Juan Pablo Vigneaux

jpvigneaux@northwestern.edu
https://jpvigneaux.github.io/ github.com/jpvigneaux

US permanent resident; no visa sponsorship required.

Research interests

Interpretability of language and speech models; representations of syntax, meaning and intonation in neural networks; information theory and mathematical foundations of AI (category theory, geometry of representations).

Education

Université Paris Cité, Paris, France.
Doctorate, Mathematics, 2019.
Dissertation: "Topology of Statistical Systems: A Cohomological Approach to Information Theory."
Supervisor: Daniel Bennequin.

Université Pierre et Marie Curie (UPMC), Paris, France.
Master, Fundamental Mathematics, *mention très bien* (highest honors), 2015.

Pontificia Universidad Católica de Chile (PUC), Santiago, Chile.
Industrial Engineer with specialization in Mathematical Engineering, *distinción máxima* (highest honors), 2014.
BS, Science of Engineering, 2013.

Research experience

Research Associate (remote), July 2025–Present
Department of Linguistics
Northwestern University, Evanston, Illinois, USA
Mentor: Jennifer Cole.

Olga Taussky and John Todd Instructor in Mathematics, October 2021–December 2024
Division of Physics, Mathematics and Astronomy.
California Institute of Technology, Pasadena, California, USA.
Mentor: Matilde Marcolli.

Postdoctoral associate, September 2019–August 2020.
Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany.
Mentor: Nihat Ay. Group: Information Theory of Cognitive Systems.

Doctoral researcher, 2015–August 2019.
Temporary Lecturer (ATER), 2018–2019.
IMJ-PRG: Mathematics Institute of Jussieu–Paris Rive Gauche, France.
Adviser: Daniel Bennequin.

Research assistant, 2012–14.
ANESTOC: Center for Stochastic Analysis and Applications, Santiago, Chile.
Adviser: Rolando Rebolledo.

Research intern, January–March, 2013.
INRIA, Paris, France.
Adviser: Nicolas Broutin. Group: RAP (Networks, Algorithms and Probabilities).

Teaching and mentoring

Department of Mathematics, California Institute of Technology, Pasadena, CA, USA.
Mentored undergraduate students:

- Stephanie Chen (then sophomore): guided reading credits in winter and spring terms 2022; undergraduate research project in Summer 2022 ("Extension of Shannon Entropy to Finite Categories").
- Eric Paul (then freshman): guided reading credits in winter and spring terms 2022; undergraduate research project in Summer 2022 ("Information-theoretic limits to learning disentangled representations of data").
- Ryan Leal (then freshman): guided reading credits in spring term 2023; undergraduate research project in Summer 2023 ("Entropy from Group Orbits").
- Bharathan Sundar (then junior): guided reading in fall term 2023.
- Aman Burman (freshman, co-mentored with Matilde Marcolli): undergraduate research project in Summer 2024 ("Probing LLMs for encoded syntactic features").

Lecturer (sole instructor, full responsibility over all aspects of the course):

- *Ma140a Probability Theory*, winter 2022 and 2024 (upper-division/graduate, 16–18 students).
- *Ma4/104 Introduction to Mathematical Chaos*, spring 2022, 2023 and 2024 (undergraduate level, 3–8 students).
- *Ma151a Algebraic and Differential Topology*, fall 2022 (upper-division/graduate, 9 students).

Organizer:

- Learning seminar on Quantum Optimal Transport, Winter–Spring 2022.

Department of Mathematics, Université Paris Diderot – Paris 7, Paris, France.
Lecturer:

- *Elementary algebra and analysis I*, Fall 2017.

Teaching assistant:

- *Elementary algebra and analysis I*, Fall 2018 and Spring 2016.
- *Elementary algebra and analysis II*, Spring 2019 and Spring 2015.

Faculty of Engineering, Pontificia Universidad Católica de Chile (PUC), Santiago, Chile.
Teaching assistant and course coordinator:

- *Introduction to Mathematical Engineering* (Prof. Rolando Rebolledo), Spring 2014.
- *Stochastic models*, Fall 2011.

