

# Luciana Dias de Macedo

📍 Cambridge MA    ✉ lucianadiasdemacedo@gse.harvard.edu    ☎ +1 (609) 473 0106

🐙 lucianadmacedo.github.io    in lucianadmacedo    🔗 lucianadmacedo

## Bio

---

I specialize in computational linguistics and NLP, with a focus on enhancing language models. My work involves fine-tuning machine learning models, annotating linguistic data, and creating tools for language analysis. I'm committed to exploring language complexity and improving natural language understanding through data-driven insights, quantitative and qualitative analyses, and corpus linguistics.

## Education

---

**Harvard Graduate School of Education** *current: 2025 - 2026*  
*Ed.M. candidate in Learning Design, Innovation, and Technology*

**Universidade Federal de Minas Gerais** *Mar 2016 – Feb 2018*  
*MA in Applied Linguistics GPA: 3.95*

- **Thesis:** ([Lexical Bundles Across Sections Of Applied Linguistics Research Articles](#) 🔗)

**Universidade Federal de Minas Gerais** *Mar 2006 – Mar 2010*  
*BA in English GPA: 3.34*

- **Coursework:** second/first language acquisition, foreign language learning/teaching, phonology, syntax, semantics, Brazilian Portuguese, French, etc

## Work Experience

---

**NLU developer** *São Paulo, SP (remote)*  
*Cerence - contractor* *10/21 - 02/23*

- development and testing of grammars for NLU (pt-BR and pt-PT)
- text manipulation and data processing
- localization of data
- accuracy and coverage improvement
- prepare and release models.

**Associate linguist** *São Paulo, SP*  
*Google - via Lionbridge* *10/18 - 08/20*

- linguistic expertise in developing NLP tools for Brazilian Portuguese.
- grammars development
- creation of templates for Natural Language Understanding
- evaluation of current system outputs
- annotation and review of linguistic data
- labelling text for disambiguation, expansion, and text normalization
- R&D, and experimental projects.

**Language assessor** *Paris, France (remote)*  
*Pipplet* *08/18 - 08/20*

- assessment comprehending the 4 skills of Brazilian Portuguese
- Common European Framework of Reference for Languages (CEFR)
- language testers are either L1 or L2 speakers

**English and Portuguese (additional languages) teacher** *Belo Horizonte, MG/*  
*Rizvi Intl Schools* *Equatorial Guinea*

- Customized classes designed for business students from diverse national backgrounds

**Technical writer***LexisNexis Risk Solutions**São Paulo, SP**08/18 - 10/18*

- Writing, translation, editing, and proofreading of IT and Software manuals and procedures
- **En-Pt** and **Pt-En**

**Publications**

---

Erjavec, T. et al. *ParlaMint: Comparable Corpora of European Parliamentary Data*. [🔗](#) CLARIN, 2021.

Chapter: Dutra, D., Queiroz, J., Macedo, L., Costa, D., Mattos, E.: *Adjectives as Nominal Pre-modifiers in Chemistry and Applied Linguistics in Research Articles* [🔗](#) - **Advances in Corpus-based Research on Academic Writing: Effects of discipline, register, and writer expertise**: edited by Ute Römer, Viviana Cortes, Eric Friginal, John Benjamins Publishing Company, 2020.

Macedo, L., Costa, D., *Lexical Bundles Conveying Modality in Argumentative Essays of EAP Students: a corpus-based study* [🔗](#). ANAIS DO X CONGRESSO INTERNACIONAL, ABRALIN, pp. 2068-2076, 2017.

**Most Recent Projects**

---

**LEX-BR***repo:LEX-BR* [🔗](#)

working on the annotation of a legal text in Brazilian Portuguese (LEX-BR-Ius) with Professor Lucia Ferrari from UFMG, using UDPipe - our aim is to investigate diachronic language change since 1943 (the creation of the document) until nowadays.

**Academic writing in the humanities***repo:aacL2024* [🔗](#)

volunteer researcher in the analysis of academic writing in the humanities under the supervision of Professor Deise Dutra.

**Skills**

---

**Programming and Technical:** Python, Git (Version Control), Linux, Regular Expressions (regex), LaTeX, Excel

**Natural Language Processing (NLP):** Corpus Linguistics, Data Annotation, Data Analysis, Model Releases

**Languages and Teaching:** Brazilian Portuguese, English

**Grammar and Modeling Frameworks:** JSGF (Java Speech Grammar Format), BNF (Backus-Naur Form)