

LUCAS WEBER

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

Education

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

Talks and Presentations

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

First author publications

- 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

Other notable papers

- Maia Polo, F., Weber, L., Choshen, L., Sun, Y., Xu, G. and Yurochkin, M. (2024) *tinyBenchmarks: evaluating LLMs with fewer examples*, ICML 2024

Pre-prints / Under review

- Weber, L., Jumelet, P., Bruni, E. and Hupkes, D. (2024) *Interpreting Language Models via Task Spaces*, 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

tinyBenchmarks	2024
<ul style="list-style-type: none">• 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 tinyBenchmarks/*.	
ICL consistency test	2023
<ul style="list-style-type: none">• A NLP task for the consistency evaluation of LLMs. Find the code on github.	
ICL consistency dataset	2023
<ul style="list-style-type: none">• A NLP dataset the consistency evaluation of LLMs. Find on the huggingface hub at LucasWeber/icl_consistency_test.	
2048 RL environment	2020
<ul style="list-style-type: none">• A OpenAI gym environment to train agents to play the game of 2048 and visualise their play. Find on the code on github.	

Teaching

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

- 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