

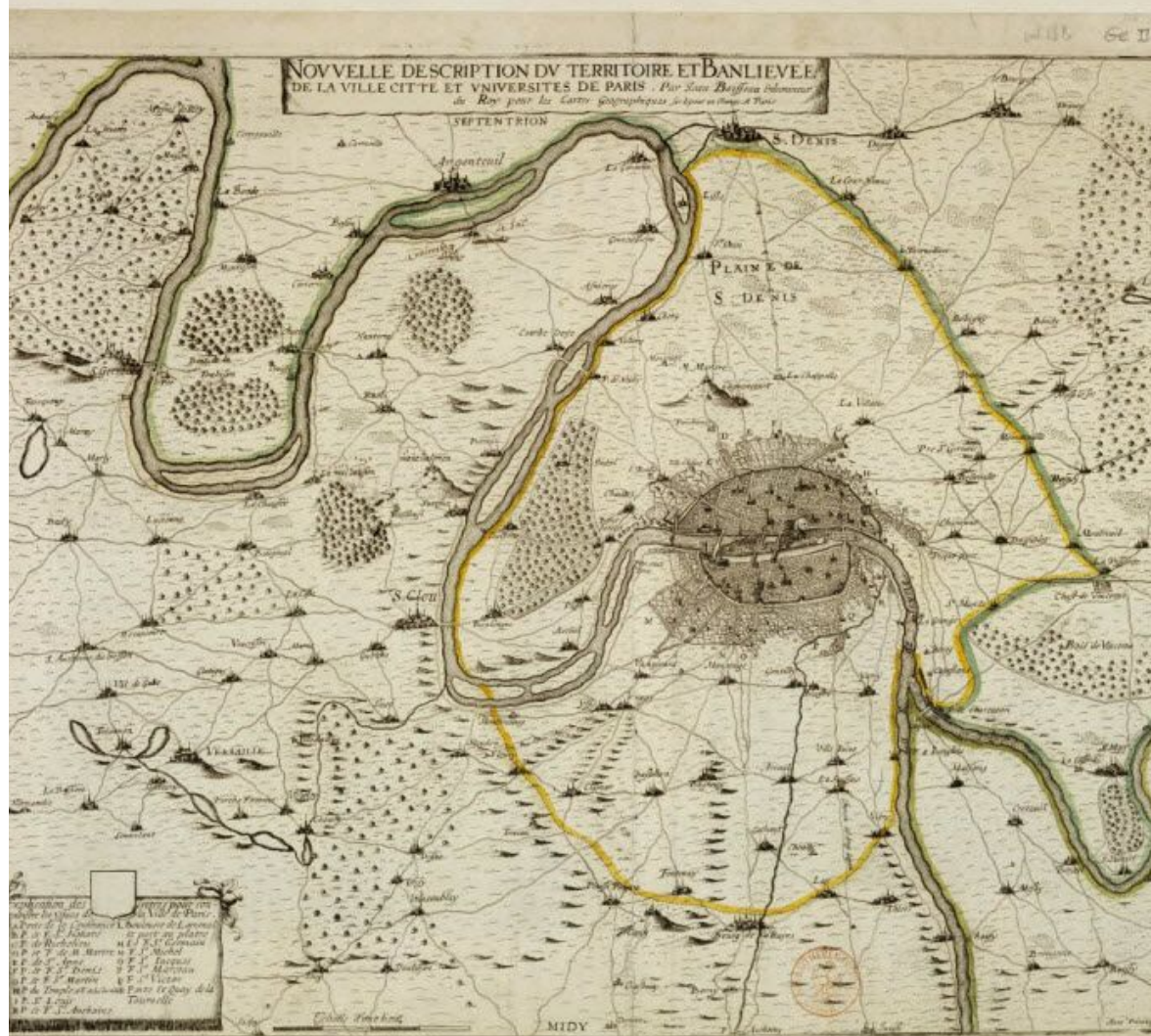
# Extraction automatique d'entités toponymiques et noms communs liés à la ville



# Plan de la présentation

- Définitions
- Reconnaissance d'entités nommées
  - identification
  - annotation
  - approches
- Ingénierie de grammaires pour l'extraction de toponymes
  - Unitex
  - CasEN
- La *ville* dans le corpus Visiautrices

# Définitions



Boisseau, Jean (16..1657?), Nouvelle description du territoire et banlieue de la ville, cité et université de Paris  
<https://images.bnf.fr/#/detail/658396>

# Définitions

Définition par Friburger 2012 :

Les **entités nommées** regroupent tous les éléments du langage définis par référence : les **noms propres** au sens classique, les noms propres dans un sens élargi mais aussi les **expressions de temps** et de **quantités**.

## Les toponymes

Les toponymes sont des entités qui définissent des **localisations**, des **lieux** et des **entités spatiales relatives** et **absolues**. Un toponyme est un nom de lieu, constitué d'un ou plusieurs mots, en rapport étroit avec un **détail géographique localisé** (Dominguès & Eshkol-Taravella, 2015).

Les toponymes se divisent en plusieurs catégories qui peuvent définir les **noms de villes**, de **pays**, de **quartiers**, de **montagnes**, les **voies**, etc.

# Toponymie urbaine (Bouvier, 1999)

## Toponymes d'usage :

- accord implicite entre les habitants du lieu sur la désignation d'un référent qui leur est commun
- caractère **référentiel** prononcé avec une relation sémantique directe entre le déterminant et le référent

**place de l'Horloge, rue de l'Eglise, place du Marché**

## Toponymes de création :

- créés par la décision d'un pouvoir, ils s'imposent en principe à l'usage des habitants et de tous les utilisateurs du lieu
- souvent sans aucune relation de sens avec leur référent
- mémoriels, fonction **symbolique**

**rue Victor Hugo, boulevard Descartes, Place de la République**

# Le statut des toponymes

Sous-ensemble de la catégorie de noms propres qui se présente sous diverses formes :

- avec ou sans contenu sémantique
- constitués d'un ou plusieurs constituants : nom commun, adjectif, mot grammatical, etc.

massif du Mont-Blanc

- association de noms communs et qualificatifs ou adjectifs relationnels et des anthroponymes ou toponymes

noms communs : bourg, île, promenade, port, etc.

qualificatifs/adjectifs relationnels : basse, bonne, grand, gros, etc.



# Reconnaissance d'entités nommées

*Vue intérieure de Paris, représentant le Port au Blé depuis le marché aux veaux jusqu'au pont Notre Dame sur la Seine. Collection de dessins sur Paris réunie par Hippolyte Destailleur.*

<https://images.bnf.fr/#/detail/672450/1>



# Reconnaissance d'entités nommées

## Deux étapes de traitement :

**Identification/reconnaissance** : repérer dans le texte les entités concernées

Le **maréchal de Biron**, qui n'épiait qu'une telle occasion, en étant averti, feint de venir avec son armée près de là pour joindre à un passage de rivière **Monsieur de Cornusson**, **sénéchal de Toulouse**, qui lui amenait des troupes, et au lieu d'aller là, tourne vers **Nérac**, et sur les **neuf heures du matin** s'y présente avec toute son armée en bataille, prêt et à la volée du canon.

**Catégorisation** : attribuer une catégorie préalablement définie

Le **maréchal de Biron**<sup>pers</sup>, qui n'épiait qu'une telle occasion, en étant averti, feint de venir avec son armée près de là pour joindre à un passage de rivière **Monsieur de Cornusson**<sup>pers</sup>, **sénéchal de Toulouse**<sup>pers</sup>, qui lui amenait des troupes, et au lieu d'aller là, tourne vers **Nérac**<sup>loc</sup>, et sur les **neuf heures du matin**<sup>temp</sup> s'y présente avec toute son armée en bataille, prêt et à la volée du canon.



