14 ° Colloque International
27 - 28 - 29 Aout 2019 - Lyon
Association Internationale pour la Recherche en Didactique du Français
Les concepts dans la recherche en didactique du français
Émergence et création d'un champ épistémique

Les corpus numériques pour l'aide à l'écriture académique

CRISTELLE CAVALLA

UNIVERSITÉ SORBONNE NOUVELLE PARIS 3
Membre de l'Université Sorbonne Paris Cité
DILTEC - EA 2288
Didactique des langues, des textes et des cultures

Article publié en version courte dans la Lettre de l'AIRDF :
Cavalla C. (2018). Exemple d'enseignement de la phraséologie transdisciplinaire à l'aide de corpus numériques en FLE. *La lettre de l'AIRDF*, Association internationale de recherche en didactique du français, La didactique du lexique, pp.43-47. ⟨http://airdf.ouvaton.org/⟩. ⟨hal-02087791⟩

Article publié en version longue :
Cavalla C. (2018),Lexique transdisciplinaire et enseignement aux étudiants allophones. A. Tutin et M-P. Jacques (Eds), *Lexique transversal et formules discursives des sciences humaines*, ISTE Editions, pp.191-214, ⟨https://iste-editions.fr/products/lexique-transversal-et-formules-discursives-des-sciences-humaines⟩. ⟨hal-01917082⟩

Présentation

1. **Hypothèse et objectifs**

2. **Outils d'analyses et pédagogiques**
   - Outil pour séquence
   - Type de corpus
   - Analyses linguistiques
   - Approche pédagogique

3. **Séquence didactique**
   - Contexte d'enseignement
   - Introduction des corpus
   - Introduction de la phraséologie

4. **Conclusion**

UNIVERSITÉ SORBONNE NOUVELLE PARIS 3
Membre de l'Université Sorbonne Paris Cité
DILTEC - EA 2288
Didactique des langues, des textes et des cultures

Objectifs généraux et hypothèse

OBJECTIFS LINGUISTIQUES DES APPRENANTS ?		OBJECTIFS PRAGMATIQUES DES APPRENANTS ?
Linguistique de corpus	⟷	Utilisation du corpus en classe de langue
Genre « écrit académique »	⟷	Appropriation des spécificités des écrits scientifiques

Discours académique
- Normes langagières du genre
- Normes professionnelles

Hypothèse didactique (affect) **:**

L'entrée par le corpus numérique aide à aborder la langue de façon sereine

Objectifs pour l'apprenant

Être capable d'utiliser des corpus numériques pour rédiger

+

Être capable d'utiliser les éléments linguistiques appropriés à cet écrit

→

Être capable de rédiger un écrit académique

Présentation

1. **Hypothèse et objectifs**

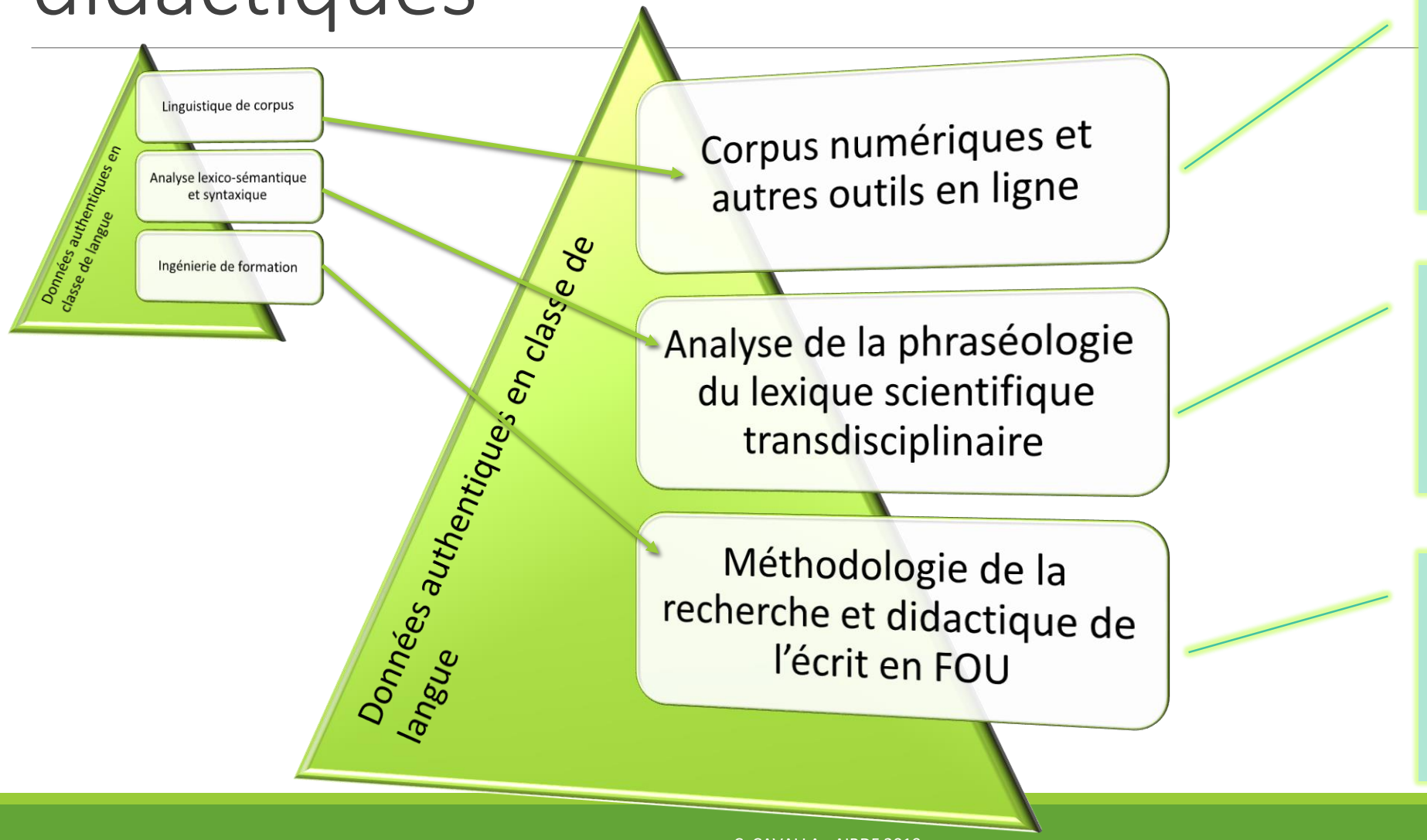
2. **Outils d'analyses et pédagogiques**
   - Outils pour séquence
   - Type de corpus
   - Analyses linguistiques
   - Approche pédagogique

