

SELEXINI

SEsemantic **LEX**icon **IN**duction for **I**nterpretability and diversity in text processing

Induction de lexiques sémantiques pour
l'interprétabilité et la diversité en traitement de textes

Projet financé par l'Agence Nationale de la Recherche
2022-2025

Le TAL : Traitement Automatique des Langues

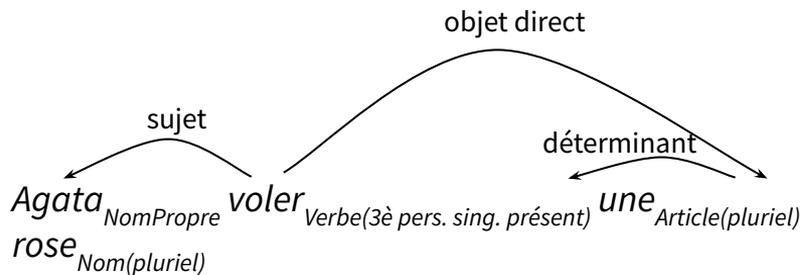
- Modèles informatiques pour traiter les langues humaines, p.ex.
 - Traduire (DeepL, Google translate)
 - Répondre à des questions (moteur de recherche)
 - Dialoguer, demander de l'aide (chatbot, assistant virtuel)
 - Classifier des messages (filtre spam)
 - Aider à mieux écrire (correcteur)
 - ...

Le TAL traditionnel

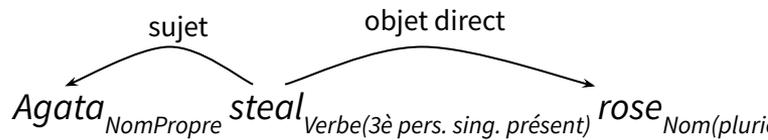
- Connaissances linguistiques : comment fonctionnent les langues ?
 - Morphologie : formation des mots incessamment = in+cesser+ment (adverbe)
 - Syntaxe : ordre des mots sujet, verbe, complément
 - Sémantique : sens des mots et phrases voler = dérober ou s'envoler ?
- Représentation informatique symbolique de ces connaissances
 - Dictionnaires électroniques, règles de grammaire, logique...
- Combinée avec statistiques pour traiter l'ambiguïté
 - Estimées sur des données voler + objet direct = dérober (96% des occurrences)

Le TAL traditionnel : exemple

- Traduction automatique français → anglais
 - Article pluriel (des) + nom pluriel (roses) → nom pluriel (roses) (89%)
 - *Voler* + objet direct → *steal* (96%) ou *fly* (4%)