# Annotations

## Trois étapes indispensables:

Guide d'annotation : description des **entités à annoter** et définition des **critères** qui serviront à la personne qui réalisera l'annotation manuelle

Outils d'annotation : outils informatiques basés sur des approches l'apprentissage automatique et/ou l'ingénierie des grammaires

Métriques d'évaluation de la qualité des annotations : série des tests qui ont comme objectif de confronter les annotations produites automatiquement et celles par les annotateurs, étape qui garantit la fiabilité de l'annotation automatique

# Comment identifier une entité nommée ?

McDonald, 1996, distingue deux indices :

**Preuves internes** : mots déclencheurs, abréviations ou à l'intérieur de l'entité elle-même

- séquences de mots de l'entité
- lettre majuscule

Par exemple : **Avenue** Longueil, **Château** de Fontainebleau, **Mont** Saint Michel,  
born / bron (= fontaine, source) : Caden**bronn**, Bett**born**, **Brognard**,  
mari / \*meri (= mare, étang) : Cambre**mer**, Longe**mer**, Gérard**mer**

**Preuves externes** : informations qui accompagnent l'entité nommée en lui attribuant des propriétés spécifiques

Par exemple : **rivière d'**Yerre, **ville de** Saint-Aster

# Comment annoter les entités nommées ?

Deux grandes approches :

## Approche symbolique

- Construction manuelle des modèles d'analyse, sous la forme de **règles contextuelles**  
*CasEN* (Friburger, Maurel, 2004)

## Approche statistique

- Méthodes statistiques ou d'apprentissage automatique qui utilisent des modèles d'analyse à partir de **gros volumes de données annotés**, des modèles d'analyse de textes sous la forme d' **arbres de décision**, d'ensembles de **règles logiques**, des modèles **probabilistes** (Ehrmann, 2008).

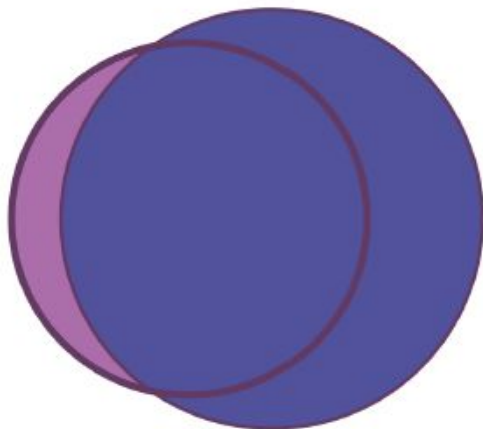
*OpeNER, SpaCy, SEM, Stanford CoreNLP*

## Approches hybrides

- Méthodes qui combinent les approches **statistiques** et **symboliques**  
*msX* (Nouvel, 2012), *Perdido* (Moncla, 2015)

# Avantages/inconvénients (Martinez, 2017)

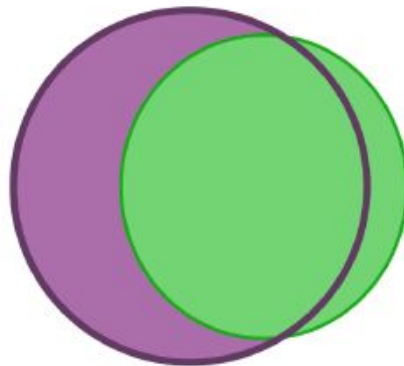
Apprentissage  
automatique  
(basé sur les mots)



+ Rapidité mise en place + Haut rappel  
- Difficiles à déboguer - Textes courts

● Entités pertinentes  
■ Entités trouvées

Ingénierie des  
grammaires  
(basée sur la structure de  
la phrase)



+ Niveau très fin + Haute précision  
- Difficiles à maintenir - Développement lent

● Entités pertinentes  
■ Entités trouvées

# Résultats de Perdido (Moncla, 2015)

- Développé pour analyser des documents en langues modernes en combinant l'utilisation de l'analyse morphosyntaxique et des automates en cascade

Evaluation sur un corpus des textes descriptifs (Moncla, 2015)

Rappel	96%
Précision	82.1%
Précision classification	<b>55.6%</b>
Précision frontières	79.5%

Evaluation pour les documents historiques (Mcdonough et al., 2018)

Rappel	55.5%
Précision	<b>75.7%</b>
F-mesure	64.1%

# Annotations - Campagnes d'évaluations

Il existe un grand nombre de typologies, dont une des plus répandues et générales est celle de la conférence **MUC 7** (Chinchor, 1998), divisée en trois catégories :

Type	Description
<b>ENAMEX</b>	Noms de personne, lieu, organisations, entreprises
<b>TIMEX</b>	Dates et heures
<b>NUMEX</b>	Expressions monétaires et pourcentages

**Deux campagnes d'évaluation pour la langue française:**

**ESTER2 (2008)** : traitement de transcriptions de parole radio ou télédiffusée

**ETAPE (Évaluations en Traitement Automatique de la Parole) 2011-2012** : avait pour objectif de mesurer les performances des technologies vocales appliquées à l'analyse des flux télévisés en langue française



# Typologies pour les toponymes - ETAPE

Exemple de la typologie toponymique du projet Quaero

le mont Ventoux => kind + name

```
<loc.phys.geo>  
  <kind>mont</kind>  
  <name>Ventoux</name>  
</loc.phys.geo>
```

	Location	
<i>Administrative</i> loc.adm.town loc.adm.reg loc.adm.nat loc.adm.sup	<i>Physical</i> loc.phys.geo loc.phys.hydro loc.phys.astro	<i>Facilities</i> loc-fac odonym loc.oro address loc.add.phys loc.add.elec

# Typologies pour les toponymes - ESTER2

Exemple de la typologie toponymique de la campagne Ester 2

La chaîne de montagnes des Alpes

[ent=loc.geo-] chaîne de montagnes des Alpes [-ent=loc.geo]

Location	
géographique naturel	loc.geo
région administrative	loc.admi
axe de circulation	loc.line
adresse	loc.addr
○ adresse postale	○ loc.addr.post
○ téléphone et fax	○ loc.addr.tel
○ adr. électronique	○ loc.addr.elec
construction humaine	loc.fac

- Manque de portabilité entre annotations des campagnes

# Annotations dans la TEI

Les lieux sont divisés en lieux **administratifs** (placeName), les lieux **géographiques** (geogName) et les lieux **relatifs** (relative place).

**placeName** : contient un nom de lieu

`<placeName type="abbaye bénédictine de tours">Saint Jullien</placeName>`

**geogName** : identifie un nom associé à une caractéristique géographique

`<geogName><geogFeat>rade</geogFeat> de Suez</geogName>`

**entité nommée relative** : lieu défini par sa situation en relation avec un autre lieu

`<placeName><measure>20 km</measure> <offset>à l'est de</offset> <settlement type="city">Paris</settlement></placeName>`

La typologie Quaero est compatible avec la TEI, le balisage s'effectue de manière quasiment similaire en utilisant seulement la norme TEI

