



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
DIPARTIMENTO DI FILOLOGIA CLASSICA
E ITALIANISTICA

“Come scrivevamo”. Costituzione di un corpus di prove scritte di italiano degli esami di maturità del periodo 2013-2026

Progetto finanziato da Zanichelli

Tutor: Prof. Matteo Viale

1. Scopo del progetto e ruolo dell’assegnista

La valutazione delle competenze linguistiche degli studenti, sia in prospettiva sincronica che diacronica, presenta sempre il forte rischio di valutazioni soggettive e permeabili a numerosi pregiudizi. Se alcune ricerche concluse o in corso (IPRASE 2020; Grandi 2025, Marroni 2022) stanno cercando di far luce sulle effettive abilità di scrittura degli studenti al termine della secondaria di secondo grado o universitari, disponiamo, però, di scarsissimi dati per le competenze del passato.

Questa lacuna è stata colmata dal progetto “Come scrivevamo”, promosso dal Dipartimento di Filologia Classica e Italianistica sotto la direzione scientifica di Matteo Viale e finanziato da Zanichelli, che ha consentito di costituire un corpus di prove scritte di italiano degli esami di maturità rappresentativo dei diversi decenni del periodo in esame (1970-2012), dei tipi di scuola superiore (liceo o istituto tecnico) e delle macroaree geografiche, per un totale di circa 4.000 prove. La costituzione del corpus è stata resa possibile da un censimento degli archivi di una trentina di scuole superiori che hanno aderito al progetto e a Zanichelli, mentre la digitalizzazione e metadatazione avverrà all’interno del laboratorio AD Lab del Dipartimento di Filologia Classica e Italianistica dell’Università di Bologna.

La presentazione dei risultati – la cui elaborazione approfondita è ancora in corso - ha suscitato un ampio interesse e portato Zanichelli a finanziare la prosecuzione del progetto per integrare il corpus di prove di maturità anche per il periodo 2013-2026, così da effettuare confronti con gli anni più recenti, caratterizzata dalla sospensione delle prove scritte per due anni scolastici a causa della pandemia legata al Covid19.

Obiettivo di questa successiva fase del progetto è digitalizzare un subcorpus di circa ulteriori 3000 prove per il periodo successivo a quello alla prima parte del progetto, arrivando a documentare la situazione dell’ultimo decennio. Ciò consentirà di comprendere se nel periodo più recente siano presenti fenomeni che tendono a stabilizzarsi oppure a scomparire rispetto ai decenni precedenti,

se gli errori rilevati siano tali da inficiare la comprensione del testo e siano sintomi di una scarsa padronanza della lingua da parte degli studenti, quali scelte prediligessero gli studenti nel passato e quali, invece, si preferiscano ora, nell'ottica di una maggiore comprensione del percorso compiuto negli anni dall'italiano scolastico.

In questo quadro, rientrano tra i **compiti del borsista**:

- provvedere alla digitalizzazione, metadatazione e anonimizzazione delle prove sotto la guida del personale di AD Lab del FICLIT;
- provvedere alla trascrizione un campione di prove e alla predisposizione dei testi digitalizzati per l'analisi quantitativa con appositi software di analisi linguistica sotto la supervisione del responsabile scientifico del progetto;
- provvedere ad altre incombenze legate alla marcatura e costituzione del corpus linguistico.

2. Fasi del progetto e piano di attività

Il progetto – e di riflesso il piano delle attività del borsista – si articola in diverse fasi scandite secondo il seguente calendario:

- A) novembre-dicembre 2026: Raccolta dei plichi delle prove presso il FICLIT.
- B) dicembre-febbraio 2026: Digitalizzazione, metadatazione e anonimizzazione delle prove.
- C) febbraio 2026: Avvio della trascrizione del campione di prove secondo le indicazioni fornite dal responsabile scientifico del progetto.

Successivamente, i ricercatori coinvolti nel progetto procederanno allo studio linguistico del corpus così costituito e al confronto con quanto emerso nei decenni precedenti.

3. Bibliografia

Berruto G. (1987), *Sociolinguistica dell'italiano contemporaneo*, Carocci, Roma.

D'Achille P. (2010), *L'italiano contemporaneo*, Il Mulino, Bologna.

De Mauro T. (1970), *Storia linguistica dell'Italia unita*, Laterza, Bari-Roma.

De Mauro T. (2014), *Storia linguistica dell'Italia repubblicana: dal 1946 ai nostri giorni*, Laterza, Bari-Roma.

De Mauro T. (2018), *L'educazione linguistica democratica*, Laterza, Bari-Roma.

Galfré M. (2017), *Tutti a scuola!*, Carocci, Roma.

Genovesi G. (1998), *Storia della scuola in Italia dal Settecento a oggi*, Laterza, Bari-Roma.

Grandi N. (a cura di) (2025), *L'italiano scritto degli studenti universitari. Quadro sociolinguistico, tendenze tipologiche, implicazioni didattiche*, Milano, Franco Angeli.

