# Proposition de plan pour le document final

# Structures dans les titres scientifiques

## Introduction Sujet : les titres scientifiques

## Problématique : peut-on trouver des structures lexico-syntaxique récurrentes dans les titres scientifiques ? Comment peut-on les caractériser dans leur nature et leur usage ?

## 1. État de l’art

1.1 Problématiques étudiées

1.2 Corpus utilisés

1.3 Caractéristiques des titres

1.4 Typologies des titres

## 2. Corpus

2.1 Présentation de HAL et extraction des données

2.2 Chaîne de traitement des données

Filtrage (titres en anglais, titres vides)

Étiquetage et lemmatisation

2.3 Exemples de titres

Titres « modèles »

2.4 Outillage

2.5 Vérifications

Retrouve-t-on le fait que plus un titre a d’auteurs, plus son titre est long ?

## 3. Structures

3.1 Définition

3.2 Axe de variabilité

Éléments optionnels dans certains syntagmes (ADJ dans syntagme nominal SN)

Éléments variables (Préposition ou Préposition + Déterminant)

3.3 Patrons

3.3.1 But et limites

*Pas de répétabilité infinie d’éléments dans le patron*

3.3.1 Langage de définition

3.4 Points d’ancrage

Où cherche-t-on des structures : après « : », début de titre.

## 4. Étude de structures après « : »

4.1 Structure SN + P + SN  
 4.1.1 Description

4.1.2 Exemples

4.1.3 Patron  
4.1.4 Statistiques

Couverture (du corpus, du titre)

Fréquences

Répartition par domaine

4.2 Structure P + SN  
 4.3 Structure V + SN

## 5. Discussions et perspectives

4.1 Discussions

4.1.1 Défauts remarqués

Talismane : ADJ pris comme NC

Outil propre : « étude de cas » vu comme 3 suites N1 P N2 différentes,

Heureusement, 2 sont très marginales.

4.1.2 Ouverture sémantique

*Sur le plan sémantique, que peut-on dire à ce stade ?*

5.1 Perspectives

5.1.1 Couverture totale du corpus

5.1.2 Couverture totale des titre

5.1.3 Émergence automatique de patrons

## Conclusion

## Table des requêtes et résultats

## Table des tableaux

## Bibliographie