3. **Séquence didactique**
   - Contexte d'enseignement
   - Introduction des corpus
   - Introduction de la phraséologie

4. **Conclusion**

UNIVERSITÉ SORBONNE NOUVELLE PARIS 3
Membre de l'Université Sorbonne Paris Cité
DILTEC - EA 2288
Didactique des langues, des textes et des cultures

Outils pour l'élaboration de séquences didactiques



Bureautique
Abord de nouveaux logiciels (comme app. tél)

Langue immédiatement réinvestie

Méthode de travail pour avenir pro

Corpus utilisés : le « déjà-là » des étudiants

Données authentiques en classe de langue

- Corpus numériques et autres outils en ligne
- Analyse de la phraséologie du lexique scientifique transdisciplinaire
- Méthodologie de la recherche et didactique de l'écrit en FOU

Outils en ligne

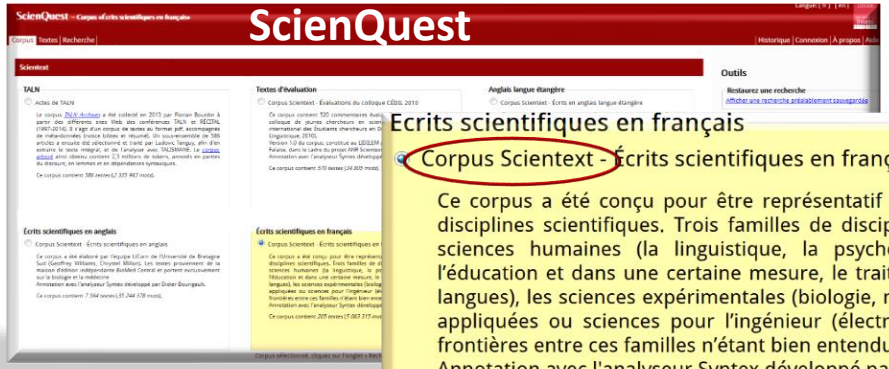
Google

Google Books Ngram Viewer

CORPORA COLLECTION
LEIPZIG UNIVERSITY

Linguee

Lexicoscope
- Exploration des profils combinatoires -



Ecrits scientifiques en français

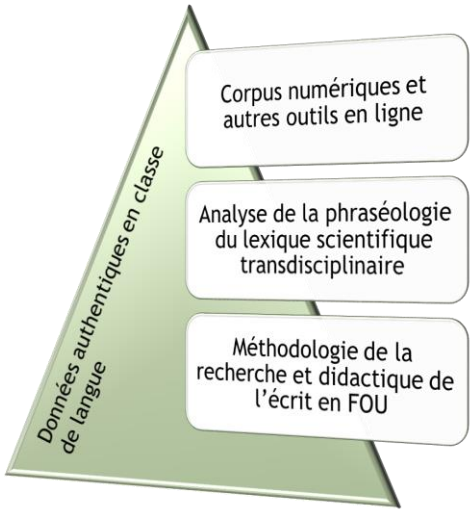
Corpus Scientext - Écrits scientifiques en français

Ce corpus a été conçu pour être représentatif des différents genres et disciplines scientifiques. Trois familles de disciplines sont incluses : les sciences humaines (la linguistique, la psychologie, les sciences de l'éducation et dans une certaine mesure, le traitement automatique des langues), les sciences expérimentales (biologie, médecine) et les sciences appliquées ou sciences pour l'ingénieur (électronique, mécanique), les frontières entre ces familles n'étant bien entendu pas étanches.
Annotation avec l'analyseur Syntex développé par Didier Bourigault.

Ce corpus contient *205 textes* (*5 063 315 mots*).

Compleat Lexical Tutor

Analyses linguistiques utilisées

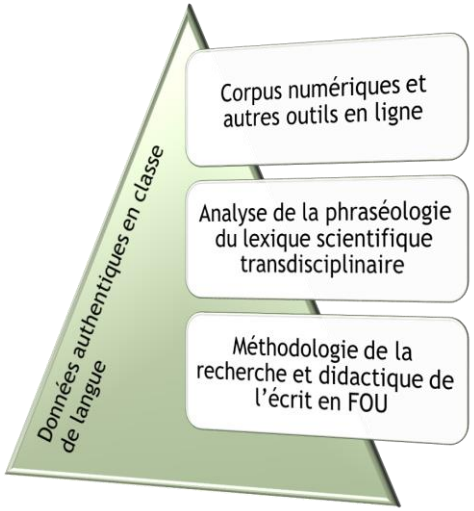


Approche par le lexique

Étude de la **phraséologie transdisciplinaire**

- Définition (Tutin et Pecman)
- Extraction
- Analyses sémantique et syntaxique
- Approche Sens-Texte (Mel'cŭk, Polguère)

Approche pédagogique



Approche directe / indirecte

Objectif : autonomie

- Directe : leurs questions
  - *comment énoncer ses questions de recherche / ses hypothèses ? Comment présenter son sujet ?…*
- Approche inductive : leurs besoins
  - *Prise de conscience du contenu de l'écrit académique*
  - *Organisation du contenu*
  - *Recherche des éléments pour énoncer le contenu*
- Onomasiologie : vers autonomie
  - *Entrées sémantiques*

Approche directe

Objectif : spécifier un élément linguistique

- Indirecte : arrêt sur image
  - *concordancier sur un seul élément ou sur une classe sémantique*
- Comprendre l'utilisation : autonomie de la méthode d'interrogation épistémologique (comment fonctionne cette langue ?)