Agata vole des roses

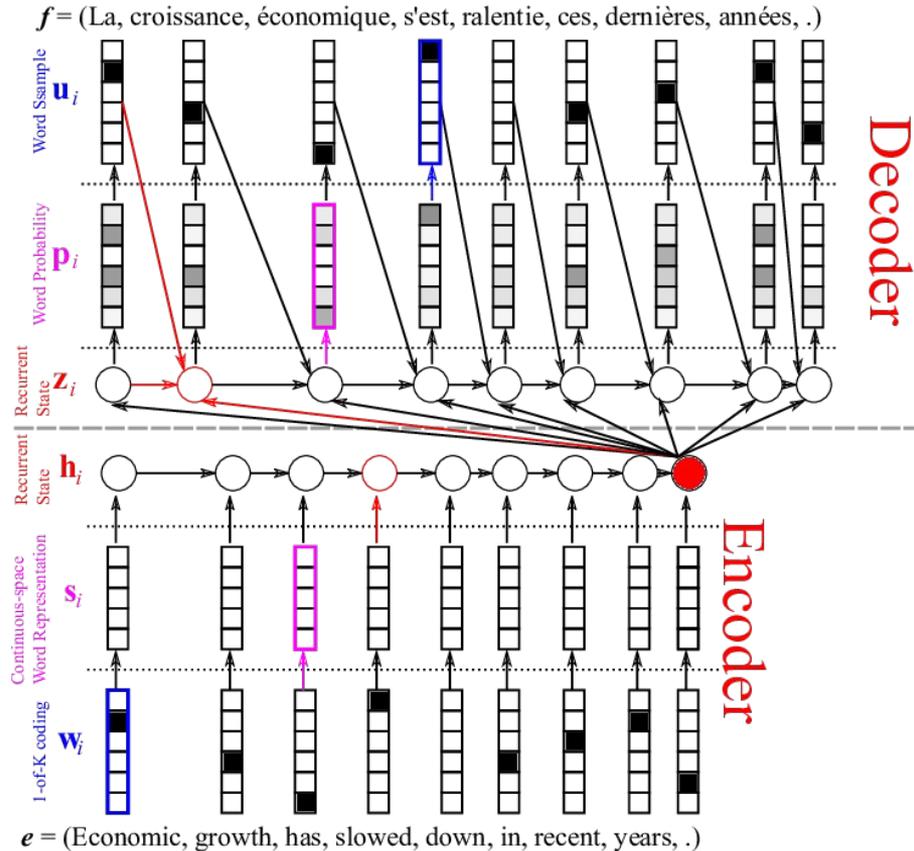


Agata steals roses

Années 2010 : la révolution du *deep learning*

- Les "réseaux de neurones"
 - Rien à voir avec les neurones du cerveau
 - Des fonctions mathématiques paramétrables (p.ex. somme pondérée)
 - Une méthode pour les combiner arbitrairement (graphes de calcul)
 - Les mots deviennent des suites de nombres (vecteurs)
 - Tout est initialisé avec des nombres tirés au hasard (aléatoires)
 - Comparer le résultat des fonctions avec le résultat attendu (loss)
 - Ajuster les paramètres pour minimiser la différence (backpropagation)
 - Répéter sur beaucoup d'exemples d'entrée/sortie (entraînement)

Traduction version *deep learning*



- Chaque mot anglais devient une suite de nombres (w_i)
- Ces nombres passent par des fonctions mathématiques (h_i, z_i, \dots)
- Les nombres résultants représentent des mots français (u_i)
- Au départ, on génère n'importe quoi (initialisation aléatoire)
- Petit à petit, le système "apprend" à convertir les entrées en sorties

Le TAL aujourd'hui: entre enthousiasme ...

- Engouement pour les modèles neuronaux
 - Amélioration des performances substantielle et régulière
 - Meilleure qualité de traduction, recherche, réponse...
 - Approches “de bout en bout” possibles, se passant de l'analyse linguistique traditionnelle
 - Bye bye morphologie, syntaxe, sémantique...
 - Apprentissage par transfert à partir de textes bruts
 - On ne démarre pas de l'aléatoire
 - Pas besoin de beaucoup d'exemples pour bien apprendre

Le TAL aujourd'hui: ... et limites

- Opacité des modèles:
 - Des milliards de paramètres (nombres réels)
 - Difficile à interpréter pour des humains
- Manque de diversité:
 - Évaluation répétée sur benchmarks comportant des **biais**
 - Les phénomènes **fréquents** sont privilégiés dans les évaluations
 - Peut induire un manque de **robustesse**
- Représentation non explicite de la composition:
 - Composition régulière: structure argumentale sujet-verbe-complément
 - Composition irrégulière: expressions prendre en compte, arme blanche

Postulat du projet SELEXINI

- Les notions de **lexique** et d'**unité lexicale** sont cognitivement importantes
- Expliciter les unités lexicales lors de traitements neuronaux permettra une meilleure **interprétabilité** des sorties



Objectif du projet (1)

Développer des techniques pour **induire** automatiquement un lexique sémantique

- à partir de corpus bruts (grandes quantités de texte brut en français)
- par clustering semi-supervisé (création de groupements)
 - Graines = unités lexicales et exemples Wiktionary (mots associés aux exemples)
 - Grouper de phrases similaires par rapport aux graines
 - Petit à petit, inclure des nouvelles phrases, créer des nouveaux groupes

Objectifs du projet (2)

... et **utiliser ce lexique** au sein de modèles de TAL *deep learning*

- Dans un systèmes de compréhension du langage (questions-réponses)
 - Évaluation ciblée (type/structure des questions/réponses selon le lexique...)
 - Intégration dans le processus d'inférence
- À quoi sert le lexique induit ?
 - Expliquer les prédictions (interprétabilité+)
 - Pousser les limites du système (diversité+)

Verrous et clés

- **Opacité des modèles**

- ⇒ Interprétabilité "par construction"

