

# LUCAS WEBER

✉ [lucas.weber@upf.edu](mailto:lucas.weber@upf.edu) | [in LinkedIn](#) | [🔗 LucWeber](#) | [🌐 webpage](#) |

---

## Education

<b>Ph.D. Computational Linguistics and NLP</b> <i>University Pompeu Fabra</i> <ul style="list-style-type: none"><li>• Thesis: <i>Interpreting the Learning Dynamics of Language Models</i></li><li>• DTCL-Scholarship</li><li>• supervision by:<ul style="list-style-type: none"><li>* Dr. Elia Bruni (University of Osnabrück)</li><li>* Dr. Dieuwke Hupkes (FAIR)</li></ul></li></ul>	2019 – 2024 Barcelona, ES
<b>RMSc. Brain and Cognitive Sciences; Track: Cognitive Science</b> <i>University of Amsterdam</i> <ul style="list-style-type: none"><li>• with distinction ('cum laude')   equiv. GPA (US): 4.0</li><li>• Thesis-project: <i>MRC-CBU, University of Cambridge</i><ul style="list-style-type: none"><li>* <a href="#">Empirical paper</a>, <a href="#">Review paper</a></li></ul></li></ul>	2016 – 2018 Amsterdam, NL
<b>BSc. Psychology</b> <i>University of Vienna</i> <ul style="list-style-type: none"><li>• equiv. GPA (US): 3.9</li><li>• ERASMUS-exchange: <i>University of Groningen</i></li></ul>	2013 – 2016 Vienna, AT

---

## Talks and Presentations

<b>Fraunhofer IIS: Invited talk</b> <ul style="list-style-type: none"><li>• Presenting work on efficient benchmarking</li></ul>	2024
<b>Samaya AI: Invited talk</b> <ul style="list-style-type: none"><li>• Presenting work on <i>tinyBenchmarks</i></li></ul>	2024
<b>DLBCN 2023</b>   <i>Poster presentation</i> <ul style="list-style-type: none"><li>• Presenting work on <i>Mind the instructions</i></li></ul>	2023
<b>GenBench @ EMNLP 2023</b>   <i>Poster presentation</i> <ul style="list-style-type: none"><li>• Presenting work on <i>Mind the instructions</i> and <i>The ICL consistency test</i></li></ul>	2023
<b>CoNLL 2023</b>   <i>Oral presentation</i> <ul style="list-style-type: none"><li>• Presenting work on <i>Mind the instructions: a holistic evaluation of consistency and interactions in prompt-based learning</i></li><li>• Honourable mention (rated as one of the 4 best papers at the conference)</li></ul>	2023
<b>ENS Paris, Invited talk at CoML</b> <ul style="list-style-type: none"><li>• Presenting work on <i>Teaching with language-commentaries</i></li></ul>	2022
<b>ELLIS-NLP Workshop 2021</b>   <i>Poster-presentation</i> <ul style="list-style-type: none"><li>• Presenting work on <i>Meeting Formal Linguistics and Optimization</i></li></ul>	2021
<b>EVIL (UPF/FAIR-internal meeting): Invited talk</b> <ul style="list-style-type: none"><li>• Presentation <i>Similarity Probing in the Linguistic Task-Space</i></li></ul>	2021
<b>EACL 2021</b>   <i>Oral-presentation</i> <ul style="list-style-type: none"><li>• Presenting work on project <i>Language Modelling as a Multi-Task Problem</i></li></ul>	2021
<b>TAB (Seminar series at UPF/UvA/CIMeC)</b>   <i>Oral presentation</i> <ul style="list-style-type: none"><li>• Presentation <i>Language Modelling as a Multi-Task Problem</i></li></ul>	2021
<b>EVIL (UPF/FAIR-internal meeting)</b>   <i>Oral presentation</i> <ul style="list-style-type: none"><li>• Presentation <i>Continual Learning in Language Emergence in artificial conversational agents</i></li></ul>	2019