# Difficultés pour la catégorisation

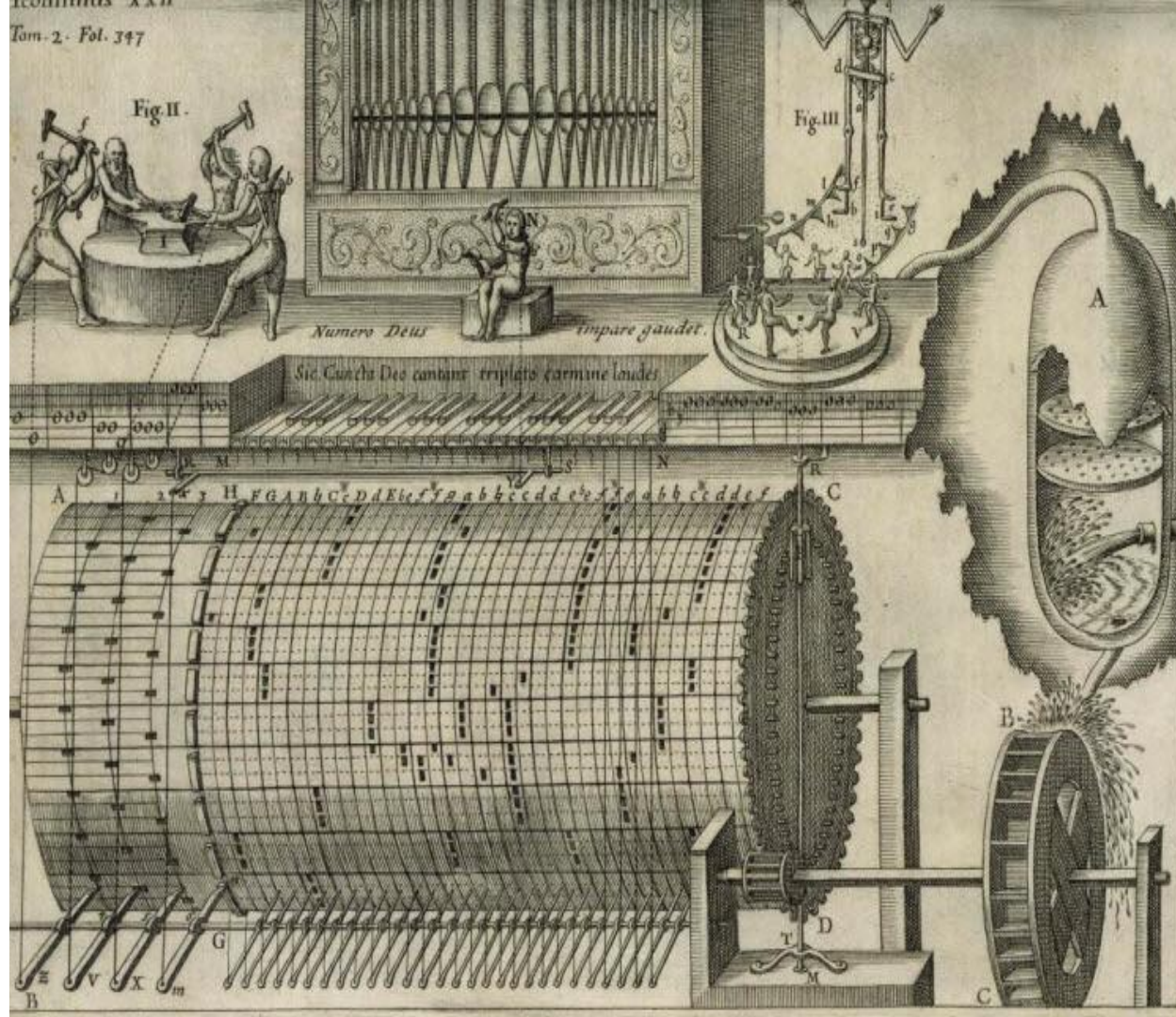
Phénomènes linguistiques qui génèrent de l'ambiguïté dans la reconnaissance d'entités nommées :

- Frontières lexicales
  - aborde le problème de **délimitation de l'entité nommée**  
*Grille royale du château de Versailles, la porte du Chapeau-Rouge*
- Facettes sémantiques (Cruse, 1999)
  - **polysémie**  
*Orange* : <Org>, <lieu>, <fruit>, <couleur>
  - **métonymie**
    - une organisation ou un nom de lieu à la place d'un ensemble d'individus  
*La France* à gagné en demi-finale.
    - un lieu d'origine à la place d'un produit  
Le *Coulommiers* est sans doute aussi ancien que le Brie.
  - **référents multiples**
    - *Ville-Neuve*

# Unitex

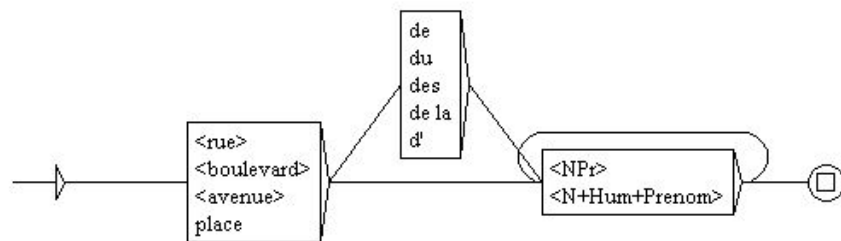
Ludovici Grignani, 1650. Orgue à cylindre :  
serinette actionnée par un système hydraulique, le  
mouvement actionnant aussi des automates (tome  
2, folio 347).

<https://images.bnf.fr/#/detail/689473/8>



# Unitex et les grammaires locales/automates

Unitex est un système de traitement de corpus, basé sur la **technologie des automates**. Le traitement fait appel à l'utilisation des ressources linguistiques.



<https://unitexgramlab.org/fr>

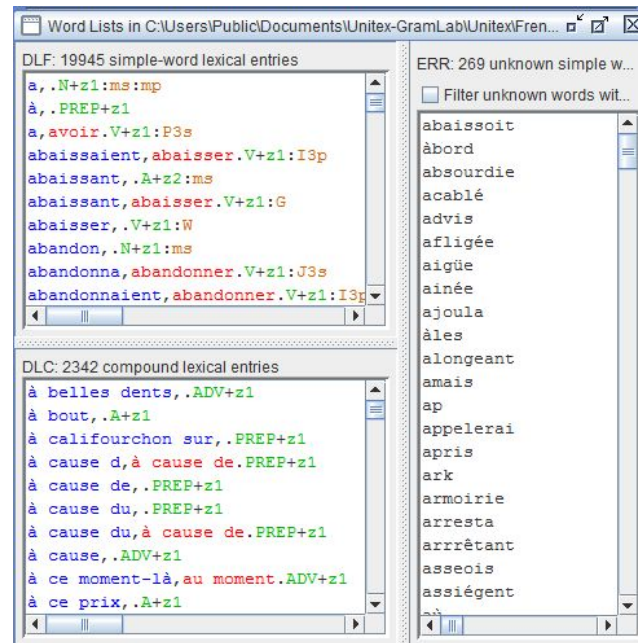
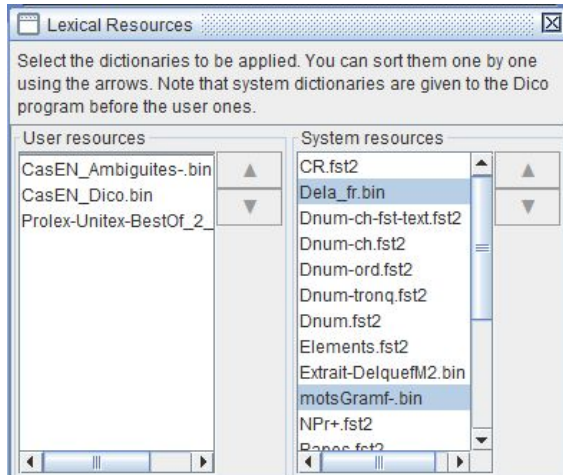
... aller d'ici au [boulevard Montmartre](#) !{S} (assait sur la [place de la Sorbonne](#), lors maison de la [rue Calade](#), qui lui appart !</p> <p>— La [rue de Catherine](#) ?...</p> <p> <p>— C'est la [rue de Catherine](#) !</p> <p> es une visite [rue de Rougemont](#), ma nour Elle habitait [rue de Rougemont](#) et nous y ne semble, ou [rue de Tournon](#), une vieille mes maisons, [rue de Valois](#), a pensé être parait de la [rue de Valois](#).</p> <p>Dire entrée de la [rue du Bouloy](#), un homme la TEUR,</p> <p><a href="#">RUE HAUTEFEUILLE, 30.</p> ère, prit la [rue Lafayette](#) et s'avança conduit à la [rue Richer](#).</p> <p>— C'est, partie de la [rue Richer](#), gagna la rue d i hôtel de la [rue Richer](#), sur un banc rus d> <p>Dans la [rue Richer](#), à Paris, il ex était située [rue Richer](#), pas loin de che ter cet hôtel [rue Richer](#), et sans avoir un beau jour, [rue Rougemont](#), et m'y lais ssait traîner [rue Rougemont](#), se montra se e n'était pas [rue Rougemont](#) que nous all:



# Unitex et les grammaires locales

Outil capable de gérer des ressources électroniques :

