

Towards a computational and pattern-based stylistics



Jean-Gabriel Ganascia
ACASA team – LIP6
University Pierre et Marie Curie
OBVIL Labex
Sorbonne University
4, place Jussieu, 75252 Paris, Cedex 05
Jean-Gabriel.Ganascia@lip6.fr

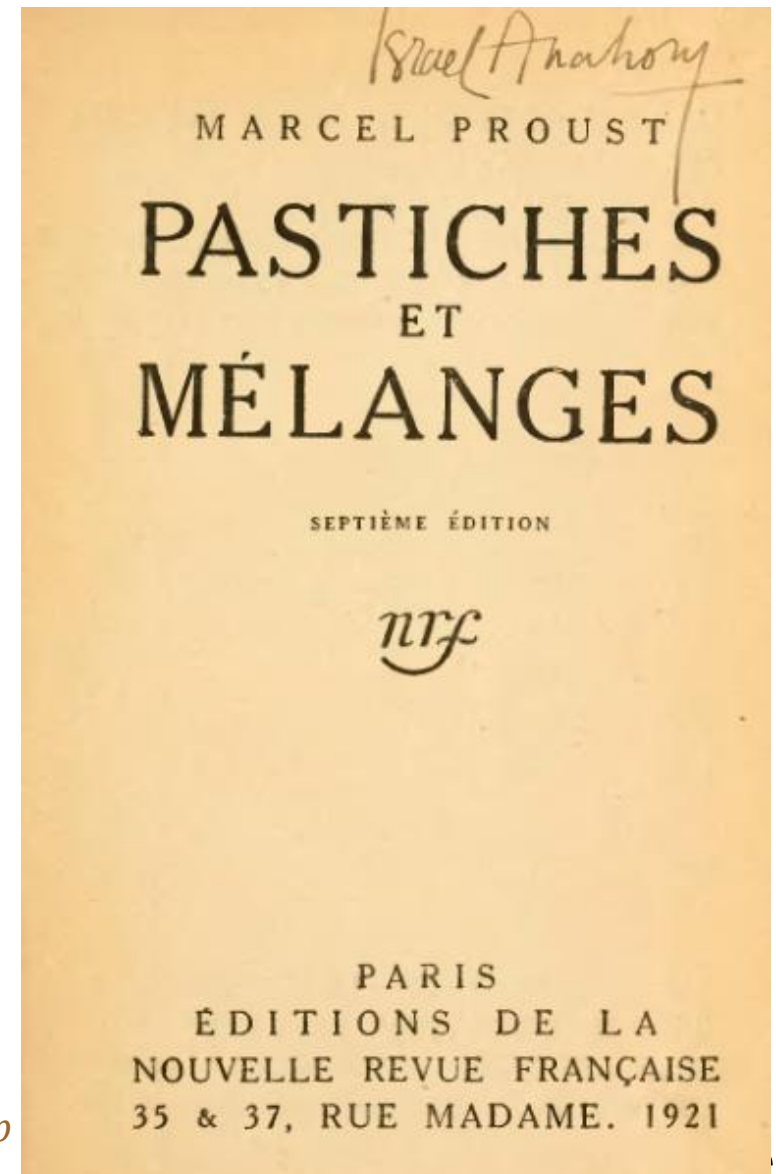


Jean-Gabriel GANASCIA



Synopsis

- What does mean “style”?
- Recurent patterns in music: FlexPat
- Possible features of literary style
- The “Littératron”
- Syntactical patterns: EReMoS
- Semantical patterns



Style - Stylistic

- Etymology: stilus
- Characteristics
 - Genre (letter, theater, novel, ...)
 - Author
 - Characters (in drama)
 - Epochs
 - Gender
 -
- Conscious vs. unconconscious?



Automatic Extraction of Patterns

UPMC

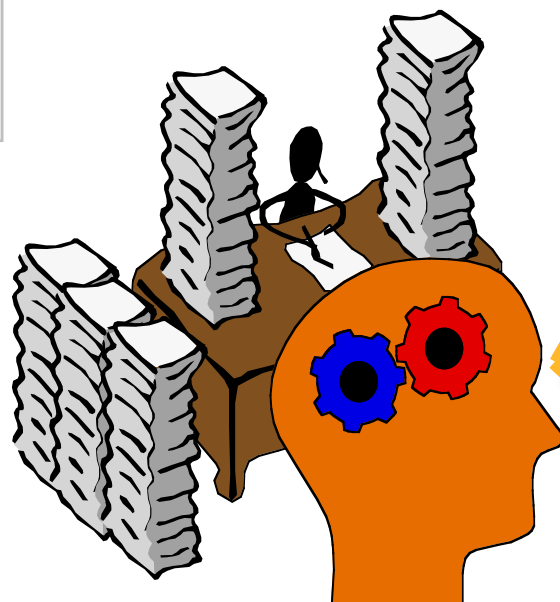
The image displays a musical score with three staves, each featuring a timeline at the top with measure numbers. The first staff has measures 9 through 14, the second staff has measures 22 through 25, and the third staff has measures 35 through 38. Each staff contains musical notation with notes, rests, and accidentals. Three specific patterns of notes are highlighted with orange ovals: the first pattern is in the first staff (measures 10-13), the second pattern is in the second staff (measures 22-25), and the third pattern is in the third staff (measures 35-38). The patterns consist of sequences of eighth and sixteenth notes, some with accidentals, and are followed by a triplet of eighth notes.

