

Research question

Considering that euphonic phenomena are strongly dependent on the tastes of an era, how can we create a program to automatically compute their frequency and therefore understand whether other factors influence this stylistic choice?

Introduction

The Italian language includes a series of euphonic phenomena used to avoid cacophony or difficulties in pronunciation. The letters involved are:

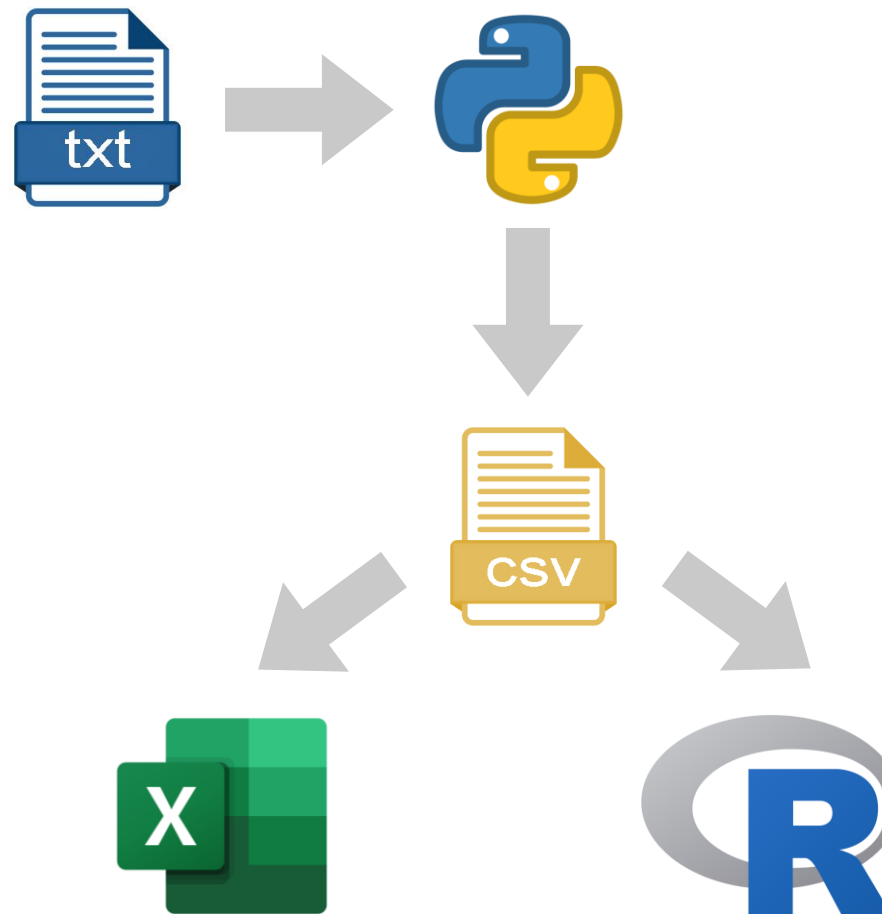
- **-d:** *ad, ed, od, ned, sed* and *ched*;
- **-i-:** “Dove vanno tutti gli altri? Prima di tutto, andremo in istrada” (Manzoni, *I promessi sposi*, cap. XXIX);
- **-r:** “[...] al lume d'una lucerna che ardeva sur un tavolino [...]” (Manzoni, *I promessi sposi*, cap. XXI).

The program

The program was coded using Python and it allows to:

- select a TXT file and automatically count each euphonic phenomenon's frequency;
- produce a CSV output that can be exported to Excel or R to conduct further analyses. It contains every collocation (b.1) along with its frequency (b.2);
- divide the results for the euphonic *d* into two different lists to separate the cases in which current general recommendations are not followed.

In this way, it is possible to double check, search for possible significant patterns and immediately see whether a text or an author employs these stylistic devices against the current recommendations.



	c		b.1		b.2		
CED	F	IED	F	ER	F	PI	F
ed ecco	15	ad essa	9	sur un	26	in ispiante	1
ad angolo	1	ad evidenza	1	sur una	33	non iscoprite	1
ed eccellentissimo	7	od avendolo	1	sur uno	2	per ispegnerla	1
ad altri	6	ad estirparlo	1			per ispiar	1
ad armare	1	ed accresciuta	1			per istinto	1
ad assicurarci	1	ad estermio	1			per iscansarli	2
ad aspettar	3	ad attestare	1			per iscoprire	3
ad arbitrio	1	ad estendere	1			un istante	4
ad altro	7	ad intendere	4			con ispavento	1
ad apparecchiare	1	ad un	3			non istaccarsi	1
ed esercitava	1	ad attaccar	1			in ismanie	1
ad alterarsi	1	ad uscire	1			per isbrigarci	1
ed entrava	1	ad ogni	3			per iscontare	1
ad aprire	5	ad implorar	3			in iscritto	3

Current and future developments

Creation of a corpus of literary works (in TXT) spanning 100 years (between the mid-18th and mid-19th centuries).

New research questions:

- What can we understand from a synchronic and diachronic perspective?
- If there are any, what other factors influence the presence or absence of these phenomena?

Bibliography

- D'Achille, P. (2011), *L'italiano contemporaneo*, Bologna: Il Mulino.
- Malagoli, G. (1912), *Ortoepia e ortografia*, Milano: Hoepli.
- Migliorini, B. and Folena, G. (1957), *Piccola guida di ortografia*, Olivetti.
- Treccani (2010), *Prontuario di dubbi e incertezze in Enciclopedia dell'Italiano*, volume II (M-Z), Istituto della Enciclopedia Italiana.
- Treccani (2012), *La grammatica italiana*, Istituto della Enciclopedia Italiana.