Approche indirecte

Lien avec l'hypothèse

Hypothèse didactique (affect) :

L'entrée par le corpus numérique aide à aborder la langue de façon sereine

Objectif : autonomie

- Directe : leurs questions
  - *comment énoncer ses questions de recherche / ses hypothèses ? Comment présenter son sujet ?…*
- Approche inductive : leurs besoins
  - *Prise de conscience du contenu de l'écrit académique*
  - *Organisation du contenu*
  - *Recherche des éléments pour énoncer le contenu*
- Onomasiologie : vers autonomie
  - *Entrées sémantiques*

Approche directe

1. Découverte individuelle : laisser le temps de la prise de conscience par manipulation
2. Poser ses propres questions en fonction de ses propres besoins
3. Voir que d'autres se posent les mêmes questions et aident à y répondre

Présentation

1. **Hypothèse et objectifs**

2. **Outils d'analyses et pédagogiques**
   - Outils pour séquence
   - Type de corpus
   - Analyses linguistiques
   - Approche pédagogique

3. **Séquence didactique**
   - Contexte d'enseignement
   - Introduction des corpus
   - Introduction de la phraséologie

4. **Conclusion**

Public – Formation

Futurs journalistes, traducteurs, enseignants, juristes dans le domaine culturel, médiateur interculturel…

Master 2e année – Niveau B2-C1 - Mémoire à rédiger en français

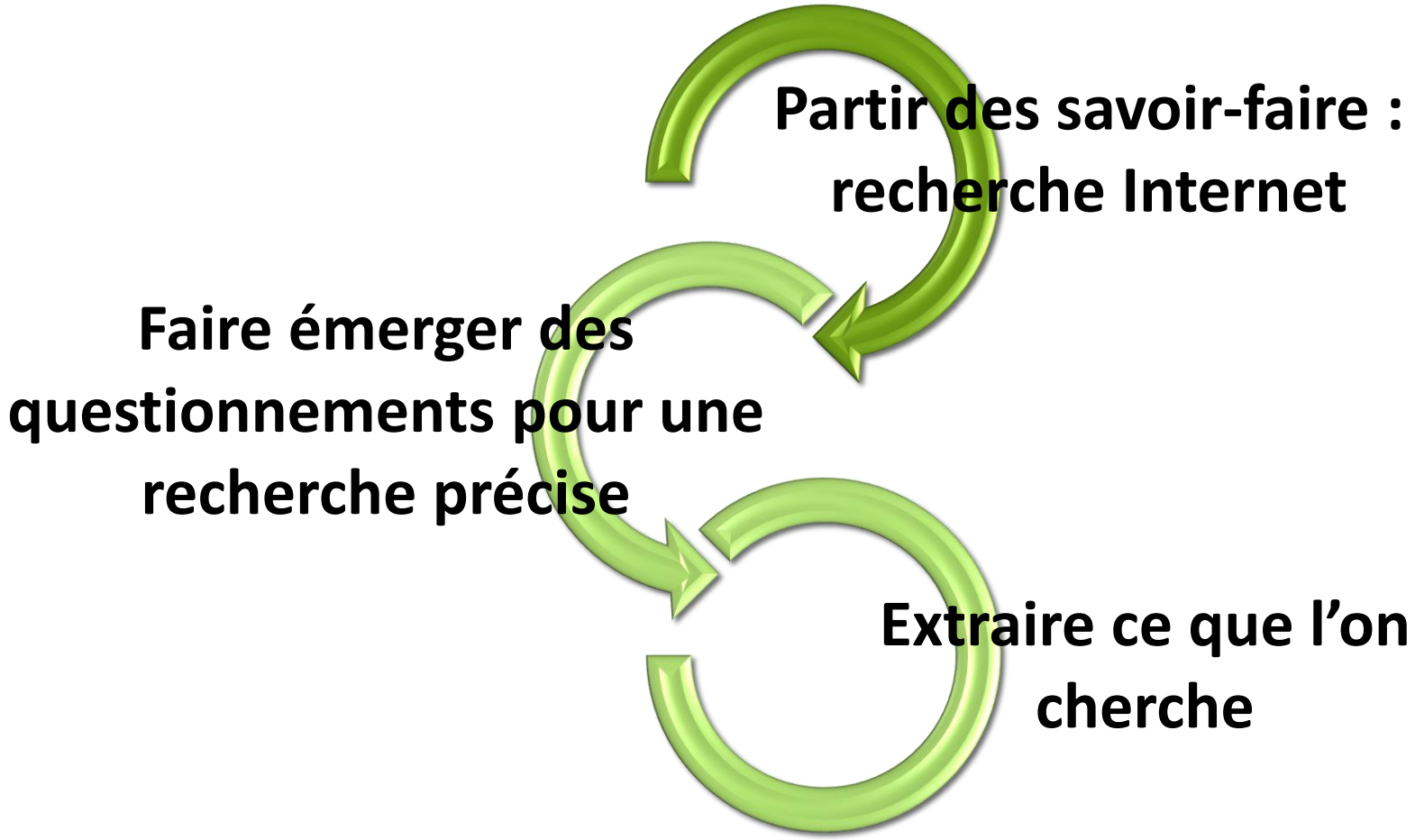
L2 : anglais, français, allemand, néerlandais, espagnol, portugais, russe, chinois…

L3 : anglais, français, allemand, néerlandais, espagnol, portugais, russe, chinois…

L1 : néerlandais, allemand, arabe, espagnol

L4 : anglais, français, allemand, néerlandais, espagnol, portugais, russe, chinois…

Comment faire découvrir ?



Kübler, N. (2014). Mettre en œuvre la linguistique de corpus à l'université - Vers une compétence utile pour l'enseignement/apprentissage des langues ? *Les Cahiers de l'Acedle, 11*(1), 37-77.