Charlie
Parker
Solos

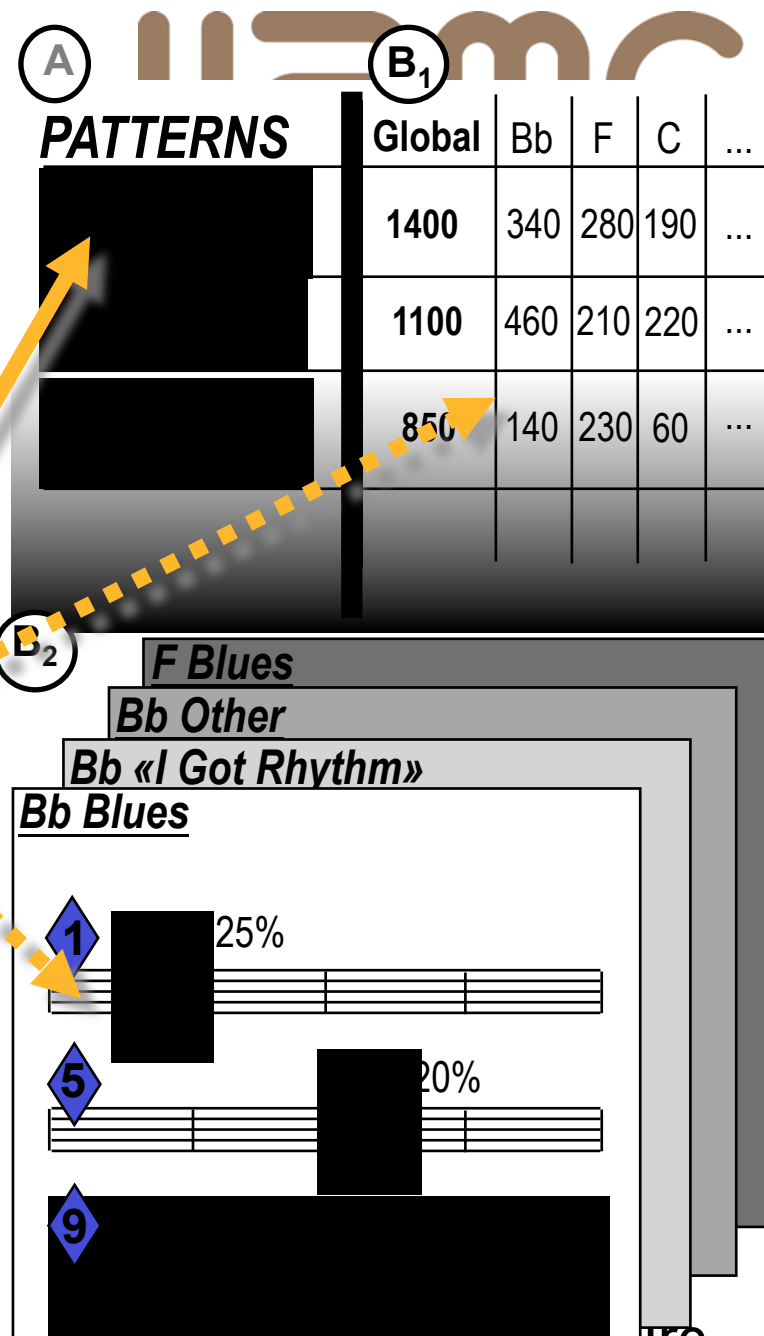
Lexicon of 200
melodic patterns



T. Owens (1974)



PhD Thesis P-Y Rolland (1998)
Multi-Valued Edit Model, i.e. linear
combination of Edit Distances

