

# Timothee MICKUS

## PERSONAL DATA

---

EMAIL: [tmickus@atilf.fr](mailto:tmickus@atilf.fr)      GITHUB: [TimotheeMickus](https://github.com/TimotheeMickus)      WEBSITE: <https://timotheemickus.github.io>  
PHONE: +33 7 69 89 33 08      SOCIAL:    

## PUBLICATIONS & COMMUNICATIONS

---

- JUL 2022 “Semeval-2022 Task 1: CODWOE – Comparing Dictionaries and Word Embeddings”  
T. MICKUS, K. VAN DEEMTER, M. CONSTANT & D. PAPERNO;  
URL: <https://aclanthology.org/2022.semeval-1.1/>
- JUN 2022 “How to Dissect a Muppet: the Structure of Transformer Embedding Space.”  
T. MICKUS, D. PAPERNO & M. CONSTANT
- DEC 2021 “About Neural Networks and Writing Definitions”  
T. MICKUS, D. PAPERNO & M. CONSTANT; oral online presentation at DSNA 2021 conference.  
URL: <https://muse.jhu.edu/article/843061>
- AUG 2021 “A Game Interface to Study Semantic Grounding in Text-Based Models”  
T. MICKUS, M. CONSTANT & D. PAPERNO; oral online presentation at Games and NLP (co-located with IEEE CoG).  
URL: <https://ieeexplore.ieee.org/document/9619149>
- DEC 2020 “What Meaning-Form Correlation Has to Compose With”  
T. MICKUS, T. BERNARD & D. PAPERNO; oral online presentation at COLING 2020 conference.  
URL: <https://www.aclweb.org/anthology/2020.coling-main.333/>
- JUN 2020 “Génération Automatique de Définitions pour le Français”  
T. MICKUS, M. CONSTANT & D. PAPERNO; oral online presentation at TALN 2020 conference.  
URL: <https://hal.archives-ouvertes.fr/hal-02784756>
- JAN 2020 “What do you mean, BERT? Assessing BERT as a Distributional Semantics Model”  
T. MICKUS, D. PAPERNO, M. CONSTANT & K. VAN DEEMTER; oral presentation at Scil 2020 conference.  
URL: <https://scholarworks.umass.edu/scil/vol3/iss1/34/>
- SEP 2019 “Mark my Word: a Sequence-to-sequence Approach to Definition Modeling”  
T. MICKUS, D. PAPERNO & M. CONSTANT; oral presentation at DL4NLP workshop (NoDALiDA 2019); presented as a poster to CLIN 2020. URL: <http://svn.nlpf.eu/outreach/dl4nlp/2019/W19-0501.pdf>
- JAN 2019 “Distributional Effects of Gender Contrasts Across Categories”  
T. MICKUS, O. BONAMI & D. PAPERNO; presented as a poster at Scil 2019 conference.  
URL: <https://scholarworks.umass.edu/scil/vol2/iss1/19/>

## WORK EXPERIENCE

---

- APR 2022–NOW Postdoctoral Researcher at University of Helsinki, Helsinki  
Machine Translation and explainable NLP, part of the FoTRAN project.
- AUG 2018–SEP 2018 Web & NLP Engineer at ELDA, Paris; *Development–Web Annotation Interface*
- NOV 2017–JUL 2018 NLP Data Scientist at ESPRITS COLLABORATIFS, Paris area; *NLP–R&D*
- NOV 2015–NOV 2017 Software Contractor at SPIRICA, Paris; *Development–Life Insurance Web Application*
- OCT 2015–NOV 2017 R&D Engineer at MIDDLEWARE FACTORY, Paris area; *Computer Programming–R&D*

## EDUCATION

---

### INTERNSHIPS

- FEB 2020–MAR 2020 NLP Intern at AIST AIRC, Tokyo  
Reinforcement learning implementation of communication games. Experiments to determine necessary factors for grounded compositional language emergence, study of existing metrics of compositionality.
- NOV 2017–JUL 2018 NLP Intern at UNIVERSITÉ PARIS DIDEROT, Paris  
Distributional Semantics Models approaches applied to French morphology of gender. Use of word embeddings and statistical analysis to explain facts relevant to gender assignment of French agent nouns.
- JAN 2016–JUN 2016 NLP Intern at UNIVERSITÉ PARIS DIDEROT, Paris  
Compute and compare average conditional entropy over n-grams for two corpora (French and Mauritian Creole).
- JUN 2015–SEP 2015 NLP Intern at LIPN, UNIVERSITÉ PARIS 13, Paris area  
Corrective maintenance for the Hybride project (ANR-11-BS002-002 2011-2015) for extracting symptoms from scientific articles on rare diseases. Hyperparameters optimization (Java & Bash).