Approche Directe : recherche sur outils connus

Google

Qwant

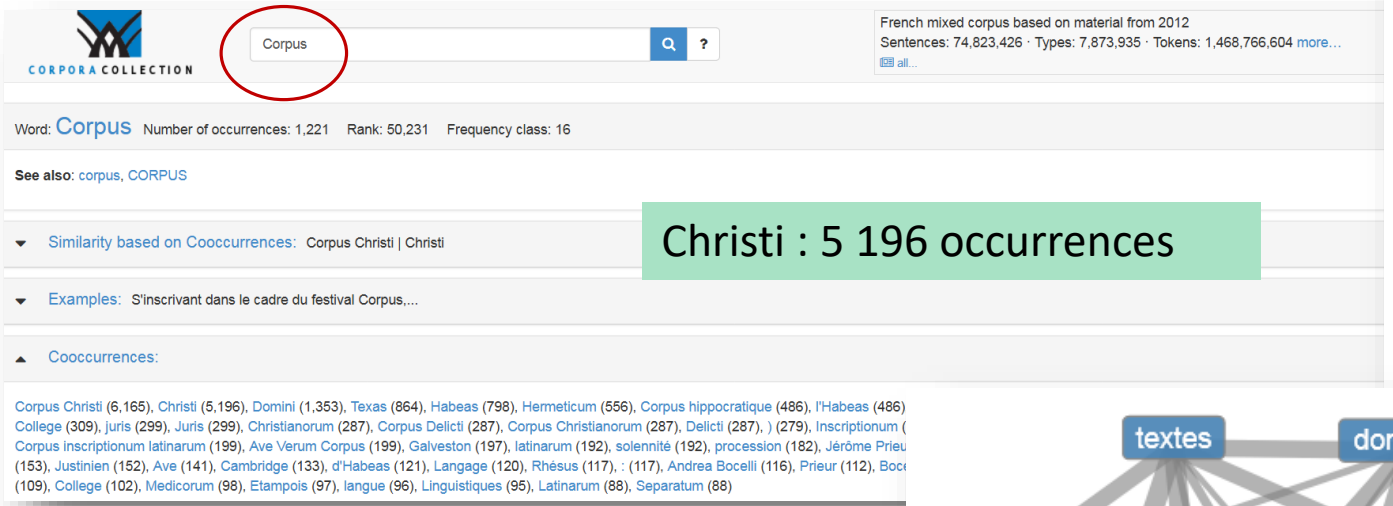
Linguee

Dans quel type de phrase trouve-t-on l'expression « en particulier » ?

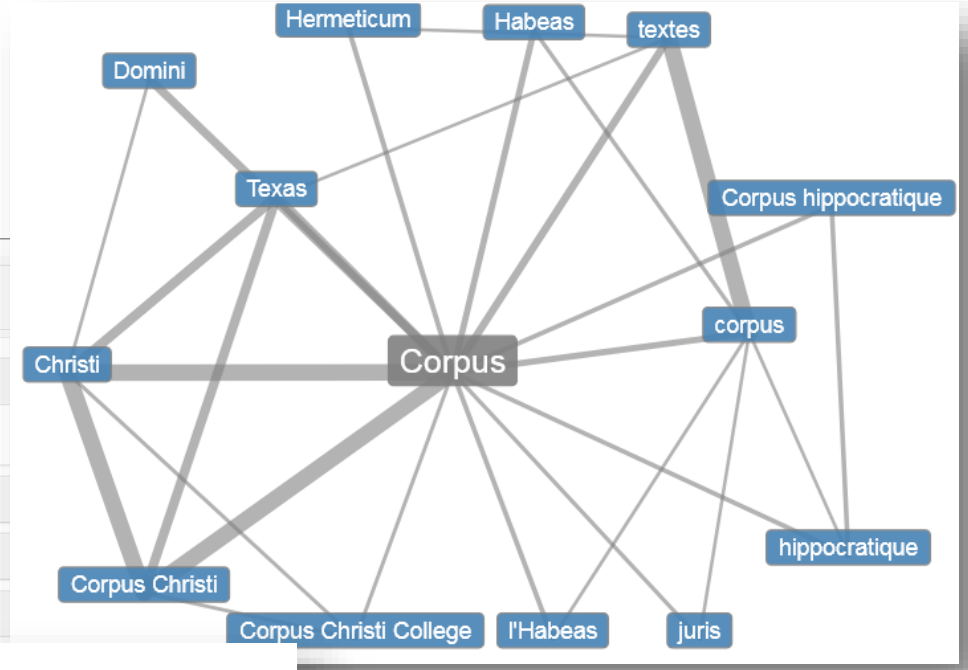
Comment sélectionner les exemples trouvés dans Google ?
Comment vérifier les contextes ?
…

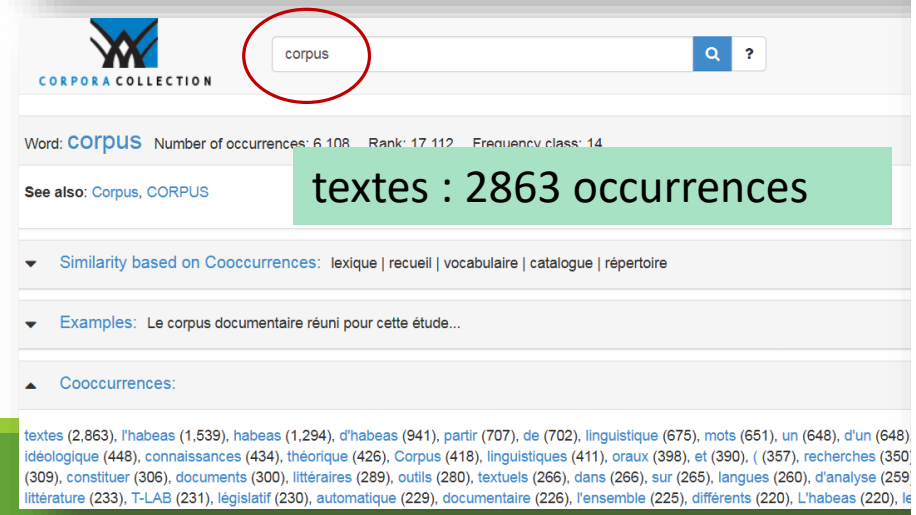
L'expression « en particulier » est utilisée dans quel contexte et par qui ?

