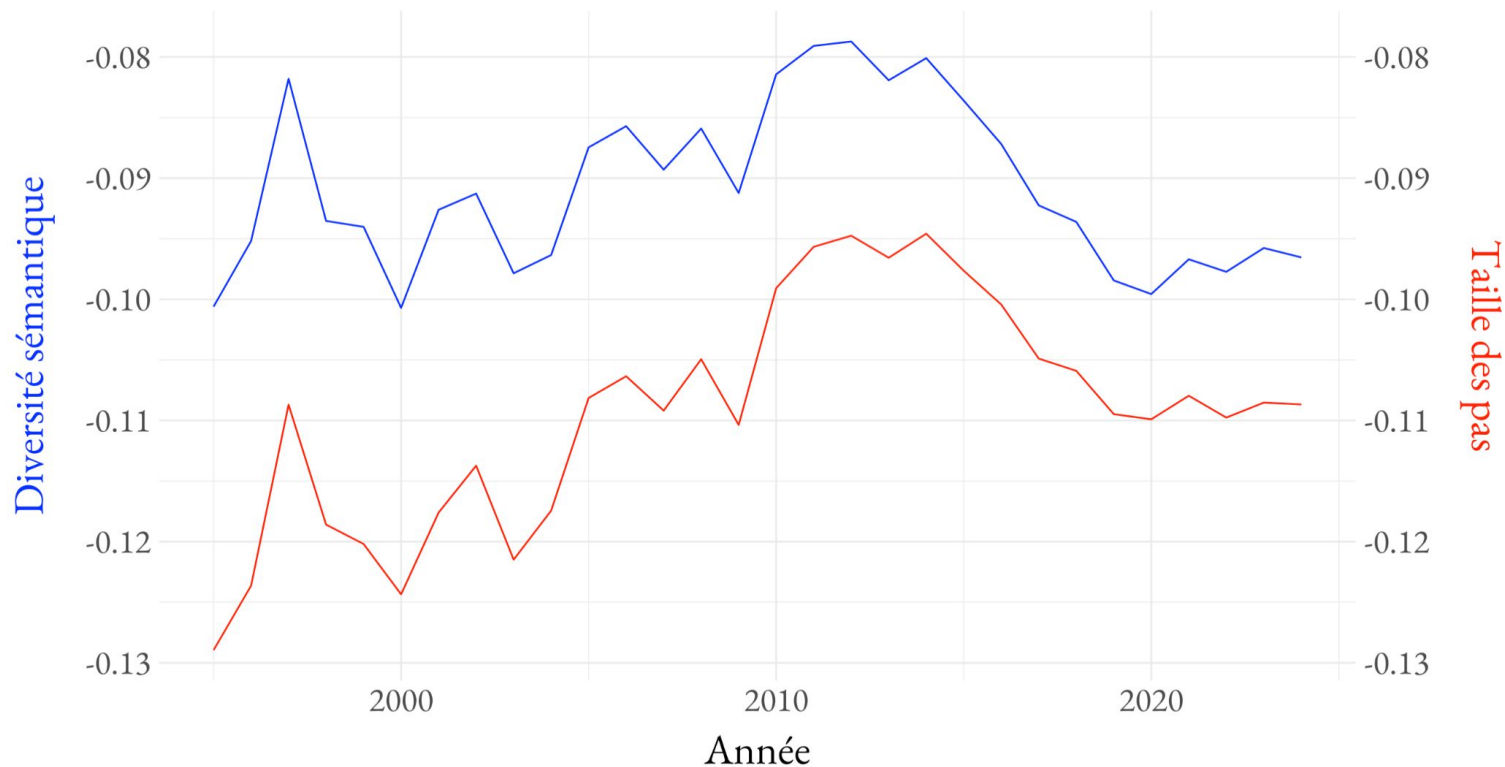
Conjonctions de coordination (log)



# Le triomphe de l'incohérence



# Le triomphe de l'incohérence



Hardcorité

3  
2  
1  
0  
-1  
-2

-3

-2

-1

0

1

2

Littérarité

Tayc

Aya Nakamura

Dadju

KeBlack

Verbes

The Shin Sekaï

Stromae

Auxiliaires

Black M

Djadja & Dinaz

KIK

Onomatopées

Interjections

Je

Jul

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino

DA Uzi

Lartiste

Kois

Dino