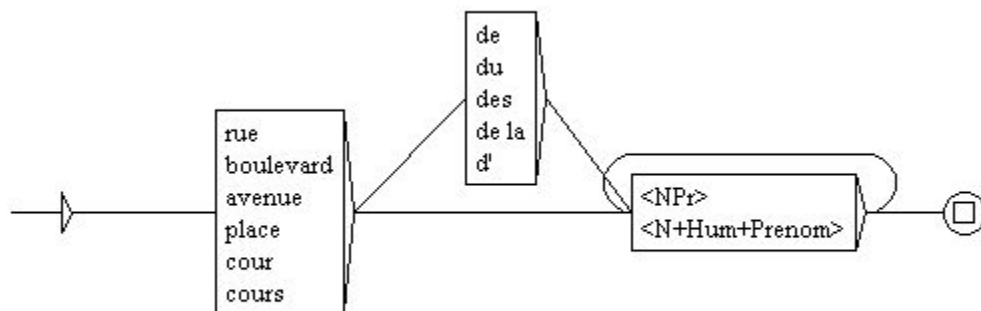
- dictionnaires électroniques
- grammaires locales/automates
- et les appliquer au niveau de la morphologie, du lexique et de la syntaxe



# Grammaires locales/automates

Les grammaires locales sont un formalisme de description (sous la forme de graphes) de règles syntaxiques ou sémantiques. Pour la phrase suivante :

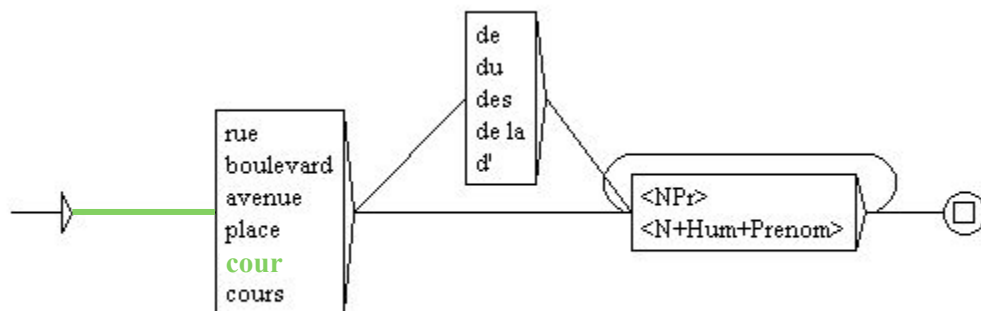
La *cour de Rome* est une voie du 8<sup>e</sup> arrondissement de Paris.



# Grammaires locales/automates

Les grammaires locales sont un formalisme de description (sous la forme de graphes) de règles syntaxiques ou sémantiques. Pour la phrase suivante :

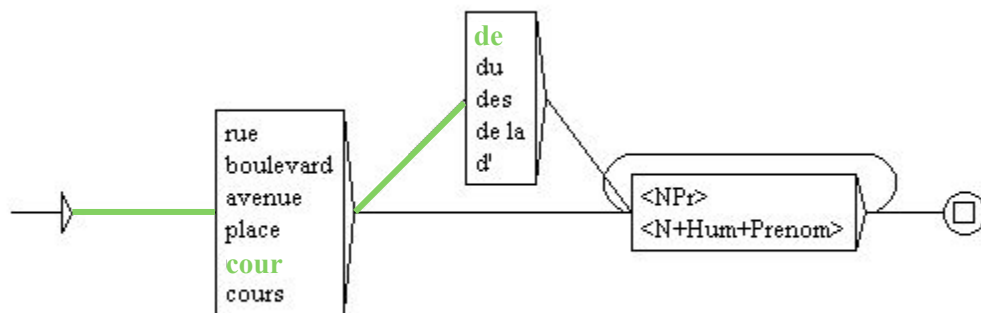
La *cour* de Rome est une voie du 8 arrondissement de Paris.



# Grammaires locales/automates

Les grammaires locales sont un formalisme de description (sous la forme de graphes) de règles syntaxiques ou sémantiques. Pour la phrase suivante :

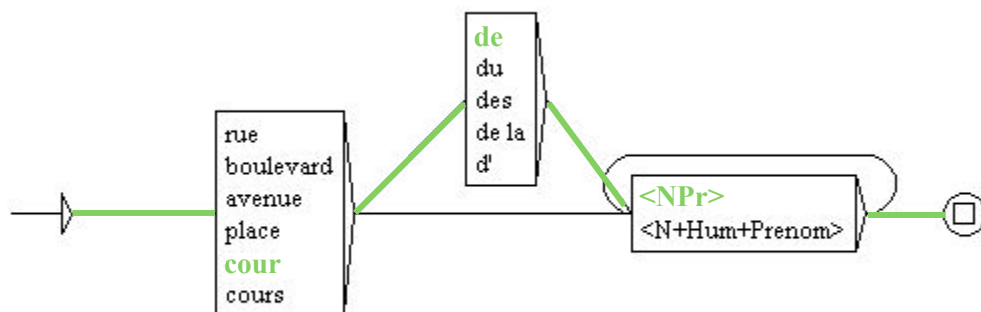
La cour de Rome est une voie du 8 arrondissement de Paris.



# Grammaires locales/automates

Les grammaires locales sont un formalisme de description (sous la forme de graphes) de règles syntaxiques ou sémantiques. Pour la phrase suivante :

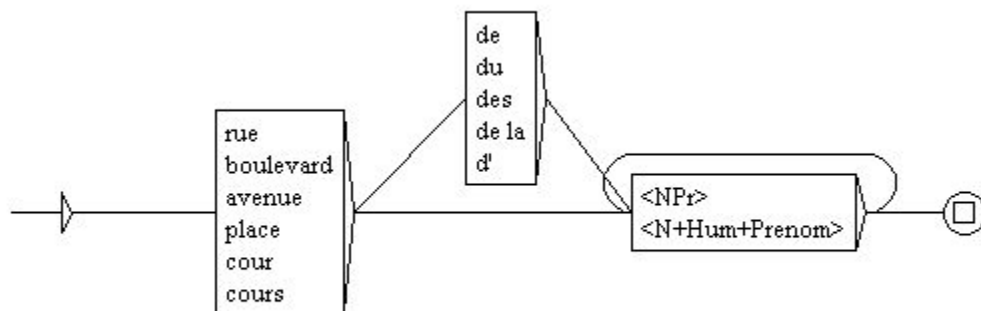
La *cour* *de* *Rome* est une voie du 8 arrondissement de Paris.



# Grammaires locales/automates

Les grammaires locales sont un formalisme de description (sous la forme de graphes) de règles syntaxiques ou sémantiques. Pour la phrase suivante :

Le *cours la Reine* est une voie du 8<sup>e</sup> arrondissement de Paris.

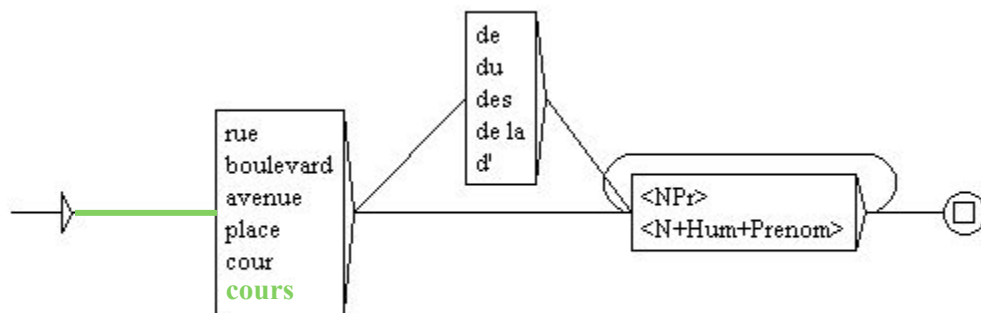




# Grammaires locales/automates

Les grammaires locales sont un formalisme de description (sous la forme de graphes) de règles syntaxiques ou sémantiques. Pour la phrase suivante :

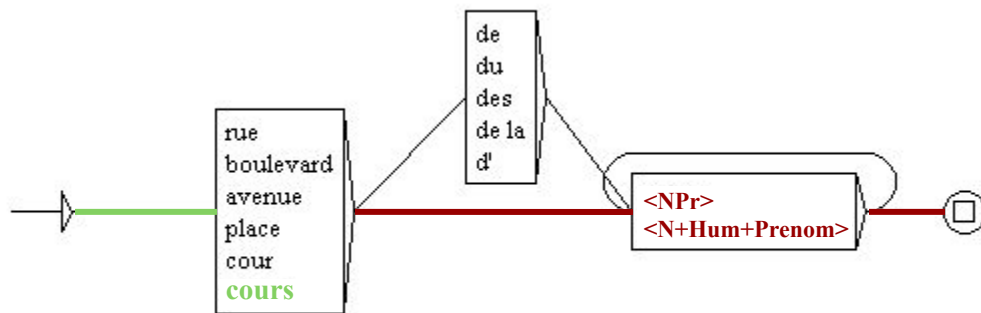
Le *cours* la Reine est une voie du 8<sup>e</sup> arrondissement de Paris.