Approche Directe
1er corpus : des chiffres

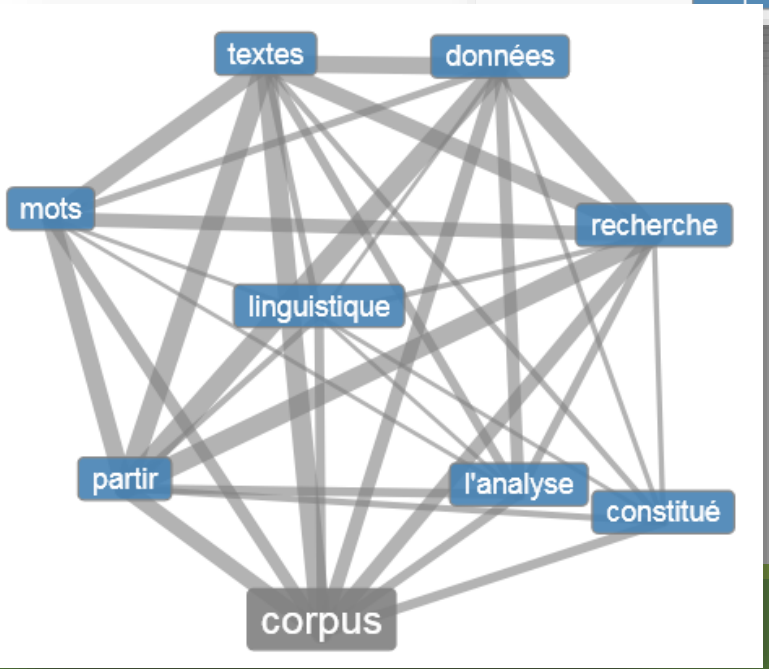


Christi : 5 196 occurrences





textes : 2863 occurrences

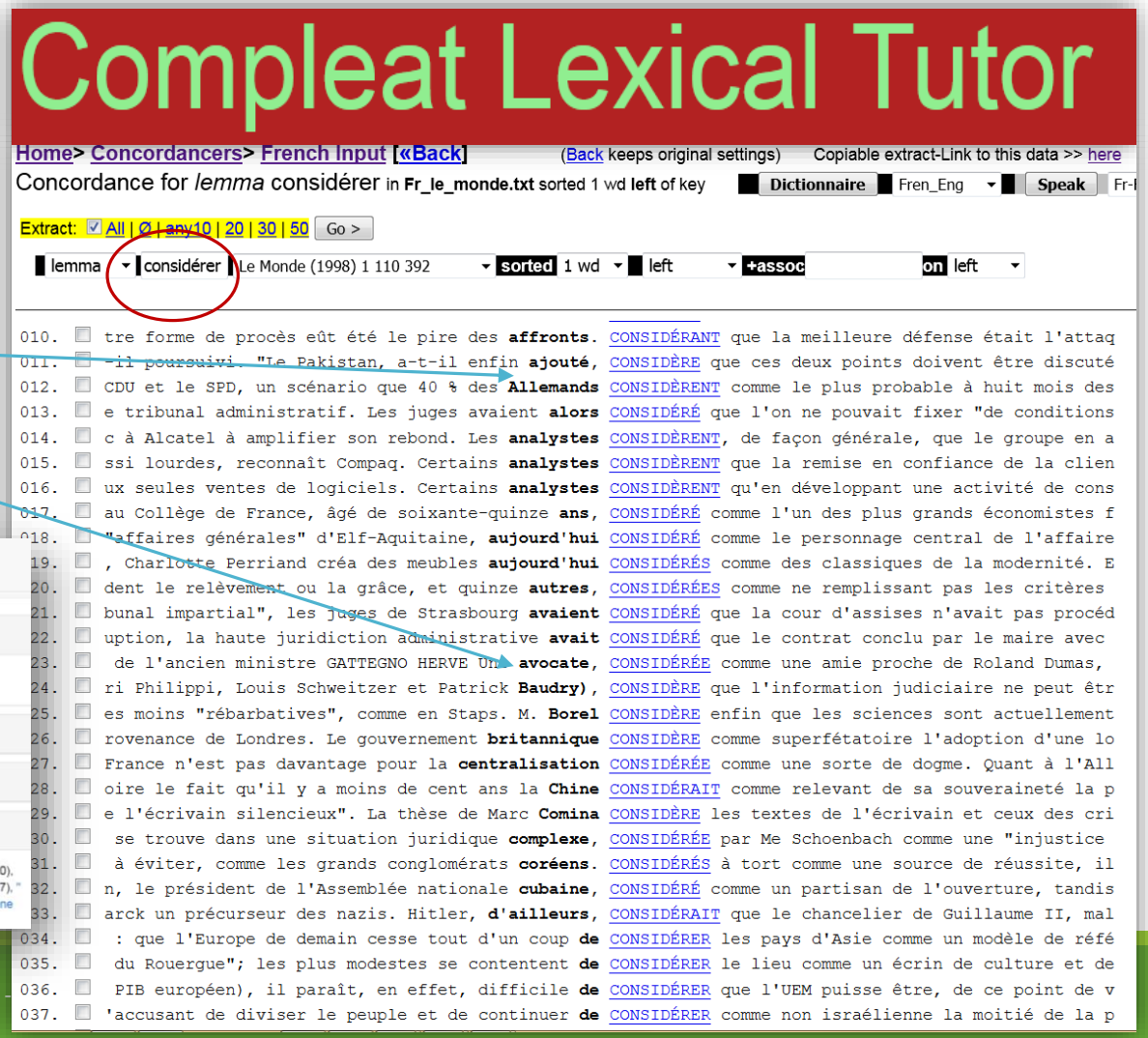


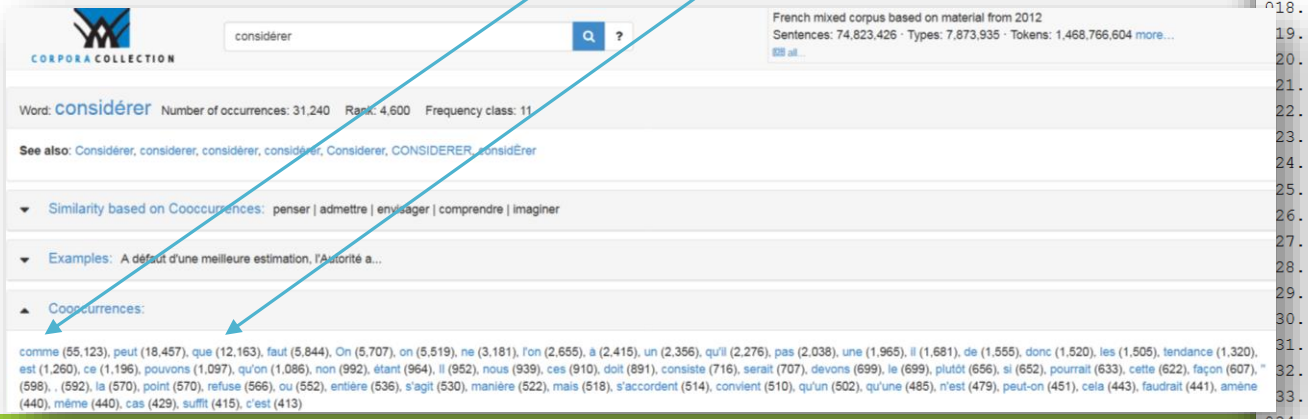
Découverte des extractions : Comment les lire ?

Approche Directe : 2e corpus : un concordancier

Lemme : *considérer*

Observer à droite : *comme / que*





Approche directe : Réflexion sur l'outil

Contextualisation

Comparer les sorties entre sites :

Leur présentation ? Quel type d'information ?...

