

LUCA DUCCESCHI

Summary

Senior Post-Doc researcher with 5 years of demonstrated experience in the field of Computational and Theoretical Linguistics, NLP, and academic teaching. My skills include advanced Python and Bash programming, natural language model training, language dataset building and development of educational resources.

Relevant Academic Experience

Post-Doc - 10/2018 to 10/2020
DIPSCO, University of Trento

I manage several research projects across various departments, and I am in charge of the HCI-NLP lab, a research laboratory aimed at increasing students' involvement and experimental experience in the fields of HCI and NLP. My research combines linguistic data collection and annotation, customization of NLP pipelines, and language models development. In the past two years, part of my time has been devoted to a dissemination effort to make such models more accessible to a non-expert via HCI practices.

of my current projects:

- Italo, a Python module for the preprocessing of Italian texts.
- Formal-IT-y: a hand-annotated Italian dataset for stylistic research and model training.
- Transcription and annotation of a CHAT compliant clinical speech database, in collaboration with CERIN and the ODF-lab on the University of Trento. The project goal is to create a hand-annotated corpus of semi-structured interviews, collected during several clinical studies, and use it to train language models capable of informing the diagnosis or the rehabilitation process.
- Development of an NLP pipeline for the Communic application, developed by FBK. Communic is a digital storytelling platform used in several educational projects like [Dico Di No](#).
- Qualitative economics: creation of a model capable of predicting the value of investments, leveraging on the linguistic content of annual reports of companies traded on NYSE and NASDAQ. In collaboration with Massironi Investments.

Post-Doc - 06/2014 to 06/2017

CIMEC, University of Trento

I co-created and developed the hands-on game Puzz-Ling, an educational game-with-a-purpose aimed at bringing syntactic research to a non-academic audience. The game is a physical implementation of a subset of x-bar rules that apply to Italian, English, and German. My role implied:

- Investigate the syntactic rules to be implemented.
- Design the mechanics of the game using a CAD software, and follow prototypes production.
- Disseminate the project in schools, conferences and other venues.

I also collaborated to the *Le Parole nel Tempo* (Words in Time) dissemination seminar at the Genova Science Fair 2014, for which I trained 15 Word2Vec models on several time framed Italian corpora.

Education

Ph.D.: **Linguistics** – 2012

University of Verona – University of Utrecht

Thesis Title: Talking about This and That – Anaphora in the Languages of India.

Focus: field-work, data collection, syntax and semantics.

MA: **Foreign Languages and Linguistics** - 2006

University of Bologna, Italy.

Thesis: A morphological analyzer for Italian using finite-state automata.

Focus: linguistics, morphology, computational linguistics.

Languages

Italian – Native

English – C2

French – C1

CS Skills

- **Python**: skilled programmer, knowledge of the main scientific modules, such as Numpy, Scipy, Scikit-Learn, Pandas, Spacy, Gensim, Pytorch, NLTK, Stanza, Flask. Currently undergoing PCAP certification.
- **Bash**: skilled programmer with extensive experience in Unix server management.
- **Git**: advanced understanding of how to manage project repositories, package modules and maintain projects:
 - <https://github.com/lucaducceschi>
 - <https://gitlab.com/luca.ducceschi>
- **Cloud computing**: expert user of Colaboratory, managing data intensive projects. Intermediate experience with AWS micro-services architecture, Azure cloud platform.

Teaching Experience

- **2017-2020:** Computational Skills for Text Analysis, Master in Language and Multimodal Interaction, CIMEC, University of Trento.
- **2015-2019:** Teorie e Tecniche di Riconoscimento, Bachelor in Interfaces and Communication Technology (Natural Language Processing and Machine Learning), DIPSCO, University of Trento.
- **2016-2019:** Computational Linguistics, Interfaces and Communication Technology, DIPSCO, University of Trento.
- **2018-2020:** Latex for Academic publishing and Python for Clinical Psychology, Ph.D. seminars, DIPSCO, University of Trento.
- **2010:** Applied Linguistics, Master in Language, Culture and Society, LLSM, University of Bologna

Publications

- Godler A., Ducceschi L., Zancanaro M., **2020**, *A Language-based Interface for Analysis of Digital Storytelling*. Paper accepted for AVI 2020 (International Conference on Advanced Visual Interfaces).
- Steltmann, J., Ducceschi, L., Herbelot, A., **2019**, *How much competence is there in performance?* Assessing the distributional hypothesis in word bigrams. In Clic-It 2019.
- Apro시오, A. P., Menini, S., Tonelli, S., Ducceschi, L., Herzog, L., **2018**, *Towards Personalized Simplification based on L2 Learners' Native Language*. In Clic-It 2018.
- Roberto Zamparelli and Luca Ducceschi, **2018**, *Puzz-Ling: a multi-lingual hands-on game to teach linguistics*. Language and Linguistic Compass.
- Poesio, M., Chamberlain J., Kruschwitz U., Robaldo L., Ducceschi L., **2013**. *Phrase Detectives: Utilizing Collective Intelligence for Internet-Scale Language Resource Creation*. ACM Transactions on Intelligent Interactive Systems, Vol 3, Issue 1.
- Robaldo, L., M. Poesio, L. Ducceschi, J. Chamberlain, U. Kruschwitz, **2011**. *Italian Anaphoric Annotation with the Phrase Detectives Game-With-A-Purpose*. In Proceedings Of 12th Congress of the Italian Association for Artificial Intelligence.