

# Resume

Paul Michel

16 Villa de Lourcine • 57014 Paris, France

CELL +33 6 46 53 57 14 • E-MAIL [pmichel31415@gmail.com](mailto:pmichel31415@gmail.com) • WEB <https://pmichel31415.github.io/>

## EDUCATION

- 08/2021 **PhD in Language Technologies**  
Carnegie Mellon University, Pittsburgh PA, USA
- 08/2018 **Master of Language Technologies**  
Carnegie Mellon University, Pittsburgh PA, USA
- 08/2016 **Master of Computer Science**  
École polytechnique, Palaiseau, France
- 09/2015 **Bachelor of science**  
École polytechnique, Palaiseau, France
- 09/2013 **Undergraduate preparatory school**  
Lycée Louis-le-Grand, Paris, France

## PROFESSIONAL EXPERIENCE

- 09/2021-present **Laplace Postdoctoral Chair in Data Science**  
École normale supérieure PSL, Center for Data Science, Paris, France  
Work on emergent communication and distributional robustness in Emmanuel Dupoux's team.
- 01/2021-05/2021 **Research internship**  
DeepMind, Language team, London UK (remote from Paris, France)  
Work on distributionally robust optimization applied to large scale generative language models.
- 05/2018-08/2018 **Research internship**  
Facebook, Language and Translation Technologies team, Menlo Park, USA  
Work on robustness of Neural Machine Translation models to noise and adversarial inputs.
- 03/2016-08/2016 **Research internship**  
École normale supérieure PSL, LSCP, Paris, France  
Under the supervision of Emmanuel Dupoux. Work on unsupervised phoneme segmentation with recurrent neural networks
- 07/2015-08/2015 **Software design intern**  
Horiba Ltd., Kyoto, Japan  
Software R&D in Kyoto, Japan: embedded systems programming, development of IoT technology, internet and security programming

## PUBLICATIONS

- 2021 **Balancing Average and Worst-case Accuracy in Multitask Learning**  
Preprint, under review  
Paul Michel, Sebastian Ruder, Dani Yogatama  
<https://arxiv.org/pdf/2110.05838>
- 2021 **Should We Be Pre-training? An Argument for End-task Aware Training as an Alternative**  
Preprint, under review  
Lucio M Dery, Paul Michel, Ameet Talwalkar, Graham Neubig  
<https://arxiv.org/pdf/2109.07437>
- 2021 **Learning Neural Models for Natural Language Processing in the Face of Distributional Shift**  
PhD Dissertation  
Paul Michel  
<https://arxiv.org/pdf/2109.01558>
- 2021 **Examining and Combating Spurious Features under Distribution Shift**  
Proceedings of ICML 2021  
Chunting Zhou, Xuezhe Ma, Paul Michel, Graham Neubig  
<https://arxiv.org/pdf/2106.07171>
- 2021 **Modeling the Second Player in Distributionally Robust Optimization**  
Proceedings of ICLR 2021  
Paul Michel, Tatsunori Hashimoto, Graham Neubig  
<https://openreview.net/pdf?id=ZDnzZrTqU9N>
- 2020 **Findings of the WMT 2020 Shared Task on Machine Translation Robustness**  
Proceedings of WMT 2020  
Lucia Specia, Zhenhao Li, Juan Pino, Vishrav Chaudhary, Francisco Guzmán, Graham Neubig, Nadir Durrani, Yonatan Belinkov, Philipp Koehn, Hassan Sajjad, Paul Michel, Xian Li  
<https://www.aclweb.org/anthology/2020.wmt-1.4.pdf>
- 2020 **Regularizing Trajectories to Mitigate Catastrophic Forgetting**  
Preprint  
Paul Michel, Elisabeth Salesky, Graham Neubig  
[http://www.cs.cmu.edu/~pmichel1/pubs/assets/conatural\\_gradient.pdf](http://www.cs.cmu.edu/~pmichel1/pubs/assets/conatural_gradient.pdf)
- 2020 **Optimizing Data Usage via Differentiable Rewards**  
Proceedings of ICML 2020  
Xinyi Wang, Hieu Pham, Paul Michel, Antonios Anastasopoulos, Graham Neubig, Jaime Carbonell  
<https://arxiv.org/pdf/1911.10088>
- 2020 **Weight Poisoning Attacks on Pre-trained Models**  
Proceedings of ACL 2020  
Keita Kurita, Paul Michel, Graham Neubig  
<https://arxiv.org/pdf/2004.06660>