- **Manque de diversité**

- ⇒ Lexique couvrant, grand corpus "annoté", évaluation ciblée

- **Composition implicite**

- ⇒ Structure argumentale, expressions polylexicales

Intérêts de la méthode proposée

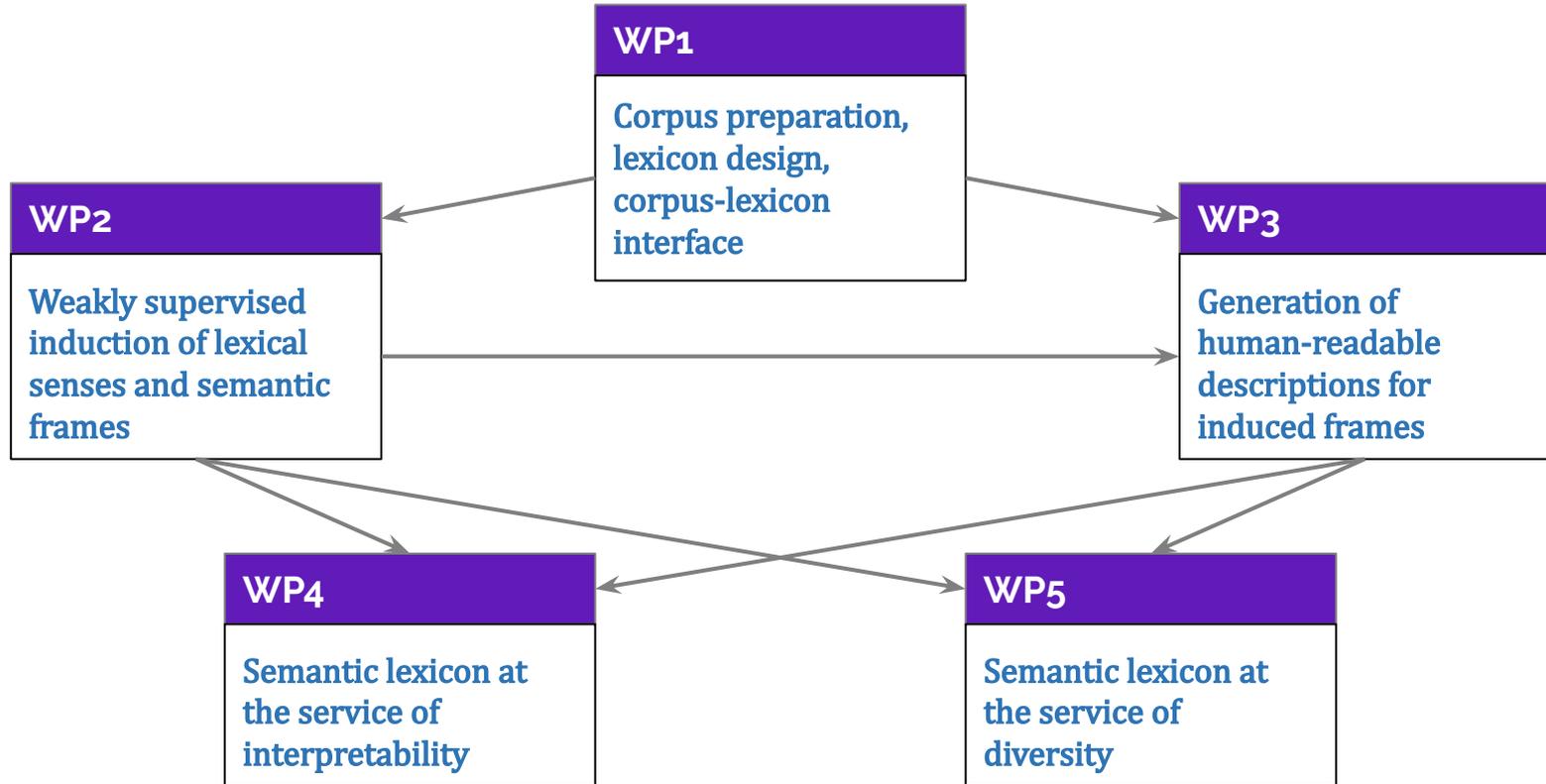
⇒ les représentations numériques des mots (vecteurs)

- Seront reliées aux entrées du **lexique** induit
- Deviendront humainement **interprétables**
- joueront un **rôle intermédiaire** entre
 - Vecteurs statiques : **un seul vecteur** pour un mot ambigu comme “*voler*”
 - Vecteurs contextuels : **un vecteur différent à chaque occurrence** de “*voler*”

Pourquoi induire un lexique?

- Interpretabilité “par construction”
 - Modèle hybride avec représentations symboliques / continues
- Corpus annoté en sens obtenu comme produit dérivé
 - Chaque mot est relié à une entrée du lexique
 - Cette entrée regroupe d'autres occurrences similaires
 - Ce groupe d'occurrences similaires représente le sens du mot
 - Ressource manquante pour la plupart des langues, dont le français
- Couverture large (mais bruité)
- S'appuyant sur ressource existante: Wiktionary
 - Lexiques couvrants existant pour de nombreuses langues
 - Méthode pourrait être facilement adaptée à d'autres langues

Work packages



Consortium

- LIS - Aix Marseille Université (C. Ramisch)
- LLF - Université de Paris (M. Candito)
- LISN - Université de Paris-Saclay (A. Savary)
- LIFAT - Université de Tours (A. Soulet)
- ATILF - CNRS Grand Est (M. Constant)

Contact

- <https://selexini.lis-lab.fr>
- selexini-core@lisn.upsaclay.fr