- Leur contenu ?
- Leur taille ?

→Textes différents donc langue différente

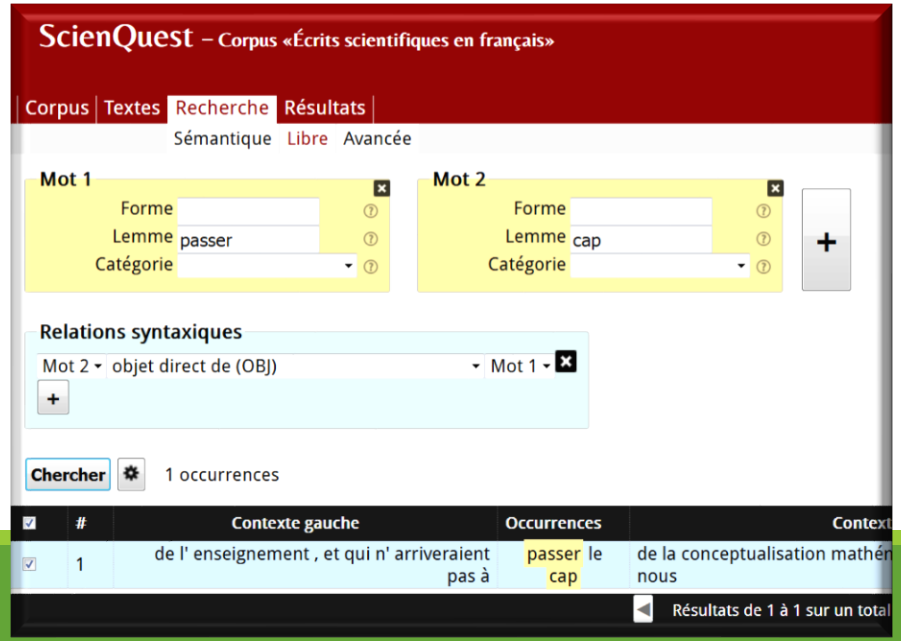
→Taille différente donc résultats à harmoniser

→Année des textes…

Introduction de la phraséologie

Pour chaque outil : recherche de mot isolé et de phrasème

- Ngram Viewer : *mettre la table / dresser la table*
- Corpus français de Leipzig : pas possible
- Lextutor : *cause / remettre en cause*
- Lexicoscope : *cap / passer le cap*
- ScienQuest : *cap / passer le cap / émettre une hypothèse*
  - Entrée libre