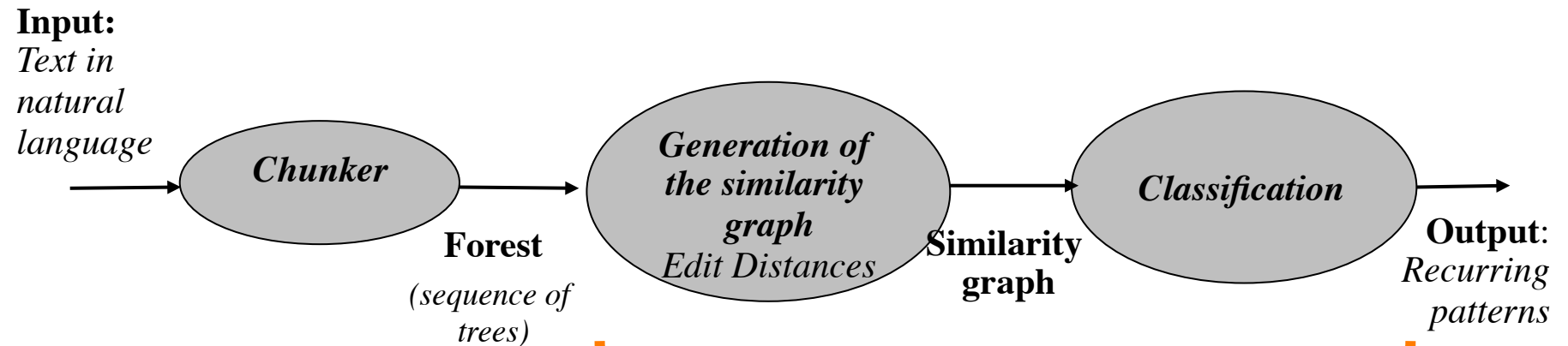


Features of Styles in Literary Studies

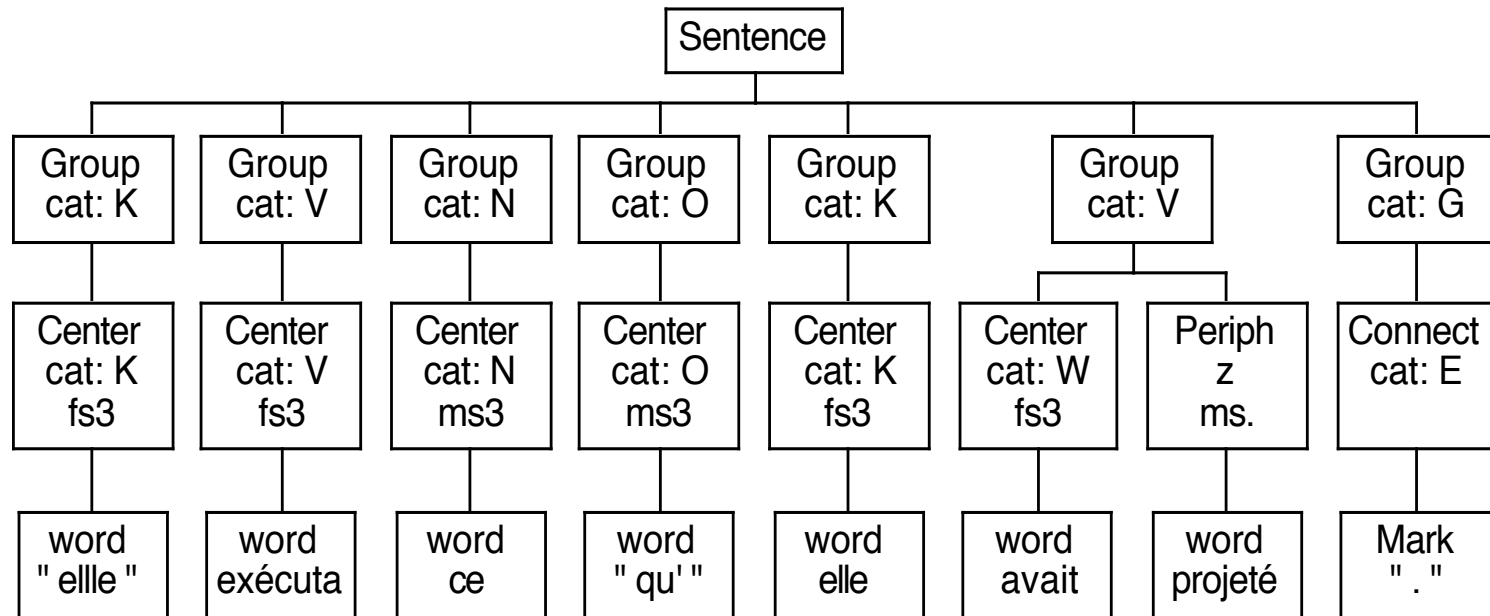


- Philology:
 - characteristics of an author: its syntax
- Lexicon
 - stop words → syntax
 - “heavy” words → semantics
- Syntactical characteristics
- Rhythm (e.g. dactyl, iamb, ...) and punctuations
- Semantical characteristics: figures

The “Littératron”*



Similarity between patterns
Use of Edit Distances



Elle exécuta ce qu'elle avait projeté.