---

## Notable publications

- Weber, L., Jumelet, J., Bruni, E. and Hupkes, D. (2024) *Interpreting Language Models via Task Spaces*, ACL 2024
- Maia Polo, F., Weber, L., Choshen, L., Sun, Y., Xu, G. and Yurochkin, M. (2024) *tinyBenchmarks: evaluating LLMs with fewer examples*, ICML 2024
- Weber, L., Bruni, E. and Hupkes, D. (2023) *The ICL consistency test*, GenBench 2023
- Weber, L., Bruni, E. and Hupkes, D. (2023) *Mind the instructions: a holistic evaluation of consistency and interactions in prompt-based learning*, CoNLL 2023
- Weber, L., Jumelet, J., Bruni, E. and Hupkes, D. (2021) *Language Modelling as a Multi-Task Problem*, EACL 2021

---

## Pre-prints / Under review

- Maia Polo, F., Xu, R., Weber, L., Mírian, S., Bhardwaj, O., Choshen, L., Flavio Melo de Oliveira, A., Sun, Y. and Yurochkin, M. (2024) *Efficient multi-prompt evaluation of LLMs*, under Review
- Weber, L., Jumelet, J., Michel, P., Bruni, E. and Hupkes, D. (2023) *Curriculum Learning with Adam: The Devil Is in the Wrong Details*, arXiv

---

## Open-source contributions

<b>tinyBenchmarks</b>	2024
<ul style="list-style-type: none"><li>• Subsamples (n=100) of the MMLU, open LLM leaderboard, helm-lite and AlpacaEval datasets paired with an item-response model to do efficient benchmarking for LLMs. Find on the huggingface hub at <a href="#">tinyBenchmarks/*</a>.</li></ul>	
<b>ICL consistency test</b>	2023
<ul style="list-style-type: none"><li>• A NLP task for the consistency evaluation of LLMs. Find the code on <a href="#">github</a>.</li></ul>	
<b>ICL consistency dataset</b>	2023
<ul style="list-style-type: none"><li>• A NLP dataset the consistency evaluation of LLMs. Find on the huggingface hub at <a href="#">LucasWeber/icl_consistency_test</a>.</li></ul>	
<b>2048 RL environment</b>	2020
<ul style="list-style-type: none"><li>• A OpenAI gym environment to train agents to play the game of 2048 and visualise their play. Find on the code on <a href="#">github</a>.</li></ul>	

---

## Teaching

<b>Computational Semantics</b>   <i>Graduate Teaching Assistance</i>	2021 - 2022
<ul style="list-style-type: none"><li>• DTCL, University Pompeu Fabra</li><li>• Language of instruction: English</li><li>• Content: Introduction to NLP methods for linguistics research</li></ul>	
<b>Recursos TIC</b>   <i>Teaching Assistance</i>	2021 - 2022
<ul style="list-style-type: none"><li>• DTCL, University Pompeu Fabra</li><li>• Language of instruction: Spanish/Catalan</li><li>• Content: IT-resources for linguists</li></ul>	
<b>Descripció i Anàlisi 2</b>   <i>Teaching Assistance</i>	2020 – 2022
<ul style="list-style-type: none"><li>• DTCL, University Pompeu Fabra</li><li>• Language of instruction: English/German/Catalan</li></ul>	

- Content: Phonology/Morphology/comp. Linguistics

**Descripció i Anàlisi 1** | *Teaching Assistance* 2020

- DTCL, University Pompeu Fabra
- Language of instruction: English/German
- Content: Syntax

**Foundations of Neural and Cognitive Modelling** | *Graduate Teaching Assistance* 2017

- IIS, University of Amsterdam
- Language of instruction: English
- Content: Dynamical systems

---

## Organization

**UR-LING workshop** | *Lead organizer* 2021

- United workshop for different sub-fields of linguistics at University Pompeu Fabra

---

## Supervision

**Thesis projects** | *Co-supervision* 2022

- Leon Schmid at IKW, University of Osnabrück
- Simon Helling at IKW, University of Osnabrück

---

## Skills

### Languages:

German (native);  
 English (fluent);  
 Dutch (fluent);  
 Spanish (fluent);  
 French (basic);  
 Catalan (basic);  
 Latin

**Programming languages:** Python, MATLAB, R

**Frameworks:** Pytorch, Tensorflow; DeepSpeed, fairscale, FSDP

**Relevant skills:** High-performance computing (HPC), CUDA, multi-GPU / multi-machine training, model sharding, data parallel