# Grammaires locales/automates

Les grammaires locales sont un formalisme de description (sous la forme de graphes) de règles syntaxiques ou sémantiques. Pour la phrase suivante :

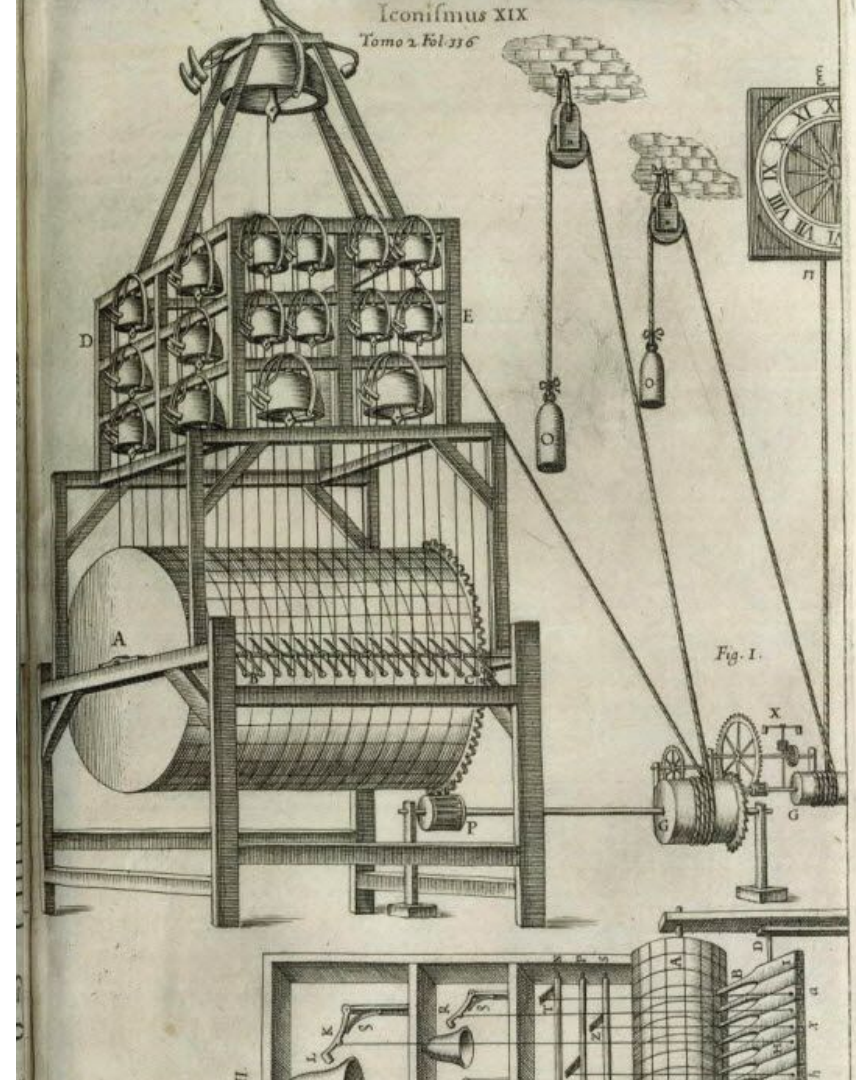
Le *cours* la *Reine* est une voie du 8 arrondissement de Paris.



# CasEN

Ludovici Grignani, 1650. Clochettes  
automatiques. Automaton campanarium  
fabricari (Liber IX : Magia consoni et dissoni)  
Musurgia universalis.

<https://images.bnf.fr/#/detail/728352>

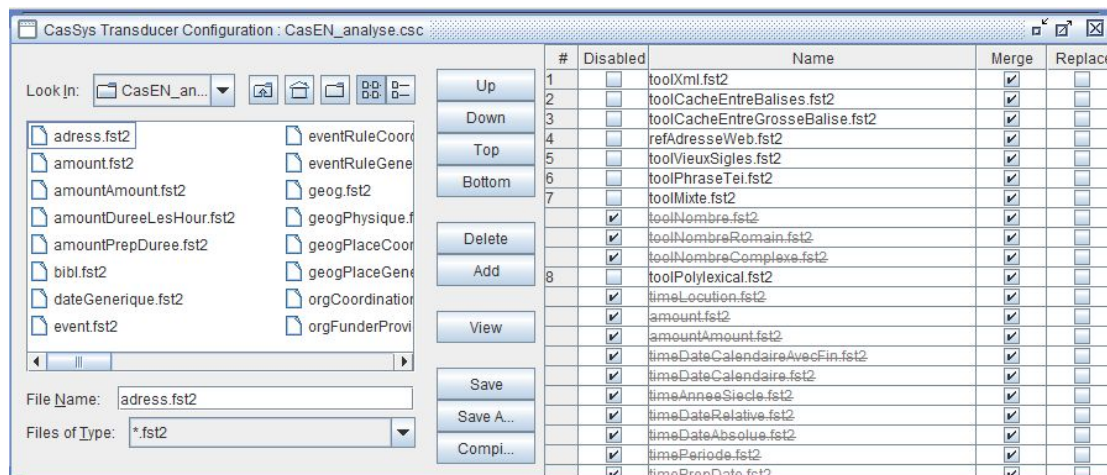


# CasEN (I)

CasEN est un système d'extraction d'entités nommées fondé sur CasSys, un module Unitex permettant l'utilisation de grammaires locales en cascade.

La cascade d'analyse CasEN utilise :

- des ressources lexicales et des descriptions locales de motifs,
- des transducteurs qui agissent sur le texte par des insertions, remplacements ou suppressions.



# CasEN (II)

Exemple de balisage avec CasEN :

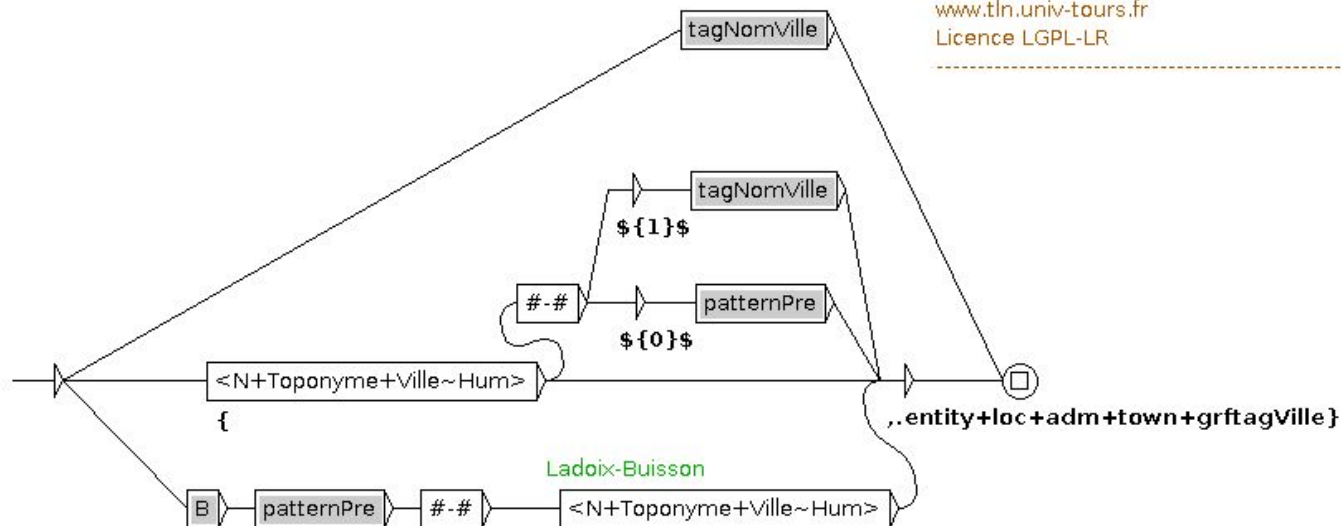
Nous gagnâmes la Bastide vis-à-vis de Bordeaux ; puis nous nous rembarquâmes sur la Garonne pour la traverser seulement, et mîmes pied à terre tout contre la porte du Chapeau-Rouge. (Catherine de La Guette, p. 184)