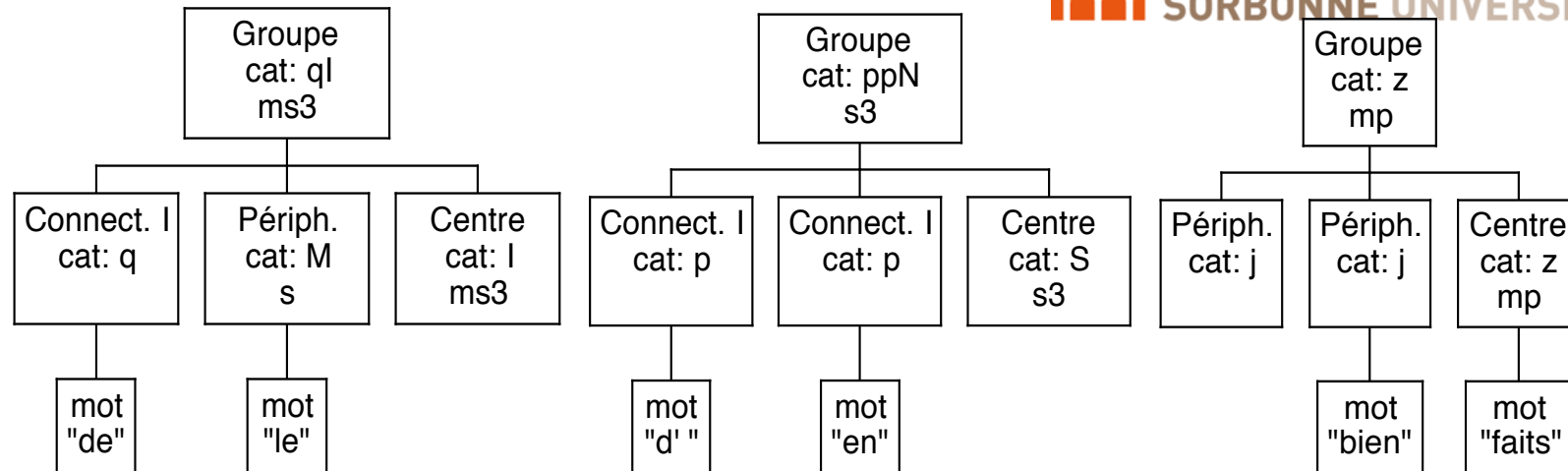
Madame de Lafayette (La comtesse de Tende)

Discrimination

- Patterns in **Madame de Lafayette** in the two following texts: “ *La comtesse de Tende* ” and “ *La princesse de Clèves* ”.
- And absent from many 19th century French authors : **Guy de Maupassant, Georges Sand et Marcel Schwob**.

Guy de Maupassant, *La peur* (1882), *La peur* (1884), *La veillée*, *La rempailleuse*, *Pierrot*, *En mer*, *Un normand*, *Ce cochon de Morin* and *Les sabots*, from **Georges Sand**, *La fée poussière*, *Le gnome des huîtres*, *Le marteau rouge*, *L'orgue du titan* and *La fée aux gros yeux*, from **Marcel Schwob**, *Arachné*, *Béatrice*, *Sur les dents*, *Le Dom*, *L'homme double*, *Le fort*, *Gabelous*, *Parabole*, *Lilith*, *Conte des œufs* and *Les portes de l'opium*.

Three specific patterns

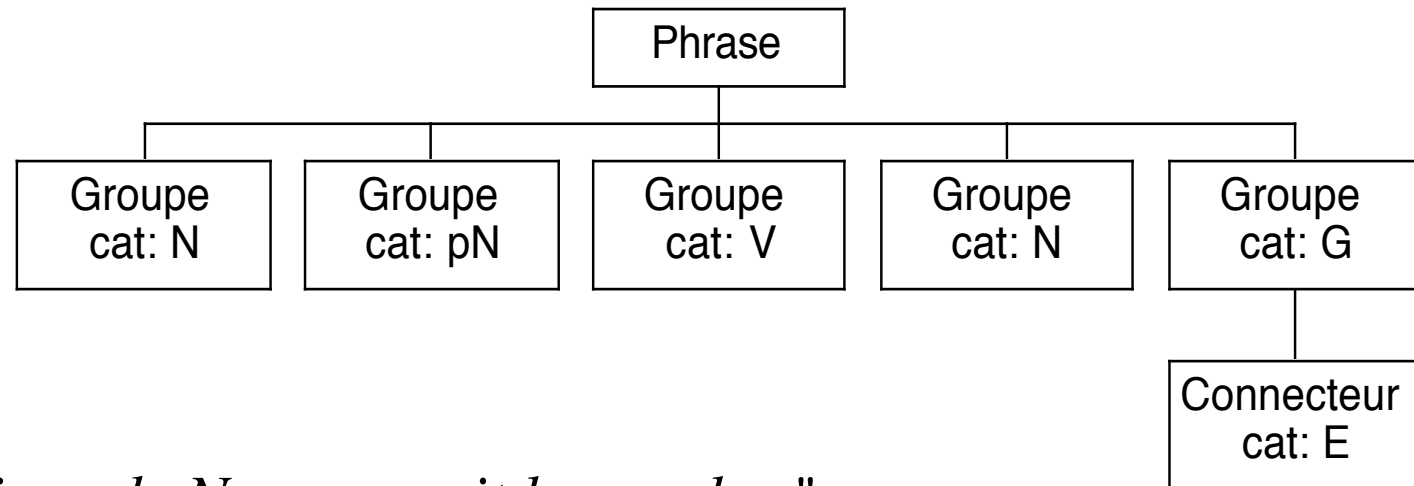


The **first** covers the following expressions : "*à le servir*", "*de le supplier*", "*de l'éviter*", "*de l'aimer*"...

The **second**, "*sans en avoir*", "*d'en avoir*", "*d'en attendre*", "*d'en garantir*", "*d'en faire*"...