Teaching assistant:

- *Numerical Modeling in Engineering*, Fall 2012.
- *Mathematical Modeling in Engineering*, Spring 2012.

Faculty of Mathematics, Pontificia Universidad Católica de Chile (PUC), Santiago, Chile.
Teaching assistant, 2009–12.

- *Calculus I* (single variable calculus), *Calculus III* (vector calculus) and *Probability and Statistics*, several times each.

Service

Department of Mathematics, California Institute of Technology, Pasadena, CA, USA.

- "Information, Geometry, and Physics" seminar co-organizer, Winter 2023–Fall 2024 (see `https://www.its.caltech.edu/~vigneaux/igps/`).
- Graduate admissions application reviewer, Winter 2022 and Winter 2023.

- PhD candidacy committee member for Sitanshu Gakkhar.

IMJ-PRG: Institut de Mathématiques de Jussieu-Paris Rive Gauche, Paris, France.

- Doctoral students representative at the Institute's administrative council, 2017–2019.
- Doctoral students' seminar co-organizer, 2016–2018.

Publications

In most articles, the order of authors is alphabetical (the norm in mathematics). The exceptions are marked by †.

Peer-reviewed journals

T.D. Bradley and J.P. Vigneaux, "The magnitude of categories of texts enriched by language models," *Theory Appl. Categ.*, 44(37), 1256–1281, 2025.

S. Chen and J.P. Vigneaux, "A formula for the categorical magnitude in terms of the Moore-Penrose pseudoinverse." *Bull. Belg. Math. Soc. Simon Stevin* 30(3): 341–353, 2023.

J. P. Vigneaux, "Typicality for stratified measures," *IEEE Trans. Inf. Theory* 69(11): 6922–6940, 2023.

J.P. Vigneaux, "A characterization of generalized multinomial coefficients related to the entropic chain rule," *Aequationes Math.*, 97(2): 231–255, 2023.

J.P. Vigneaux, "Information structures and their cohomology," in *Theory Appl. Categ.*, 35(38): 1476–1529, 2020.

D. Bennequin and J.P. Vigneaux, "A functional equation related to generalized entropies and the modular group," *Aequationes Math.*, 94(6): 1201–1212, 2020.

J.P. Vigneaux, "Information theory with finite vector spaces," in *IEEE Trans. Inf. Theory*, 65(9): 5674–5687, Sept. 2019.

Peer-reviewed conference proceedings

† J.P. Vigneaux and J. Cole, "Probing neural audio codecs for distinctions among English nuclear tunes," In: *Proc. SpeechProsody 2026*, pp. 23–27, 2026.

S. Chen and J.P. Vigneaux, "Categorical magnitude and entropy." In: F. Nielsen and F. Barbaresco (eds), *Geom. Sci. Inform. GSI 2023.* Lect. Notes Comput. Sci., vol 14071. Springer, Cham., pp. 278–282, 2023.

J.P. Vigneaux, "On the entropy of rectifiable and stratified measures." In: F. Nielsen and F. Barbaresco (eds), *Geom. Sci. Inform. GSI 2023.* Lect. Notes Comput. Sci., vol 14071. Springer, Cham., pp. 338–346, 2023.

J.P. Vigneaux, "Entropy under disintegrations." In: F. Nielsen and F. Barbaresco (eds), *Geom. Sci. Inform. GSI 2021.* Lect. Notes Comput. Sci., vol 12829. Springer, Cham., pp. 340–349, 2021.

J.P. Vigneaux, "Information cohomology of classical vector-valued observables." In: F. Nielsen and F. Barbaresco (eds), *Geom. Sci. Inform. GSI 2021.* Lect. Notes Comput. Sci., vol 12829. Springer, Cham., pp. 537–546, 2021.

Preprints

† J.P. Vigneaux and J. Cole, "Steering intonational tunes in text-to-speech synthesis." Submitted to ICLR 2027; arXiv:2610.27403, 2026.

† J.P. Vigneaux, M. Kennedy, K. Iskarous, R. Frank and M. Marcolli, "Representation of syntax in LLMs through the lens of linear distance and similarity-aware entropy," arXiv:2608.27813, 2026. Under review at ACL Rolling Review.