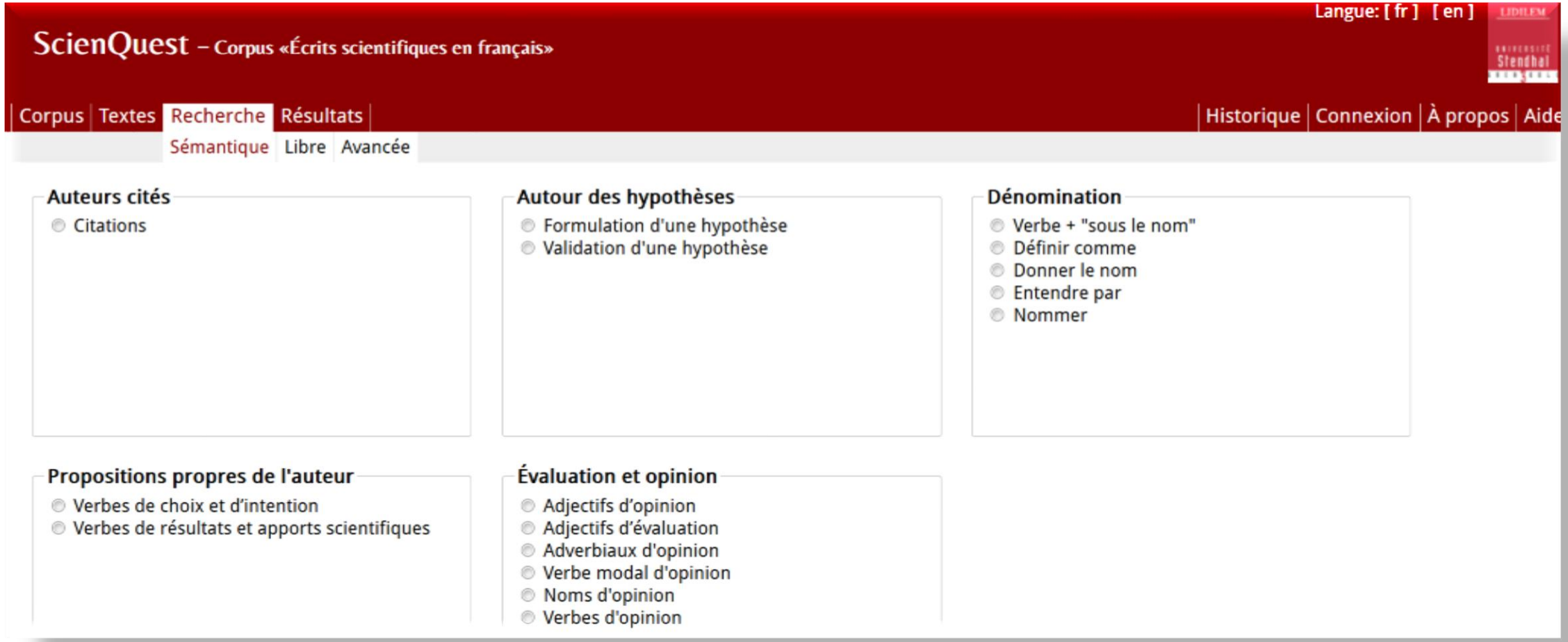


S’interrogent sur leurs besoins

Questions sur leurs écrits

- Comment écrire mes questions de recherche?
- Comment citer un auteur ?
- …

Ils découvrent alors la recherche sémantique



Vers l'écriture

1

épondez aux questions suivantes

1. D'après vous, cet extrait vient de quelle partie de l'article ? Pourquoi ?
   A. résumé // B. introduction // C. théorie // D. méthodologie // E. conclusion
2. Découpez l'extrait en paragraphes et justifiez votre choix en précisant les mots et les thématiques de votre découpage à l'aide du tableau suivant :

Contexte	Sujet	Objectif	Question

3. Trouvez des expressions synonymes pour dire « on tente de… » à l'aide de l'outil CNTRL en ligne : http://www.cnrtl.fr/synonymie/

2

Découpage

Contexte

yse linguistique que nous proposons s'inscrit en réponse à une demande de la SNCF relative à la perception du confort global en train. Les objectifs du projet sont d'identifier les propriétés sémantiques du confort en train et leurs relations de dépendance, à partir d'analyses linguistiques et cognitives.

Sujet

- **Cet article porte sur** la manière dont les formes linguistiques en contexte, utilisant les ressources de la langue mises en discours, renseignent sur les structures cognitives construites à partir des perceptions sensorielles.

Objectif

- **On tente ici d'**identifier, à partir de l'analyse des formes adjectivales, les représentations individuelles et partagées qui se construisent dans les discours des voyageurs, lorsqu'on les interroge sur leur expérience sensible du confort.

Question

- **La première question que l'on se pose alors est la suivante** : le confort en train est-il une catégorie cognitive de ce type ?

46

3

Liste des éléments

- Présenter le sujet : → **Cet article porte sur**
- Introduire les questions pour cet article : → **La première question que l'on se pose alors est la suivante**
- Présenter le contexte : → **L'analyse linguistique que nous proposons s'inscrit (en / dans)**
- Présenter les objectifs de l'article : → **On tente ici de**