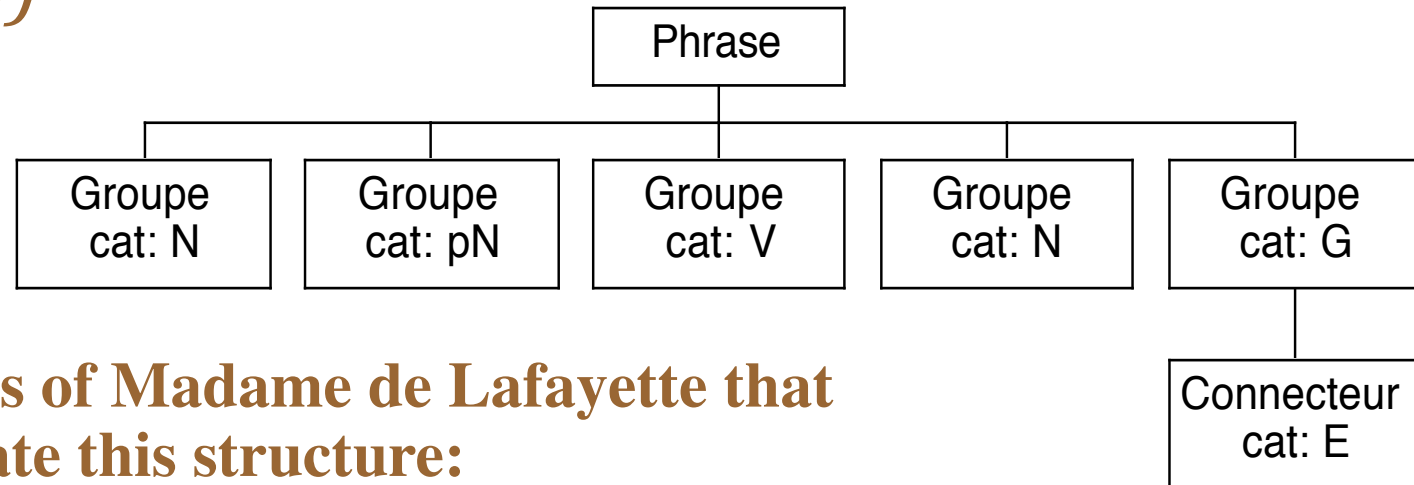
The **third** covers three fragments : "*admirablement bien faits*", "*parfaitement bien faits*", "*très bien fait*"

Typical Structure of Madame de Lafayette



*"Le prince de Navarre prit la parole :",
"La reine de Navarre avait ses favorites",
"Madame de Clèves ne répondit rien",
"Monsieur de Nemours prit la reine dauphine",...*

Typical Structure of Madame de Lafayette *(following)*



Fragments of Madame de Lafayette that approximate this structure:

"Madame de Chartres avait une opinion opposée ;",

"Le comte de Tende aimait déjà le chevalier de Navarre ;",

"Le comte de Tende sentit son procédé dans toute sa dureté ;",

"La comtesse reçut ce billet avec joie", ...

EReMoS

Acronym for “*Extraction et Recherche de Motifs Syntaxiques*”



- Lemmatization of text
- Extraction of frequent sequences of lemmas (i.e. POS tags)
- <http://eremos.lip6.fr/>
- Available for French and German and soon for Spanish and English
- “Le silence profond régnait nuit et jour dans la maison.”
 - <DET, NOM, ADJ, VER, NOM, KON, NOM, PRP, DET, NOM, SENT>
- Sequential patterns:
 - <DET><NOM><ADJ>
 - <NOM><ADJ><VER><NOM>
 - <KON><NOM><*><DET><NOM>

Dictionnaire Le Robert
Le Grand Robert de la langue fr...
EReMoS - Accueil

eremos.lip6.fr
Rechercher

EReMoS
EReMoS -(GERMAN)
AIDE +
À PROPOS
Observatoire de la vie littéraire

Aucun fichier

PARCOURIR

Veillez charger un texte (seul les fichiers txt en encodage UTF-8 sont acceptés), le nom du fichier ne doit pas contenir des caractères spéciaux ou d'espaces.