## SCHOOL CURRICULUM

- OCT 2018–MAR 2022 PHD. in COMPUTER SCIENCE, ATILF, **Université de Lorraine**, Nancy  
TITLE: “On the Status of Word Embeddings as Implementations of the Distributional Hypothesis”  
Investigations on the linguistic value of word embeddings. Using NLG to study the relation between distributional semantic representations and traditional dictionary definitions.
- JAN 2022 WINTER SCHOOL in MACHINE LEARNING, **ALPS 2022**, Online
- JUL 2019 SUMMER SCHOOL in MACHINE LEARNING, **DeepLearn 2019**, Warsaw
- JUL 2018 Master of Science in COMPUTATIONAL LINGUISTICS, **Université Paris Diderot**, Paris  
GPA: 16.95/20, *summa cum laude*, rank not available. Thesis subject: “Grammatical Gender Contrasts in French: a distributional semantic approach”.
- JUL 2016 Undergraduate Degree in COMPUTATIONAL LINGUISTICS, **Université Paris Diderot**, Paris  
GPA: 15.287/20, *cum laude*, ranked 1<sup>st</sup>.
- JUL 2015 Bachelor of Science in COMPUTATIONAL LINGUISTICS, **Université Paris Diderot**, Paris  
GPA: 16.66/20, *summa cum laude*, ranked 1<sup>st</sup>.

## SCIENTIFIC OUTREACH

### TEACHING

- NOV 2020–FEB 2021 Lectures on **Word Embeddings and Data Mining**, at UNIVERSITÉ DE LORRAINE  
as part of the “Lexical Resources” class of the M.Sc. in NLP.
- OCT 2020–JAN 2021 Practical Lectures on **Programming** (Python), at UNIVERSITÉ DE LORRAINE  
as part of the “Algorithmics & Programming” class of the M.Sc. in NLP.
- OCT 2020–DEC 2020 Practical Lectures on **Web Development**, at UNIVERSITÉ DE LORRAINE  
HTML, CSS, JavaScript.
- NOV 2020–DEC 2020 Practical Lectures on **Office Automation**, at UNIVERSITÉ DE LORRAINE  
Topics covered: Word processor software, Web search.
- NOV 2020 Practical Lectures on **Shell and  $\text{\LaTeX}$** , at UNIVERSITÉ DE LORRAINE  
as part of the “Lexical Resources” class of the M.Sc. in NLP.
- NOV 2019–JAN 2020 Lectures on **Word Embeddings and Data Mining**, at UNIVERSITÉ DE LORRAINE  
as part of the “Lexical Resources” class of the M.Sc. in NLP.
- OCT 2019–JAN 2020 Practical Lectures on **Programming** (Python), at UNIVERSITÉ DE LORRAINE  
as part of the “Algorithmics & Programming” class of the M.Sc. in NLP.
- AUG 2019 Tutor for PYTHON4NLP Summer School, Nancy  
Providing help and guidance to attendees on presented python APIs & libraries (NLTK, BeautifulSoup, spacy, sklearn...).
- NOV 2018–DEC 2018 Lectures on **Word Embeddings and Multilingual Resources**, at UNIVERSITÉ DE LORRAINE  
as part of the “Lexical Resources” class of the M.Sc. in NLP.

### OTHER ADMINISTRATIVE DUTIES

- JUN 2021–JUL 2022 Organizer for the SemEval 2022 Shared Task on Comparing Embeddings & Dictionaries  
URL: <https://github.com/TimotheeMickus/codwoe>
- APR 2021–NOW Member of the strategic committee “Université – Recherche – Santé”
- MAR 2021–JUL 2021 Organizer for the Caf  TAL NLP Reading Group, at ATILF
- JUN 2020–AUG 2020 Supervised two NLP interns (Masters 1 students)  
Internship topic: preliminary work for establishing comparable sets of word embeddings.
- FEB 2019–JUL 2019 Organizer for the NLP Reading Group, at UNIVERSITÉ DE LORRAINE

### PUBLIC OUTREACH

- NOV 2021 Demo stand for the BlankCrack game at the [FORUM DES SCIENCES COGNITIVES 2021    NANCY](#).
- JUL 2021–NOW Personal blog: “[My PhD, explained to my folks](#)” (On hiatus)

## PRACTICAL SKILLS

- COMPUTER SKILLS Python 3 (pytorch, NLTK, django...), Java, JavaScript, Bash, R, XMLHTML/CSS, SQL, Unix (Ubuntu), OpenOffice,  $\text{\LaTeX}$
- LANGUAGES French & English (native), German & Japanese (basic knowledge); notions of Classical Nahuatl, Latin, Ancient Greek, Arabic...