Trouvez-en d'autres

Précisions ensuite sur introduction / questions de recherche / citer un auteur…

Présentation

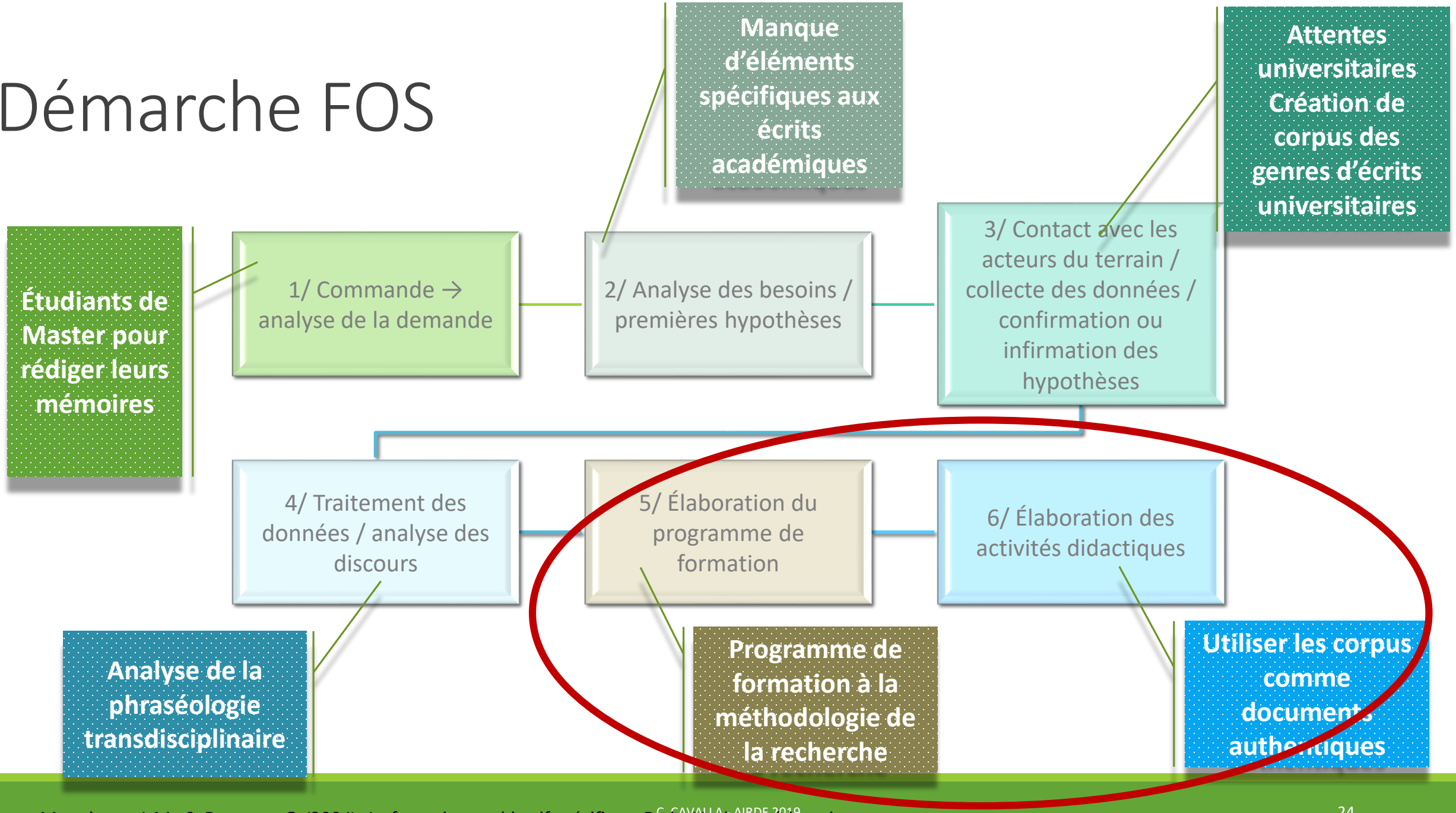
1. **Hypothèse et objectifs**

2. **Outils d'analyses et pédagogiques**
   - Outils pour séquence
   - Type de corpus
   - Analyses linguistiques
   - Approche pédagogique

3. **Séquence didactique**
   - Contexte d'enseignement
   - Introduction des corpus
   - Introduction de la phraséologie

4. **Conclusion**

Démarche FOS



Mangiante, J-M., & Parpette, C. (2004). *Le français sur objectif spécifique*. Paris: Hachette éducation.

Synopsis de la séquence

Objectifs généraux	Objectifs spécifiques	Déroulement	Temps – Matériel
Apprendre à rédiger des écrits académiques à l'aide **1. de corpus numériques** **2. de la phraséologie transdisciplinaire**	**Découverte de l'outil corpus. Comprendre les apports des corpus numériques par rapport à l'offre du net en général.**	Échange et discussion sur leurs représentations de l'écrit académique. Présentation d'outils numériques à partir d'énigmes à résoudre pour en comprendre les limites. Jeu d'intuition linguistique.	4h : dépend des prérequis des apprenants. Ngram Viewer Linguee et/ou Tradooit Logiciel d'interrogation du net.
	Appropriation des corpus numériques. Prise de conscience du rôle de la phraséologie transdisciplinaire dans les discours académiques.	Manipulation de corpus à l'aide de requêtes guidées. 1. Entrée lexicale : chercher le mot « hypothèse » et les éléments associés. 2. Entrée sémantique : chercher comment énoncer une opinion dans un écrit académique. Élargissement à d'autres corpus. Réflexion métalinguistique sur les extractions de concordances.	6h : dépend des profils disciplinaires des apprenants. Le Corpus Français de Liepzig Lextutor Lexicoscope ScienQuest
	Préparation, à l'aide des corpus et des éléments phraséologiques, d'une présentation de leur sujet de mémoire.	Sélection et manipulation d'éléments de la phraséologie transdisciplinaire : 1. à partir de questions générales (rédiger une introduction, des hypothèses…) 2. à partir de leurs interrogations personnelles	10h : dépend des attentes pour le produit final. Corpus au choix Diaporama

Pour conclure

Pas de stress sur la langue

Pas de traduction

La recherche personnelle motive

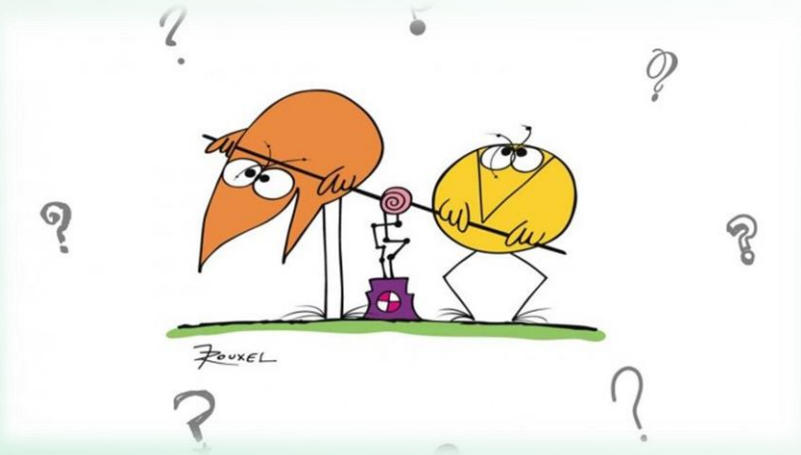
Entrée sémantique rassurante

En somme, une fois que l'apprenant sait où chercher l'information utile à sa rédaction personnelle, une fois qu'il a pris conscience de l'existence de la phraséologie transdisciplinaire, alors son autonomie scripturale dans la production de textes scientifiques s'en trouve accrue.

14 ° Colloque International
27 - 28 - 29 Aout 2019 - Lyon

Association Internationale pour la Recherche en Didactique du Français

Les concepts dans la recherche en didactique du français
Émergence et création d'un champ épistémique



Merci

Les corpus numériques pour l'aide à l'écriture académique

CRISTELLE CAVALLA

cristelle.cavalla@sorbonne-nouvelle.fr

UNIVERSITÉ SORBONNE NOUVELLE PARIS 3
Membre de l'Université Sorbonne Paris Cité

DILTEC - EA 2288
Didactique des langues, des textes et des cultures