

# Hee-Soo CHOI

Nancy, France — hee-soo.choi@loria.fr — (+33) 618137231 — <https://hschoi4.github.io/>

## FORMATION

---

**Doctorat mention Sciences du Langage** 2021 — Aujourd'hui

Laboratoires ATILF - LORIA, Université de Lorraine, Nancy, France  
Lier des ressources lexicales du français en vue d'une interopérabilité entre niveaux linguistiques

**Master Langue et Informatique** 2019 — 2021

Sorbonne Université, Paris, France  
Mémoire M2 : Vérification d'universaux linguistiques sur des corpus multilingues annotés en syntaxe de dépendances

**Licence Langue Française et Techniques Informatiques (LFTI)** 2016 — 2019

Sorbonne Université, Paris, France

**Première Année Commune aux Etudes de Santé (PACES)** 2014 — 2016

Université Pierre et Marie Curie, Paris, France

## EXPÉRIENCES PROFESSIONELLES

---

**Doctorante invitée** février 2024 — Aujourd'hui

SapienzaNLP, Sapienza University of Rome, Rome, Italie  
Sous la direction du Pr. Roberto Navigli.  
Projet de recherche sur la levée d'ambiguïté lexicale en français par intégration de ressources lexico-sémantiques.

**Stage de recherche** avril 2021 — septembre 2021

Laboratoire LORIA, Nancy, France

- Vérification d'universaux linguistiques sur des corpus multilingue annotés en syntaxe de dépendances
- Étude de quatre ordres de mots sur 74 langues/141 corpus

**Stage de recherche** juillet 2020 — août 2020

Laboratoire LORIA, Nancy, France

- Études bibliographiques en typologie des langues sur les universaux linguistiques
- Études statistiques multilingues sur les corpus Universal Dependencies 2.7 par réécriture de graphes
- Détermination d'un ordre dominant sujet - verbe - objet sur 74 langues/141 corpus

## ENSEIGNEMENTS

---

**Bases de données - TD 4h/TP 8h** 2023

1ère année, Télécom Nancy

- Modélisation conceptuelle et logique des données
- Algèbre relationnel et calcul relationnel de tuples
- Langages SQL et PL/PgSQL

**Introduction à Python - TD 24h/an** 2021 — 2022

L2 Sciences du Langage, Université de Lorraine

Création de la totalité des supports du cours, des TD et d'évaluation (voir : <https://hschoi4.github.io/teaching/>)

- Introduction à l'algorithmique
- Structures de données et de contrôle
- Expressions régulières
- Manipulation de fichiers et fonctions
- Traitements et analyses statistiques de textes

## PUBLICATIONS

---

### Conférences internationales avec comité de lecture

- **Beyond Model Performance: Can Link Prediction Enrich French Lexical Graphs?**  
Hee-Soo Choi, Priyansh Trivedi, Mathieu Constant, Karèn Fort and Bruno Guillaume  
Accepté à LREC-COLING 2024, mai 2024, Turin, Italie

- **Corpus-based Language Universals Analysis using Universal Dependencies**  
Hee-Soo Choi, Bruno Guillaume and Karèn Fort  
Quantitative Syntax (QUASY), SyntaxFest 2021, mars 2022, Sofia, Bulgarie
- **Investigating Dominant Word Order on Universal Dependencies with Graph Rewriting**  
Hee-Soo Choi, Bruno Guillaume, Karèn Fort and Guy Perrier  
Recent Advances in Natural Language Processing (RANLP) 2021, septembre 2021, en distanciel, Bulgarie

### Conférences nationales avec comité de lecture

- **Des ressources lexicales du français et de leur utilisation en TAL : étude des actes de TALN**  
Hee-Soo Choi, Karèn Fort, Bruno Guillaume, Mathieu Constant  
Actes de la 30e Conférence sur le Traitement Automatique des Langues Naturelles (TALN), juin 2023, Paris, France
- **État de l'art : Liage de ressources lexicales du français**  
Hee-Soo Choi  
Actes de la 24e Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL), juin 2022, Avignon, France

### Séminaires invités

- **Corpus-based Language Universals Analysis using Universal Dependencies**  
Séminaire Traitement Automatique du Langage Ecrit et Parlé (TALEP)  
Laboratoire d'Informatique et Systèmes (LIS), mars 2023, Marseille, France
- **Investigating Dominant Word Order on Universal Dependencies with Graph Rewriting**  
Séminaire Research Group in Mathematical Linguistics (GRLMC), mai 2022, Tarragona/distanciel, Espagne

### Conférences nationales sans comité de lecture

- **Analyse orientée corpus d'universaux de Greenberg sur Universal Dependencies**  
Hee-Soo Choi, Bruno Guillaume, Karèn Fort  
GDR LIFT Linguistique Informatique, Formelle et de Terrain 2021, décembre 2021, Grenoble, France

## ORGANISATION ET RESPONSABILITÉS

---

<b>Co-encadrement de Master 1</b>	février 2024 - juin 2024
Anisa Amourat et Assia Bouznad, M1 Langue et Informatique, Sorbonne Université Création de ressources annotées pour la levée d'ambiguïté lexicale en français (co-encadrement avec Karèn Fort)	
<b>Relectrice et sous-relectrice</b>	2024
LREC-COLING 2024	
<b>Co-responsable du CafÉTAL</b>	2023
Séminaire de l'équipe TAL de l'ATILF	
<b>Membre du comité d'organisation de la conférence IWCS 2023</b>	2023
IDMC, Nancy, France	

## MÉDIATION

---

### Intervention CHICHE - 1 scientifique, 1 classe

- 4 classes - Lycée Saint-Exupéry, Fameck, France
  - 3 classes - Lycée Varoquaux, Tomblaine, France
  - Classe Scientifilles - Lycée Claude Gellée, Épinal, France
- janvier 2024  
décembre 2023  
janvier 2023

## COMPÉTENCES INFORMATIQUES

---

- **Langages de programmation** : Python, bash, L<sup>A</sup>T<sub>E</sub>X, SQL, PL/PgSQL, Java
- **Autres** : Git, Environnement Linux

## LANGUES

---

- **Français** : langue maternelle
- **Anglais, Coréen** : niveau C2
- **Espagnol** : niveau B1
- **Chinois, Grec moderne, Italien** : notions