R. Leal, J. Sun and J.P. Vigneaux, "Entropies associated with orbits of finite groups," arXiv:2512.02257, 2025. Submitted to *Information and Inference*.

J.P. Vigneaux, "A combinatorial approach to categorical Möbius inversion and pseudoinversion," arXiv:2407.14647, 2024.

D. Bennequin, O. Peltre, G. Sergeant-Perthuis, and J.P. Vigneaux, "Extra-fine sheaves and interaction decompositions", arXiv:2009.12646, 2020. Submitted to *Algebraic and Geometric Topology*; awaiting final decision by managing editors.

In preparation

† J.P. Vigneaux and J. Cole, "Intonational tunes in neural audio representations."

J.P. Vigneaux, "Curvature corrects contextual embeddings: second-order neighbourhood geometry and the categorical–continuous structure of word meaning."

Presentations

Contributed talk "Ungrading measure-theoretic probability: weekly reports, portfolios, and self-assigned grades in an upper-division course" in the AMS Special Session on Alternative Grading and Assessment in Math Education at the *Joint Mathematics Meetings 2027*, Chicago, USA, January 15, 2027.

Contributed talk "Stabilizer codes and entropies associated with orbits of finite groups" in the AMS Contributed Paper Session on Group Theory at the *Joint Mathematics Meetings 2027*, Chicago, USA, January 14, 2027.

Contributed talk "Curvature corrects contextual embeddings: Ollivier–Ricci flow on the hypersphere and the categorical–continuous structure of word meaning" in the AMS Special Session on Algebraic, Topological and Geometric Methods in Machine Learning at the *Joint Mathematics Meetings 2027*, Chicago, USA, January 14, 2027.

Contributed talk "Probing neural audio codecs for distinctions among English nuclear tunes" at the conference *Speech Prosody 2026*, Philadelphia, USA, May 26, 2026.

Invited talk "The geometry of syntax processing in transformer-based language models" at the conference *Geometry, topology and machine learning*, Sorbonne Université, Paris, France, May 20, 2026.

Invited talk "Modelos de prosodia y significado en lingüística y su representación en redes neuronales" ("Models of prosody and meaning in linguistics and their representation in neural networks") at Playa Ancha University, Valparaiso, Chile, December 16, 2025.

Invited talk "Mathematics of LLMs" at the workshop *Mathematical models of generative linguistics*, Caltech, Pasadena, USA, May 31, 2025.

Invited talk, "A combinatorial approach to categorical Möbius inversion and magnitude" at the Applied Algebraic Topology Research Network's seminar (online), April 17, 2024.

Invited talk, "Cohomological aspects of information" at the Topos Institute's Colloquium (online), January 25, 2024.

Contributed talk, "A combinatorial approach to Möbius inversion and pseudoinversion" at the conference *Magnitude 2023*, Osaka University, Osaka, Japan, December 7, 2023.

Invited talk "New links between information and geometry" at Centre Lagrange, Paris, France, September 11, 2023.

Invited talk "New links between information and geometry" at the Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany, September 8, 2023.

Invited talk "Typicality for stratified measures" at ETH, Zurich, Switzerland, September 6, 2023.

Contributed talk "On the entropy of rectifiable and stratified measures" at the conference *Geometric Science of Information 2023*, Saint Malo, France, August 30, 2023.

Contributed talk "Categorical Magnitude and Entropy" at the conference *Geometric Science of Information 2023*, Saint Malo, France, August 30, 2023.

Contributed talk "Typicality for stratified measures" at the Information, Geometry and Physics Seminar, Caltech, Pasadena, USA, April 19, 2023.

Invited talk "Quantum optimal transport" at the CMS SIAM Chapter, Caltech, Pasadena, USA, April 7, 2023.

Contributed talk "Categorical magnitude and entropy" at course Ma 20: Frontiers in Mathematics, Caltech, Pasadena, USA, November 1, 2022.

Contributed talk "Optimal transport on certain principal bundles" at the conference *PIMS-IFDS- NSF Summer School on Optimal Transport*, University of Washington, Seattle, USA, June 23, 2022.