Résultat après l'application de la cascade *analyse.csc* :

```
id à {terre,.geogFeat+grfgeogPhysique} tout contre la porte
la {{\{porte\,.w\+type="noun"+quaero="kind"\} du Chapeau-Rouge\,.adrLine\+grftagAdresseRue\},.adress+grfadress}.
la {Garonne,.geogName+type="hydronym"+grfgeogPlaceDico} po
la {{\{Bastide\,.w\+type="noun"+quaero="kind"\},.place+type="building"+grfplaceBatiment}
de {Bordeaux,.placeName+type="settlement"+grftagVille} ; p
```

# CasEN (III)

- Graphe qui étiquette les toponymes *Ville*



CasEN

Auteurs : Nathalie Friburger, Denis Maurel

LI, Université François Rabelais Tours

{nathalie.friburger, denis.maurel}@univ-tours.fr

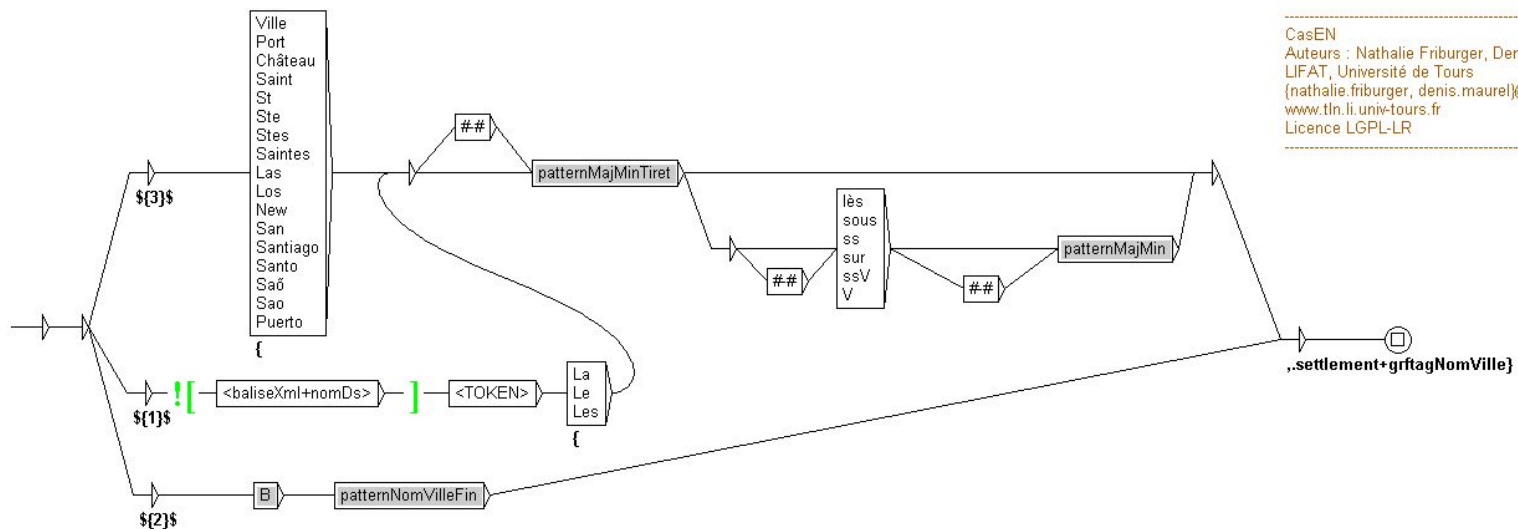
www.tln.univ-tours.fr

Licence LGPL-LR



# CasEN (IV)

Pour les entités toponymiques telles que : *Chateau Gontier*, *Chateau-Thierry*, *Chateau du Loir* le graphe ne parvient pas à reconnaître le mot **Chateau**



CasEN  
Auteurs : Nathalie Friburger, Denis Maurel  
LIFAT, Université de Tours  
{nathalie.friburger, denis.maurel}@univ-tours  
www.tln.li.univ-tours.fr  
Licence LGPL-LR

# La visualisation de la ville

*Langlois, Nicolas (1640-1703). L'église de  
Notre-Dame de Paris.*

<https://images.bnf.fr/#/detail/694938>



# La ville dans le corpus VisiAtrices

- L'exploration du nom commun *ville* a été effectuée à l'aide de l'outil TXM

The screenshot displays the TXM software interface. The top menu bar includes 'Fichier', 'Édition', 'Affichage', and 'Aide'. Below the menu is a toolbar with various icons. The left sidebar shows a project tree with folders like 'Corpus' and 'Fichier', and a list of files including 'CORPUS20012020', 'VISAUTRICESANSXML', and 'ville'. The main window is titled 'Requête' and shows search results for the word 'ville'. The results are organized into three columns: 'Contexte gauche', 'Pivot', and 'Contexte droit'. The 'Pivot' column lists the word 'ville' for each result. The 'Contexte gauche' and 'Contexte droit' columns show the surrounding text for each occurrence. The results are sorted by 'text\_id' and include tags like 'No-tags\_visiatrices\_01' and 'No-tags\_visiatrices\_02'.

text_id	Contexte gauche	Pivot	Contexte droit
No-tags_visiatrices_01	les Espagnols. Il y a dans la	ville	, les faubourgs et même les provinces, des endroits désignés par
No-tags_visiatrices_01	brûler avec eux dans l'incendie de cette	ville	plutôt que de les abandonner. Les tendres soins que ces oiseaux
No-tags_visiatrices_01	huile est extraite et déjà expédiée pour la	ville	la plus prochaine. Hérodote, qui visita l'Égypte près de
No-tags_visiatrices_01	quatre cent millions par année, et la	ville	de Gothembourg, en Suède, plus de sept cents millions.
No-tags_visiatrices_01	d'Augsbourg, en 1687, la seule	ville	d'Honfleur envoyait annuellement quarante vaisseaux à la pêche de la morue
No-tags_visiatrices_01	de l'Arabie Heureuse, près de la	ville	de Carisa, dans le golfe de Manaar, l'île de
No-tags_visiatrices_01	imiter les fourmis. Une fourmière est une	ville	en miniature, avec ses rues, ses carrefours, ses maisons
No-tags_visiatrices_01	et au midi. C'est aussi une	ville	sous le rapport de l'administration intérieure. Il y a des
No-tags_visiatrices_01	sanguinaire, arriver et s'établir dans la	ville	assiégée, y apportant à plusieurs reprises tout ce qui faisait sa
No-tags_visiatrices_01	qu'elle est commune aux environs de la	ville	de Tarente, est devenue célèbre par les fables dont son histoire
No-tags_visiatrices_02	, à dix kilomètres de la jolie petite	ville	de Chantilly. Cette cabane s'appelle la Maison de l'Aveugle
No-tags_visiatrices_02	petit Pierre, quand il va à la	ville	? — C'est pas ça non plus, grand'mère. —
No-tags_visiatrices_02	, choisis dans les meilleures familles de la	ville	de Saint-Denis, où logeait le roi de France avec son fils
No-tags_visiatrices_02	; elle était sombre et triste ; la	ville	dans ce temps n'était point éclairée ; heureusement que les torches
No-tags_visiatrices_02	d'armes et de ces flambeaux, la	ville	s'émeut : les portes s'ouvrent, les fenêtres se garnissent
No-tags_visiatrices_02	France ? — On ne réveillerait pas la	ville	pour cela, dit un vieux soldat. — C'est peut-être
No-tags_visiatrices_02	dit-il, nous cherchons en toute la	ville	notre bien-aimé prince royal Dagobert que le démon a fait sortir du

- Corpus issu du projet de recherche *VisiAtrices* composé par des textes de Charlotte Duplessis Mornay, Catherine Des Roches, Marguerite de Navarre, Marguerite de Valois, Jeanne-Marie Leprince de Beaumont et Anne Wiazemsky.