INVALSI (2012), *Rilevazione degli errori più diffusi nella padronanza della lingua italiana nella prima prova di italiano*, Rapporti Invalsi.

IPRASE (2020), *Come cambia la scrittura a scuola. Rapporto di ricerca*, a cura di Ruele M. e Zuin E., Iprase, Trento.

ISTAT (2012), *L'Italia in 150 anni. Sommario di statistiche storiche 1861-2010*, Istat, Roma, pp. 339-398.

Marroni S. (2022), *La forza della norma*, Firenze, Cesati, 2022

“Come scrivevamo” ["How we wrote"]. Creation of a corpus of written Italian tests from the final exams for the period 2013-2026

Project financed by Zanichelli

Tutor: Prof. Matteo Viale

1. Purpose of the project and role of the research fellow

Assessing students' language skills, whether synchronic or diachronic, always carries a high risk of being subjective and susceptible to numerous biases. While some completed or ongoing research (IPRASE 2020; UNIVERS-ITA Marroni 2022) is attempting to shed light on the actual writing skills of students completing secondary school or university, we have very little data on past skills.

This gap was filled by the "Come scrivevamo" project, promoted by the Department of Classical Philology and Italian Studies under the scientific direction of Matteo Viale and funded by Zanichelli, which resulted in the creation of a corpus of written Italian exam tests representative of the various decades of the period under review (1970-2012), types of secondary school (lyceum or technical institute), and geographic macro-areas, for a total of approximately 4,000 tests. The corpus was created through a census of the archives of around thirty high schools that participated in the project and Zanichelli. Digitization and metadata collection will be carried out within the AD Lab of the Department of Classical Philology and Italian Studies at the University of Bologna.

The presentation of the results—the in-depth analysis of which is still ongoing—aroused widespread interest and led Zanichelli to fund the continuation of the project to integrate the corpus of high school diploma exams for the period 2013-2026, allowing comparisons with more recent years, which were characterized by the suspension of written exams for two school years due to the Covid-19 pandemic.

The goal of this next phase of the project is to digitize a subcorpus of approximately 3,000 additional tests for the period following the first part of the project, thus documenting the situation over the last decade. This will allow us to understand whether, in the most recent period, there are phenomena that tend to stabilize or disappear compared to previous decades; whether the errors detected are such as to impair text comprehension and are symptoms of students' poor command of the language; and what choices students favored in the past and what they prefer now, with a view to better understanding the evolution of school Italian over the years.

Within this framework, the tasks of the fellow include:

- provide for the digitization, metadating and anonymization of the tests under the guidance of the FICLIT AD Lab staff;
- provide for the transcription of a sample of tests and the preparation of digitized texts for quantitative analysis with special linguistic analysis software under the supervision of the scientific manager of the project;
- provide for other duties related to the marking and constitution of the linguistic corpus.

2. Description of the project and activity plan

The project—and consequently the fellow's activity plan—is divided into several phases, organized according to the following schedule:

A) November-December 2026: Collection of test packages at FICLIT.

B) December-February 2026: Digitization, metadating, and anonymization of the tests.

C) February 2026: Beginning of transcription of the test sample according to the instructions provided by the project's scientific director.

Subsequently, the researchers involved in the project will proceed with the linguistic study of the corpus thus constituted and compare it with what emerged in previous decades.

3. Bibliography

Berruto G. (1987), *Sociolinguistica dell'italiano contemporaneo*, Carocci, Roma.

D'Achille P. (2010), *L'italiano contemporaneo*, Il Mulino, Bologna.

De Mauro T. (1970), *Storia linguistica dell'Italia unita*, Laterza, Bari-Roma.

De Mauro T. (2014), *Storia linguistica dell'Italia repubblicana: dal 1946 ai nostri giorni*, Laterza, Bari-Roma.

De Mauro T. (2018), *L'educazione linguistica democratica*, Laterza, Bari-Roma.

Galfré M. (2017), *Tutti a scuola!*, Carocci, Roma.

Genovesi G. (1998), *Storia della scuola in Italia dal Settecento a oggi*, Laterza, Bari-Roma.

Grandi N. (a cura di) (2025), *L'italiano scritto degli studenti universitari. Quadro sociolinguistico, tendenze tipologiche, implicazioni didattiche*, Milano, Franco Angeli.

INVALSI (2012), *Rilevazione degli errori più diffusi nella padronanza della lingua italiana nella prima prova di italiano*, Rapporti Invalsi.

IPRASE (2020), *Come cambia la scrittura a scuola. Rapporto di ricerca*, a cura di Ruele M. e Zuin E., Iprase, Trento.

ISTAT (2012), *L'Italia in 150 anni. Sommario di statistiche storiche 1861-2010*, Istat, Roma, pp. 339-398.

Marroni S. (2022), *La forza della norma*, Firenze, Cesati, 2022