- 2019 **Are Sixteen Heads Really Better than One?**  
Proceedings of NeurIPS 2019  
Paul Michel, Omer Levy, Graham Neubig  
<https://arxiv.org/pdf/1905.10650>
- 2019 **Findings of the First Shared Task on Machine Translation Robustness**  
Proceedings of WMT 2019  
Xian Li, Paul Michel, Antonios Anastasopoulos, Yonatan Belinkov, Nadir Durrani, Orhan Firat, Philipp Koehn, Graham Neubig, Juan Pino, Hassan Sajjad  
<https://arxiv.org/pdf/1906.11943>
- 2019 **compare-mt: A Tool for Holistic Comparison of Language Generation Systems**  
Proceedings of NAACL 2019, Demo Paper  
Graham Neubig, Zi-Yi Dou, Junjie Hu, Paul Michel, Danish Pruthi, Xinyi Wang  
<https://arxiv.org/pdf/1903.07926>
- 2019 **On Evaluation of Adversarial Perturbations for Sequence-to-Sequence Models**  
Proceedings of NAACL 2019, Long Paper  
Paul Michel, Xian Li, Graham Neubig, Juan Miguel Pino  
<https://arxiv.org/pdf/1903.06620>
- 2018 **MTNT: A Testbed for Machine Translation of Noisy Text**  
Proceedings of EMNLP 2018, Long Paper  
Paul Michel, Graham Neubig  
<https://arxiv.org/pdf/1809.00388>
- 2018 **Extreme Adaptation for Personalized Neural Machine Translation**  
Proceedings of ACL 2018, Short Paper  
Paul Michel, Graham Neubig  
<https://arxiv.org/pdf/1805.01817>
- 2017 **DyNet: The Dynamic Neural Network Toolkit**  
Arxiv preprint  
Graham Neubig, Chris Dyer, Yoav Goldberg, Austin Matthews, Waleed Ammar, Antonios Anastasopoulos, Miguel Ballesteros, David Chiang, Daniel Clothiaux, Trevor Cohn, Kevin Duh, Manaal Faruqui, Cynthia Gan, Dan Garrette, Yangfeng Ji, Lingpeng Kong, Adhiguna Kuncoro, Gaurav Kumar, Chaitanya Malaviya, Paul Michel, Yusuke Oda, Matthew Richardson, Naomi Saphra, Swabha Swayamdipta, Pengcheng Yin  
<https://arxiv.org/pdf/1701.03980>
- 2017 **Blind phoneme segmentation with temporal prediction errors**  
Proceedings of ACL 2017, Student Research Workshop  
Paul Michel, Okko Räsänen, Roland Thiolliere, Emmanuel Dupoux  
<https://arxiv.org/pdf/1608.00508>
- 2017 **Does the Geometry of Word Embeddings Help Document Classification?  
A Case Study on Persistent Homology Based Representations**  
Proceedings of the 2nd Workshop on Representation Learning for NLP  
Paul Michel, Abhilasha Ravichander, Shruti Rijhwani  
<https://arxiv.org/pdf/1705.10900>

## TEACHING EXPERIENCE

09/2019-12/2019 **CS 11-731: Machine Translation and Sequence-to-sequence Models**  
**Teaching Assistant**

Language Technologies Institute, CMU

Office hours, teaching and grading. Held lecture on “Robust Machine Translation” <http://www.phontron.com/class/mtandseq2seq2019>

07/2019 **Machine translation**  
**Lab Instructor**

JSALT 2019 Summer School,

Taught lab in neural machine translation.

<https://www.clsp.jhu.edu/workshops/19-workshop/2019-hlt/>

01/2018-06/2018 **CS 11-747: Neural Networks for NLP**  
**Teaching Assistant**

Language Technologies Institute, CMU

Office hours, teaching and grading. Held lecture on “Using/Evaluating Sentence Representations” <http://www.phontron.com/class/nn4nlp2018>

## INVITED TALKS

12/2021 **LIMSI, CNRS** Departmental Seminar  
09/2021 **Google Research** NLP-discuss talk series  
07/2021 **MILA** Deep learning theory reading group  
01/2021 **Deepmind** Language reading group  
10/2020 **NLP With Friends** Online seminar  
01/2019 **INRIA** Almanach Team  
07/2018 **Google Translate**

## SERVICE

**Area Chair** NAACL (21)  
**Reviewer** ACL (20), NAACL (19,20), EMNLP (20, Outstanding reviewer), ICLR(20,21), NeurIPS (20,21), WMT (18), ICML (20, Top reviewer, 21, Expert reviewer), W-NUT (19), ACL-SRW (18)  
**(Co-)Organizer** Second WMT Robustness Shared Task (20), First WMT Robustness Shared Task (19)  
**Department** Editor, CMU Machine Learning Blog (20-21), Co-organizer, CSD Seminar talk series at École normale supérieure (21)

## SKILLS

**Computer programming:** Main language Python, past experience in C++, C, Java, PHP and OCaml. Experience in Dynet (contributor), Pytorch (user) and Jax (user). Limited experience in Tensorflow.

**Languages:** French (native), English (fluent), German (proficient), Japanese beginner (1 year of study)