# Cooccurrences de la ville

En utilisant l'ensemble des dictionnaires distribués avec Unitex, nous avons obtenu la liste de cooccurents de la ville avec une **étiquette grammaticale** et un **trait sémantique**.

Nous distinguons ainsi 5 groupes de mots liés à la ville :

- les adjectifs qualificatifs
- les noms communs
- les noms de profession
- les verbes
- les toponymes

ambassadeurs	ambassadeur	N+z1+f3+Profession:fp
empereurs	empereur	N+z1+f3+Profession:fp
imprimeurs	imprimeur	N+z1+f3+Profession:fp
musiciens	musicien	N+z1+f3+Profession:fp
sergents	sergent	N+z1+f3+Profession:fp
cadet		N+z1+f3+Profession:fs
marchand		N+z1+f3+Profession:fs
pâtissier		N+z1+f3+Profession:fs
sergent		N+z1+f3+Profession:fs

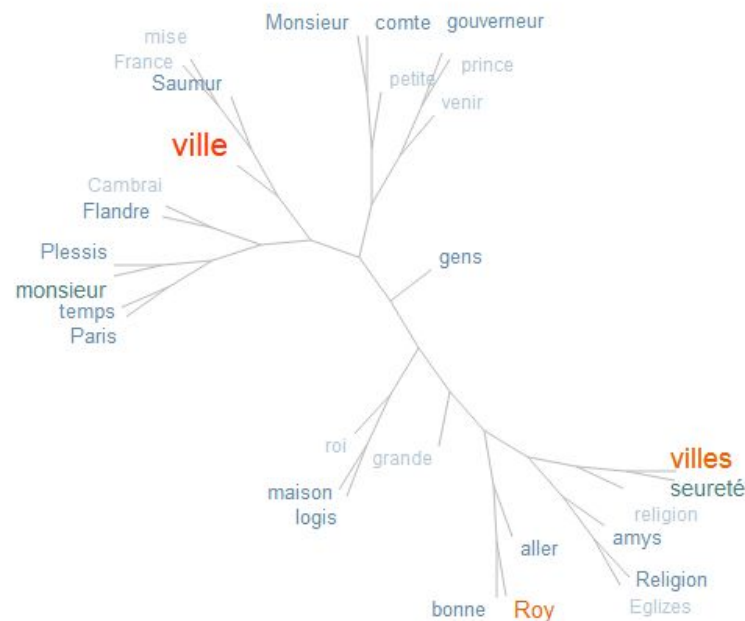
# Catégorisation de cooccurrents

adjectifs qualificatifs	noms communs	noms de profession	verbes	toponymes
belle, célèbre, grande, lointaine, jolie, petite, voisine, natale, magnifique, bâtie, peuplée	brasserie, maison, muraille, rue, capitale, place, porte, digue, lieue, sommité, enceinte, entrée, hôtel chateau, chemin, cloches, édifices, bruit, jardins, magasin, port, police, statue, luxe, théâtre, quartier, boulevard, temple, fondement, incendie, département	ambassadeur, empereur, imprimeur, musicien, sergent, pâtissier, jurisconsulte	habiter, défendre, présider, combattre, figurer, forcer, ravager arriver, quitter, atteindre, emparer, entrer, monter, ouvrir, renverser, acheter, construire, envoyer, étendre, dévorer, bâtir, dîner	Francfort, Grenade, Olympie, Angers, Anvers, Athènes, Babylone, Bayreuth, Bergen, Bordeaux, Bruxelles, Corinthe, Édimbourg, Épidaure, Francfort, Fribourg, Genève ...

# Visualisation de la ville avec TreeCloud

Visualisation du nuage arboré du mot **ville** et **villes**  
à partir du texte de Mémoires de Mme De La Guette,  
Mémoires de Marguerite De Valois, Mémoires de  
Charlotte Duplessis Mornay.

- les nuages de mots disposés autour d'un arbre indiquent leur proximité dans le texte





# Visualisation de l'évolution de fréquences I

- *VisuLexique*, un script de visualisation de fréquences de groupes de mots dans des textes

Trouver les mots dans le texte !

### Paramètres

Découper le texte en  blocs.

☒ Chercher les mots ci-dessous en minuscules, dans le texte converti en minuscules.

Première liste des mots à chercher dans le texte (un mot par ligne, utiliser \_ dans la liste ci-dessous, et dans le texte, pour les mots composés) :

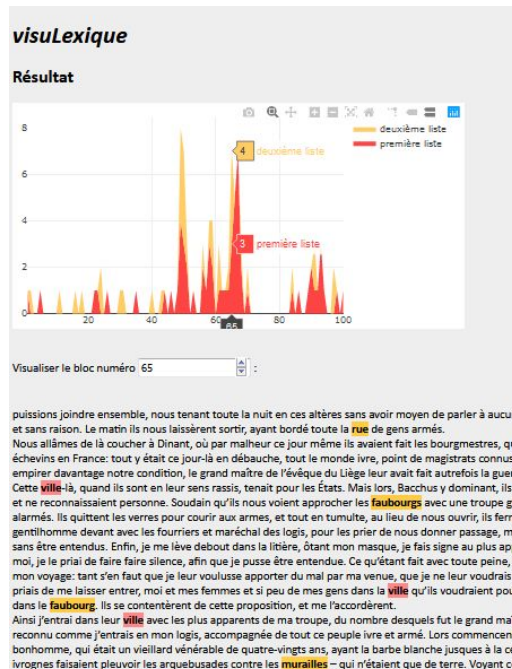
citadines  
métropole  
métropoles  
urbain  
urbaine  
urbains

Deuxième liste des mots à chercher dans le texte (un mot par ligne, utiliser \_ dans la liste ci-dessous, et dans le texte, pour les mots composés) :

immeuble  
immeubles  
appartement  
appartements  
citadelle  
citadelles

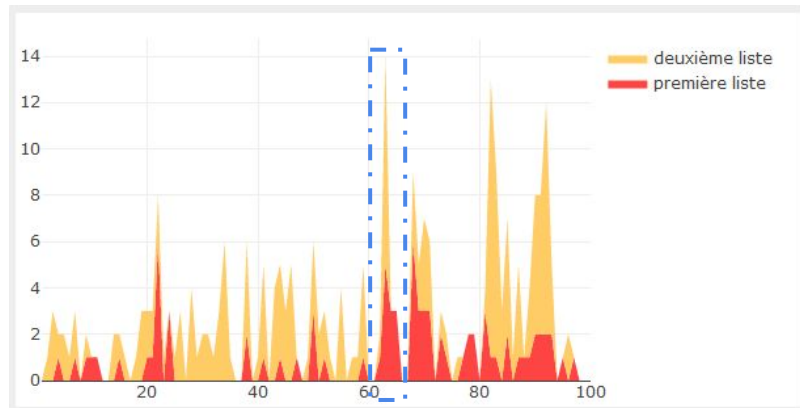
Texte à explorer :

Mémoires de Marguerite de Valois  
Je louerais davantage votre œuvre, si elle ne me  
louait tant, ne voulant qu'on attribue la louange  
que j'en ferais plutôt à la philautie qu'à la  
raison, ni que l'on pense que, comme Thémistocle,



# Visualisation de l'évolution de fréquences II

Visualisation de l'évolution des termes liées à ville contenus au livre d'Emma Morel "Mythologie épurée à l'usage des maisons d'éducation pour les deux sexes".



l'oracle de Delphes, qui lui conseilla de s'arrêter là où le conduirait un bœuf qu'il devait rencontrer à la **porte** du **temple**, et de fonder une **ville** en cet endroit. Cadmus trouva en effet un bœuf qui se mit à marcher devant lui, et il le suivit jusque dans une contrée à laquelle il donna le nom de Béotie.