ANNULER

ENVOYER

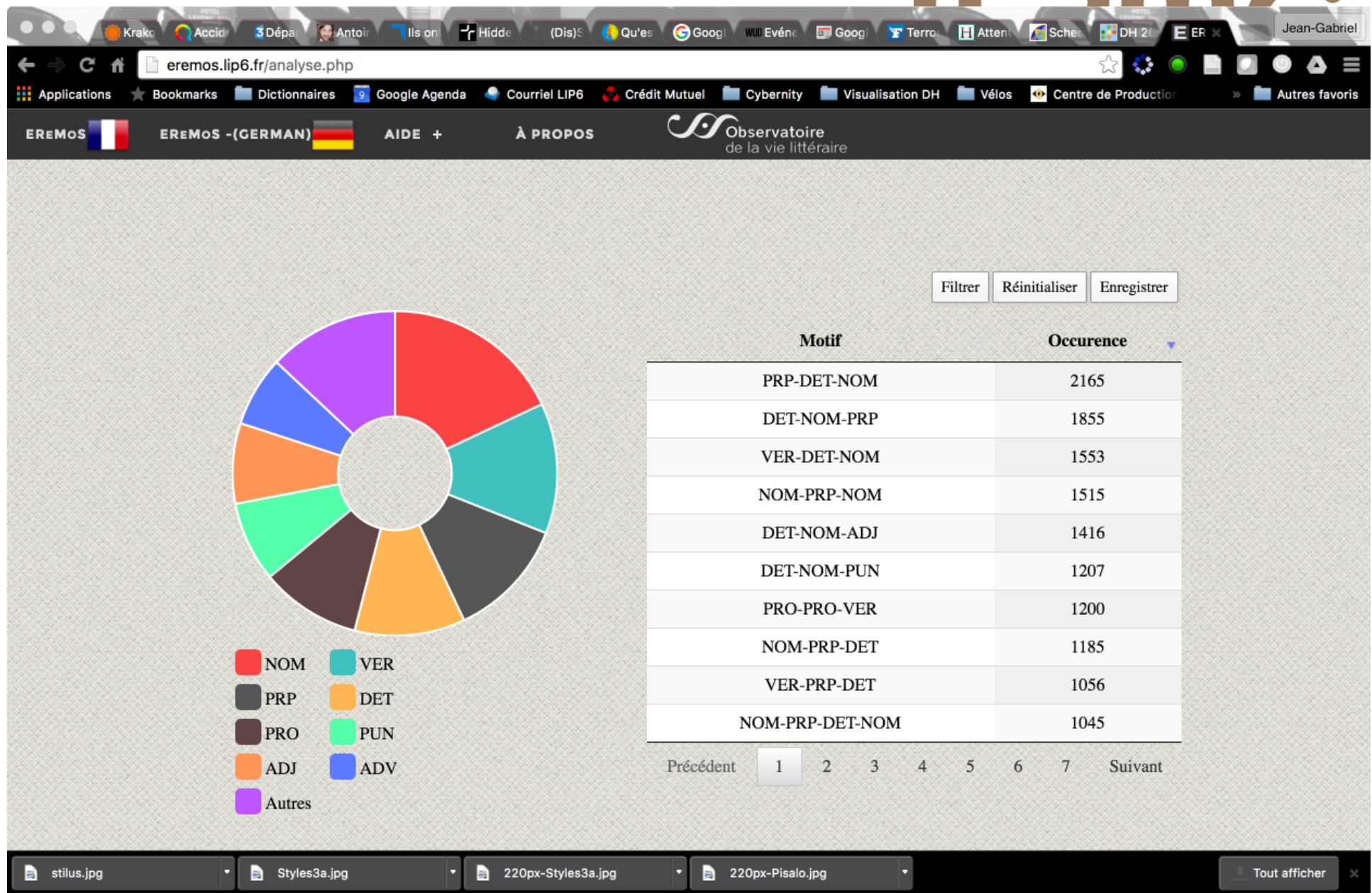
Fréquence d'apparition relative des motifs
10 %

Fréquence d'apparition absolue des motifs
10

Longueur des motifs

minimale
3

maximale
5



ere mos.lip6.fr/analyse.php

Applications Bookmarks Dictionnaires Google Agenda Courriel LIP6 Crédit Mutuel Cybernity Visualisation DH Vélos Centre de Production Autres favoris