Contributed talk "Information cohomology of classical vector-valued observables" at the conference *Geometric Science of Information 2021*, Paris, France, July 23, 2021.

Contributed talk "Entropy under disintegrations" at the conference *Geometric Science of Information 2021*, Paris, France, July 22, 2021.

Invited talk "Functors on posets, extra-fine sheaves, and interaction decompositions" at the seminar *Algebraic and Combinatorial Perspectives in the Mathematical Sciences* (online), October 30, 2020.

Invited talk "Cohomology of statistical systems" at Robert Ghrist's seminar, University of Pennsylvania, Philadelphia, PA, USA, October 23, 2019.

Invited talk "Cohomology of statistical systems" at the *Category Theory Seminar*, Johns Hopkins University, Baltimore, MD, USA, October 22, 2019.

Invited talk "Information cohomology: an overview" at *OASIS: The Oxford Advanced Seminar on Informatic Structures*, Oxford, England, January 25, 2019.

Invited talk "Une introduction à la topologie de l'information" at the seminar *Higher categories, polygraphs and homotopy*, Université Paris Diderot, Paris, France, September 28, 2018.

Contributed talk "Information theory associated to Tsallis' 2-entropy" at the conference *Latin American Week on Coding and Information*, Campinas, Brazil, July 26, 2018.

Invited talk, "Entropy and combinatorics" at the colloquium of the Mathematics Institute, Universidad de Talca, Talca, Chile, April 23, 2018.

Contributed talk "Information topology and probabilistic graphical models" at the conference *Applied Algebraic Topology*, Sapporo, Japan, August 8, 2017.

Invited talk "Cohomologie de l'information" at the seminar *Geometry and mathematical physics*, Université Paris Diderot, Paris, France, November 4, 2016.

Grant proposals

Sony Research Award Program 2026 (Focused Research Award), "Intonation steering in speech generation models." J. Cole (PI), J.P. Vigneaux (co-PI). Submitted September 2026; under review.

Open Philanthropy (AI safety), "Features from internal operations." Submitted 2025; not funded. Judged by the funder to "represent a valuable contribution to AI safety" and entered into a randomized allocation of the remaining grants, in which it was not selected.

Awards and scholarships

PGSM Masters Scholarship, Fondation Sciences Mathématiques de Paris, France, 2014.
An international academic scholarship that covered the full cost of studying one year in Paris.

Padre Hurtado award, Pontificia Universidad Católica de Chile, Chile, 2008.
A university-wide academic scholarship that covered the tuition fees of undergraduate education.

Additional research-related experiences (selected)

Workshop "Neural Coding and Combinatorics,"
ICERM, Providence, RI, USA, 2023.

International Congress of Mathematicians,
Rio de Janeiro, 2018.

European Talbot Workshop: "Topological aspects of quantum field theories,"
Winterberg, Germany, 2016.

Course "Disruptive technologies and public policies,"
Master 2 École d'Affaires Publiques, Sciences Po, Paris, France, Fall 2015.

Complex Systems Summer School - Chile,
Santa Fe Institute – Universidad del Desarrollo, Zapallar, Chile, 2013.

Languages

First language: Spanish

Other languages: English (level C1), French (level C1), German (level B1)

Programming Skills

- **Programming:** Python (advanced), Bash, Git, LaTeX.
- **Deep learning:** PyTorch, Hugging Face Transformers and Hub, TransformerLens. Worked with pretrained audio encoders (Mimi, WavLM, SpeechTokenizer), text-to-speech models (Sesame CSM, Fish Audio S2) and LLMs (BERT, RoBERTa, GPT-2/J, ...).
- **Speech processing:** measuring pitch (F0) and prosody with Praat/Parselmouth, forced alignment with the Montreal Forced Aligner, handling audio files with torchaudio.
- **Data analysis and statistics:** NumPy, pandas, SciPy, statsmodels, HDF5 (h5py); Matplotlib, Plotly, seaborn.
- **High-performance computing:** SLURM batch scheduling on a university cluster (Northwestern Quest), GPU jobs on A100s, conda/mamba environments.