Avant de s'occuper de la **ville** qu'il devait construire, il voulut offrir un sacrifice aux Dieux et envoya ses compagnons puiser de l'eau à une fontaine consacrée au dieu Mars ; mais ils furent tous dévorés par un énorme dragon ; Cadmus les vengea, et, après avoir tué le monstre, il s'empara de ses dents, suivant le conseil de Minerve, et les sema sur la terre. Ces dents donnèrent aussitôt naissance à des hommes armés qui combattirent les uns contre les autres, et s'entre-tuèrent, à l'exception de cinq qui aidèrent à Cadmus à jeter les **fondements** d'une **ville** qu'il nomma Thèbes. Thèbes, située au pied du lac Capais, jouit, au temps de Pélopidas et d'Épaminondas, d'une grande puissance et d'une prospérité remarquable. Elle fut ruinée, vers l'an 419 de Rome, par Alexandre, qui n'épargna que la **maison** où le poète Pindare

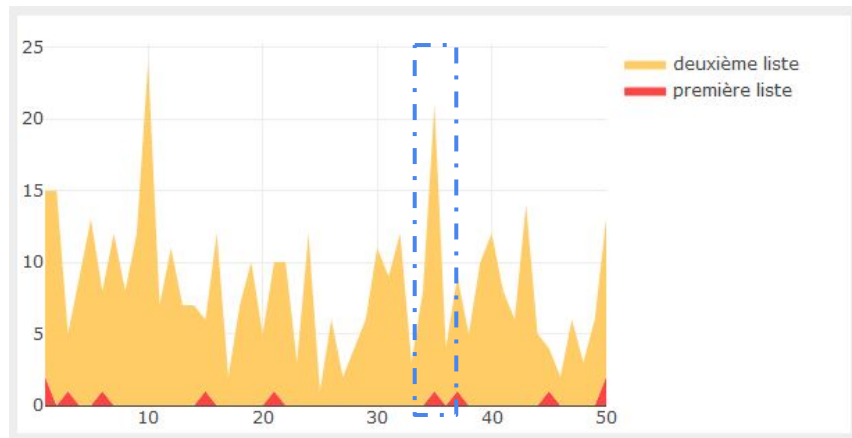
était né. Sa **citadelle** se nommait Cadmée. Un village nommé Thiva occupe aujourd'hui sa place ; quelques inscriptions sont les seuls vestiges des magnifiques monuments qui la décorèrent autrefois. Il y eut en Égypte une autre Thèbes, qui fut fondée par Busiris, laquelle fut également célèbre par ses richesses et par les cent **portes** ouvertes dans ses **murailles**. Plus tard, Amphion éleva les **murailles** de cette **ville**, et il savait tirer de sa lyre des sons si ravissants, que les pierres, sensibles à cette harmonie, venaient d'elles-mêmes se ranger à leur **place**. Niobé, femme d'Amphion et fille de Tantale, ayant eu quatorze enfants, se mettait orgueilleusement au-dessus de Latone, mère d'Apollon et de Diane. Cette déesse en fut tellement irritée, qu'elle fit tuer par ses enfants les sept fils et les cinq filles de Niobé ; cette dernière fut métamorphosée en rocher.

Cadmus s'occupa ensuite de polir les mœurs des sauvages habitants des campagnes qu'il avait réunis dans sa **cité**. Il les assujettit à des lois pleines de sagesse, et leur enseigna des arts utiles. On lui attribue



# Visualisation de l'évolution de fréquences III

Visualisation de l'évolution des termes liées à ville contenus au livre de Suzanne Duval "L'Agente".



Une autre fois la promenade l'avait emmenée plus loin. Elle avait commencé par contourner la colline et s'était bientôt retrouvée dans les **rues** élégantes des alentours du cimetière de Montmartre, déboucha sur l'**avenue** de Saint-Ouen. S'arrêtant, elle hésitait, s'appuyait sur une jambe puis l'autre, essayant de retrouver la direction du Nord sans oser demander son chemin aux passants. Les voitures défilaient. Elle repartit au hasard, accélérant le pas, emprunta une **rue** encombrée et tomba sur l'**avenue** de Clichy, sans comprendre la direction qu'elle avait prise. Elle marcha un temps sur le même trottoir, passant devant deux restaurants turcs et un dépôt-vente ; regarda un moment une rangée de chaussures usées, essayait de retrouver ses esprits. Elle respirait de nouveau normalement, marcha à une allure plus lente. Elle venait d'emprunter une **rue** pavée. De jolies maisons d'un à deux étages se succédaient, chacune ornée d'un jardinet, et elle était presque sûre qu'elle avait déjà visité l'une d'entre elles, autrefois ; s'amusa à regarder les noms sur les boîtes aux lettres. Elle finit par retrouver un tronçon de la petite ceinture, à une centaine de mètres du **boulevard** périphérique ; reconnaissait le mur concave de cette **cité** monumentale, en brique blanche, chaque fenêtre encadrée d'une mosaïque de fleurs vertes. On était tout près de la porte Pouchet, elle n'avait plus qu'à longer les rails vers l'Est pour rentrer, elle serait chez elle dans une dizaine de minutes. Il faisait nuit, elle pressait le pas et eut envie, arrivée devant chez elle, de pousser jusqu'au café de l'ancienne **gare**, d'appeler Chafik. Renonçant, finalement, hésitant à nouveau, immobile dans le hall d'entrée. Elle regarda son téléphone. Paul lui avait encore écrit. Il était passé à Central immobilier pour la saluer, avait été déçu de ne pas l'y trouver. Si elle était à Paris cette semaine, on pourrait peut-être se revoir ? Elle décida de lui répondre : OK, quand ? Le rendez-vous avait été fixé quelques jours plus tard, sur la **place** de la Sorbonne. Elle descendit au **métro** Saint-Michel, évita les **boulevards**, escaladant la montagne Sainte-Genève par la **rue** Saint-Jacques. Devant l'imposante bâtisse du **lycée** Louis-Le-Grand, des groupes d'adolescents prenaient leur pause entre deux cours. Certains, peu nombreux, fumaient. Par la vitre grillagée, on voyait le hall, des

86 lycéens qui circulaient entre les colonnes de marbre, la cour d'honneur, le bassin entre les rosiers sans fleurs. Judith avait été élève ici, autrefois. Il était arrivé qu'elle vienne la chercher, à la sortie des cours, attendant sur le même trottoir, guettant son arrivée, craignant le regard des élèves du même âge qu'elle. Judith finissait par apparaître, se hâtant, le pas festif, pour la rejoindre et l'embrasser. Elles descendaient ensuite en direction des quais, s'asseyaient côte à côte au bord du fleuve, entre les crottes de pigeon, sur les pavés inconfortables. Judith fumait ses premières cigarettes.

# Références

- Bouvier, J. (1999). Odonymes d'agglomération entre l'écrit et l'oral. *Nouvelle Revue d'Onomastique*, n° 33-34, 303–310.
- Chinchor N. (1998), MUC-7 Named Entity Task Definition (version 3.5), in *Proceedings of the 7th Message Understanding Conference (MUC-7)*, 19 April-1 May 1998, Fairfax, VA.
- Cruse, D. A., 1999. *Meaning in language. An introduction to semantics and pragmatics*, Oxford University Press, Oxford. Deuxième édition : 2004.
- Ehrmann, M. (2008). *Les entités nommées, de la linguistique au TAL : Statut théorique et méthodes de désambiguïsation*. Université PARIS 7 - DENIS DIDEROT.
- Friburger, N. (2002). *Reconnaissance automatique des noms propres; application à la classification automatique de textes journalistiques*. Thesis. Université de Tours.
- Friburger, N., & Maurel, D. (2004). Finite-state transducer cascades to extract named entities in texts. *Theoretical Computer Science*, 313(1), 93-104.
- Martinez, C. (2017). *Grammaires locales étendues : principes, mise en œuvre et applications pour l'extraction de l'information*. Thesis. Université Paris-Est.
- McDonald, D. (1996). Internal and external evidence in the identification and semantic categorization of proper names. *Corpus Processing for Lexical Acquisition*, (Cambridge, MA: MIT Press).
- Mcdonough, K., & Moncla, L. (2018). Expérimentation de méthodes d'extraction d'informations géographiques pour les documents historiques. *Atelier Humanités Numériques Spatialisées (HumaNS'2018)-SAGEO 2018.*, 1–7.
- Moncla, L. (2015). *Automatic reconstruction of itineraries from descriptive texts*. Université de Pau et des Pays de l'Adour; Universidad de Zaragoza.
- Nouvel, D. (2012). Reconnaissance des entités nommées par exploration de règles d'annotation - Interpréter les marqueurs d'annotation comme instructions de structuration locale. Apprentissage . Université François Rabelais - Tours, 2012.