DET-NOM-PRP-DET-NOM	653
PUN-KON-PRO	644
KON-PRO-VER	641
ADV-DET-NOM	624
PRP-NOM-ADJ	609

Précédent 1 2 3 4 5 6 7 Suivant

Précédent Suivant Distribution Enregistrer

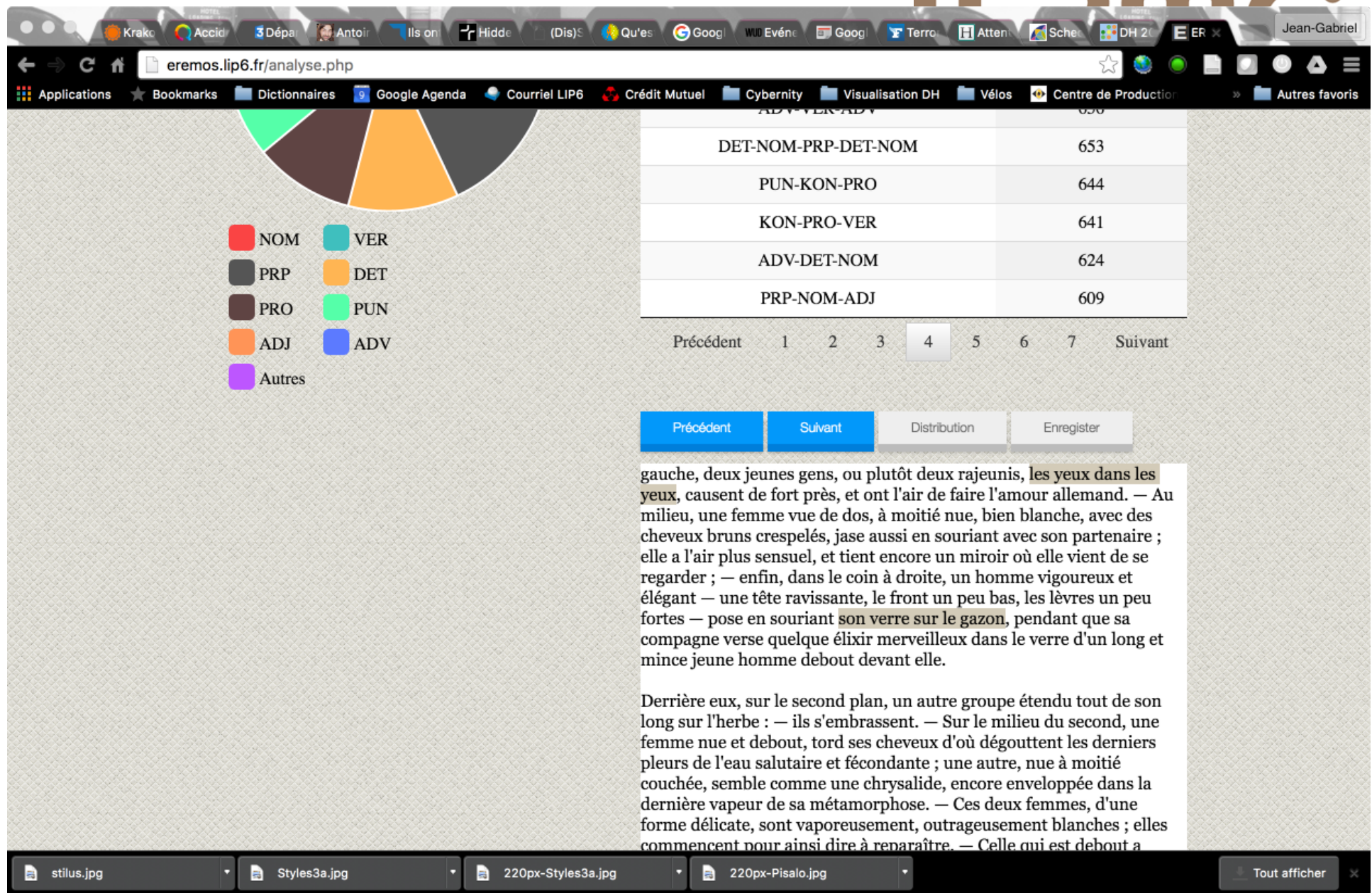
Dessins, le supplément de la galerie Française, le musée Espagnol, le musée Standish, le musée de Versailles, le musée de Marine), un esprit juste verra toujours qu'un grand artiste n'y peut que gagner, vu sa fécondité naturelle, et qu'un médiocre n'y peut trouver que le châ timent mérité.

Nous parlerons de tout ce qui attire les yeux de la foule et des artistes ; — la conscience de notre métier nous y oblige. — Tout ce qui plaît a une raison de plaire, et mépriser les attroupements de ceux qui s'égarent n'est pas le moyen de les ramener où ils devraient être.

Notre méthode de discours consistera simplement à diviser notre travail en tableaux d'histoire et portraits — tableaux de genre et paysages — sculpture — gravure et dessins, et à ranger les artistes suivant l'ordre et le grade que leur a assignés l'estime publique.

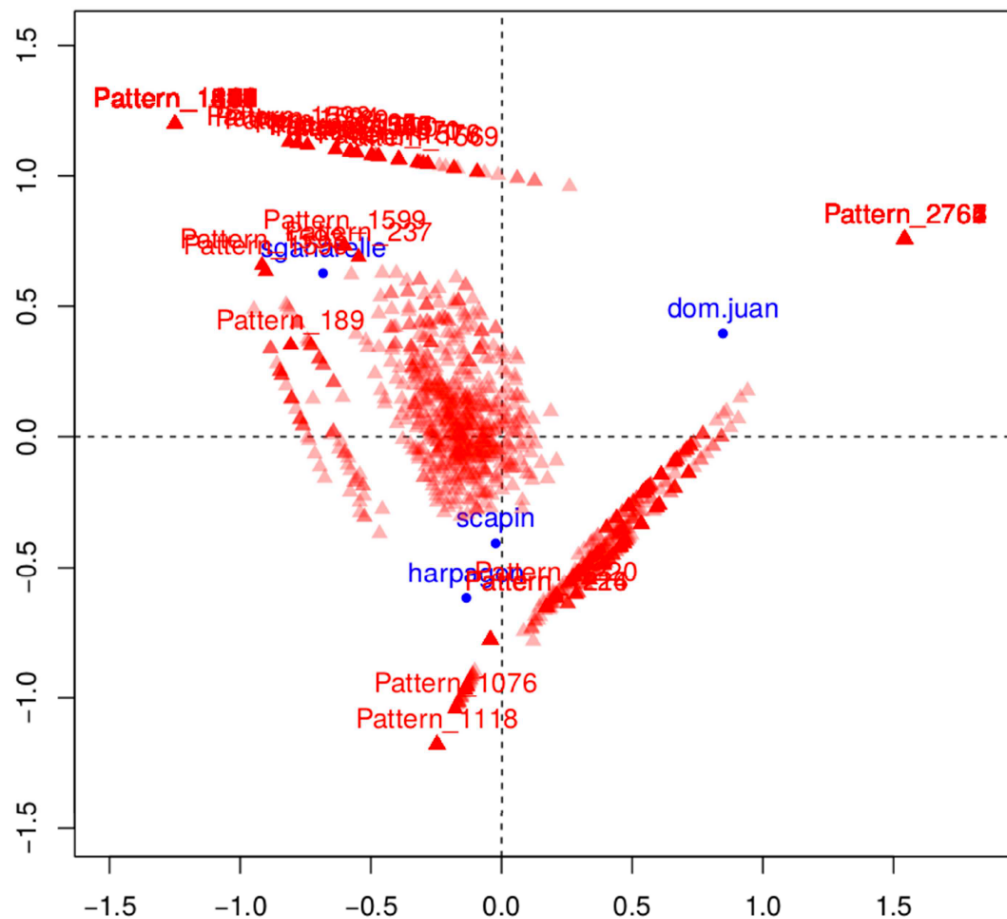
== II. Tableaux d'histoire ==

stilus.jpg Styles3a.jpg 220px-Styles3a.jpg 220px-Pisalo.jpg Tout afficher



Memorable Molière's Protagonists

Boukhaled, Besnard, Frontini 2015



- Don Juan
- Sganarelle
- Scapin
- Harpagon

“interestingness” measure

Boukhaled 2016

- Distribution of patterns in the text

- $D = (d_1, d_2, \dots, d_{n-1})$

- d_i distance between two occurrences

- $\sigma_{norm} = \sigma / \sigma_{geo}$

$$\sigma = \frac{\sqrt{E(D^2) - E(D)^2}}{E(D)}$$

$$\sigma_{geo} = \sqrt{1 - \frac{1}{E(D)}}$$

<PUN><ADV><PRO> (119 occurrences)



<PUN><NOM><NAM> (120 occurrences)



Positions of the pattern in the text

Distribution of patterns

<ADV> <PRO> <NOM> (SUPPORT= 109)



<NAM> <PRP> <NAM> <PUN> (SUPPORT= 118)



<PRO> <VER> <PRO> <PRO> (SUPPORT= 128)



<NUM> <NOM> <PRP> <NOM> (SUPPORT= 117)



<DET> <NOM> <SENT> (SUPPORT= 2480)

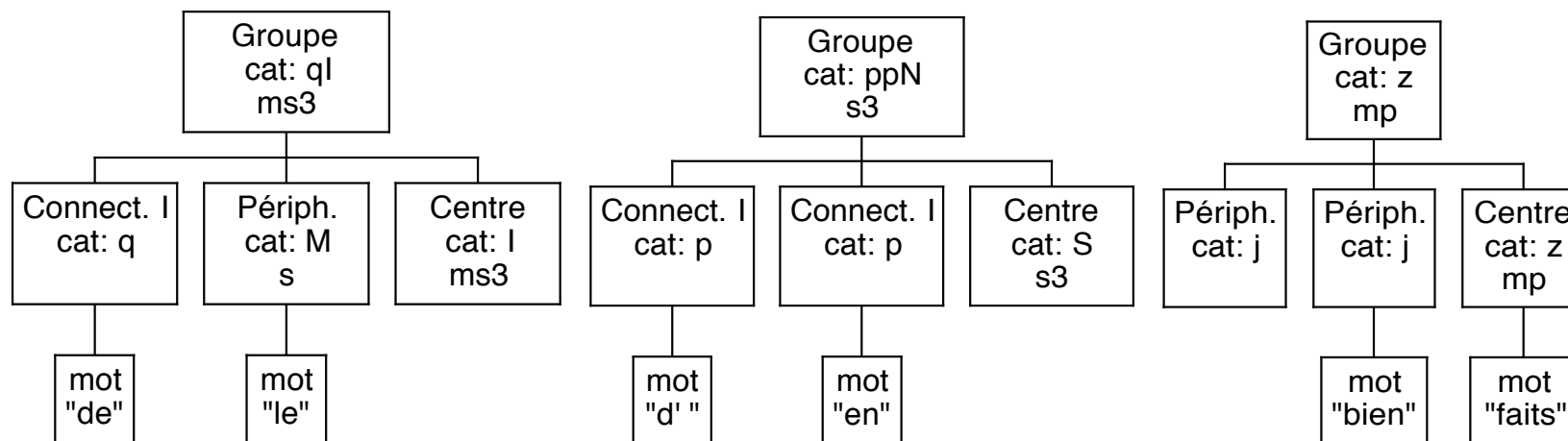


<NAM> <PRP> <NAM> <SUP> (SUPPORT= 215)



Perspective: extraction of “non flat” patterns

- Sequential flat patterns:
 - <DET><NOM><ADJ>
 - <NOM><ADJ><VER><NOM>
 - <KON><NOM><*><DET><NOM>
- Non flat patterns: mixing
lemmas, words, categories, variables
 - <DET><“le”><NOM><ADJ>
 - <KON><NOM><*><DET><NOM>



Detection of Figures

e.g. Extraction of simile

Suzanne Mpouli

And then in the fall the crackling of cornleaves, the smells, sights and sounds.

The corn stood up like armies in the shocks.

When I was a boy I went into the cornfields at night. I said words I had not dared to say to people, defying the New Englanders' gods, trying to find honest, mid-western American gods.

Sherwood Anderson (1876-1941) — "Mid-American Prayer".

