

1. Annotation

In the majority of cases, Latin grammars currently used by the educational institutions provide detailed descriptions of grammar and lexicon. However, a comparative analysis of these grammars reveals a significant amount of inconsistencies and discrepancies between the descriptions of the same linguistic features.

The present paper aims to resolve several inconsistencies present in a selection of grammars that was compiled specifically for this study. For a number of reasons (mostly simplicity and coherency), the paper concentrates only on the most prominent inconsistencies in the descriptions of nouns.

A corpus study is proposed as a primary method of solving these problems. The results of the study show how the supposedly optimal description of a feature may be written.

The paper shows that a corpus study is a viable method of resolving inconsistencies. Its results may also be used in the future grammars of Latin.

2. Description discrepancies

A. Example of discrepancies

➤ Latin 3rd declension: several inflexional types. The main ones: consonant stems, i-stems, adapted consonant stems (Bennett 1895).

➤ Consonant stems are more frequent: therefore, they have influenced i-stems. For example, evident with the **-em** ending in Acc. Sg. present in both types.

➤ However, some i-stem nouns may retain the **i** vowel in some endings, incl. Acc. Sg. (**-im**).

➤ Discrepancies: some grammars do not mention the **-im** ending. Different authors provide different lists of nouns that may have it.

B. Problems and objectives chosen for the study

Three significant noun discrepancies were chosen for the study:

- 3rd declension nouns: **-im** vs. **-em** in Acc. Sg.
- 4th declension nouns: **-u** vs. **-ui** in Dat. Sg.
- 4th declension nouns: **-ubus** vs. **-ibus** in Dat. Pl.

Objectives:

- Define the single optimal description for the feature.
- Define the hypothetic reasons behind the discrepancy

3. The corpus study

A. The corpus

➤ Based on “Classical Latin Texts” project originally created by the “Packard Humanities Institute”.

➤ CLT: includes every written Latin texts created before III AD and some late works, like Maurus Servius Honoratus. Only texts bigger than 16 KB were included in the corpus.

➤ The corpus: 98 text files, 7400000 tokens.

➤ Not optimal: not balanced, quotes, no addendum. However, further improvements are impossible.

B. Study tools: Python

Python 2 programming language was chosen as a primary tool to create the software needed for the corpus study.

The advantages:

- Easy to learn syntax, close to a natural language
- A significant amount of ready-to-use functions and libraries available
- Open-source

C. The descriptions selection

A representative selection of grammars with different languages and study schools included.

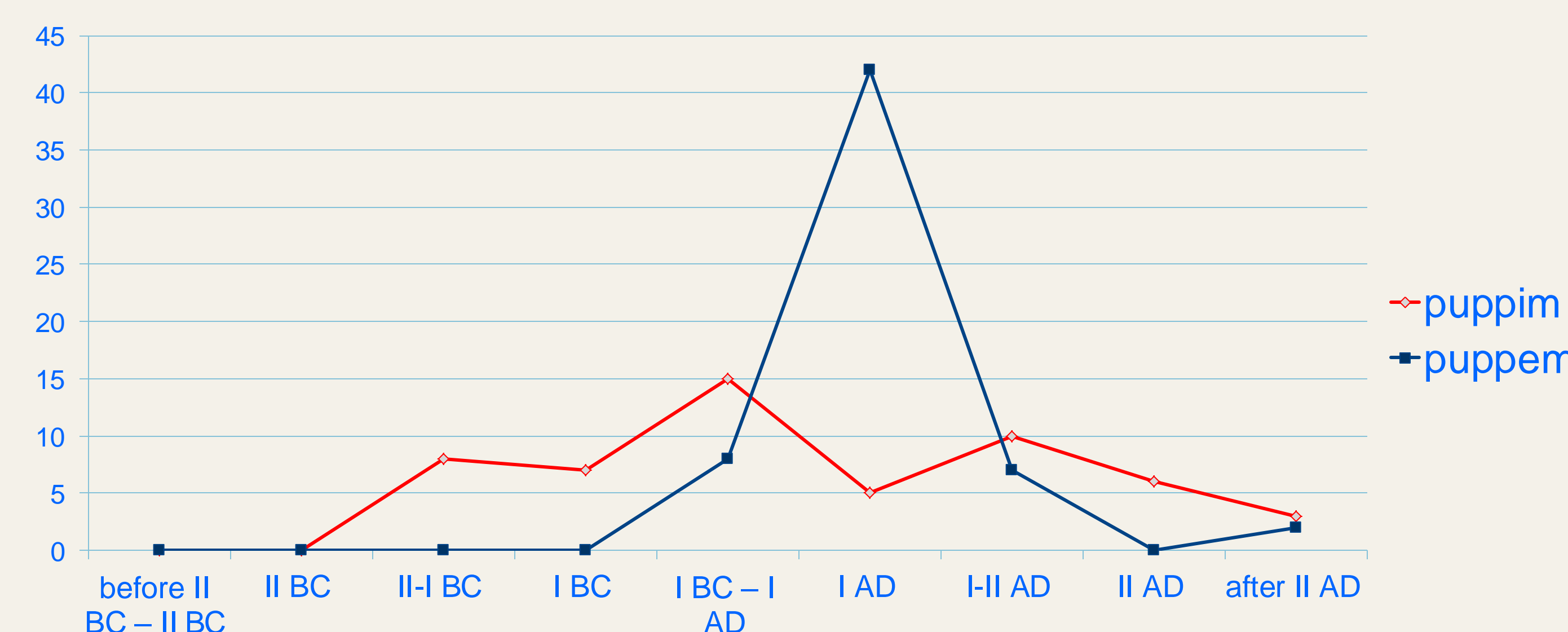
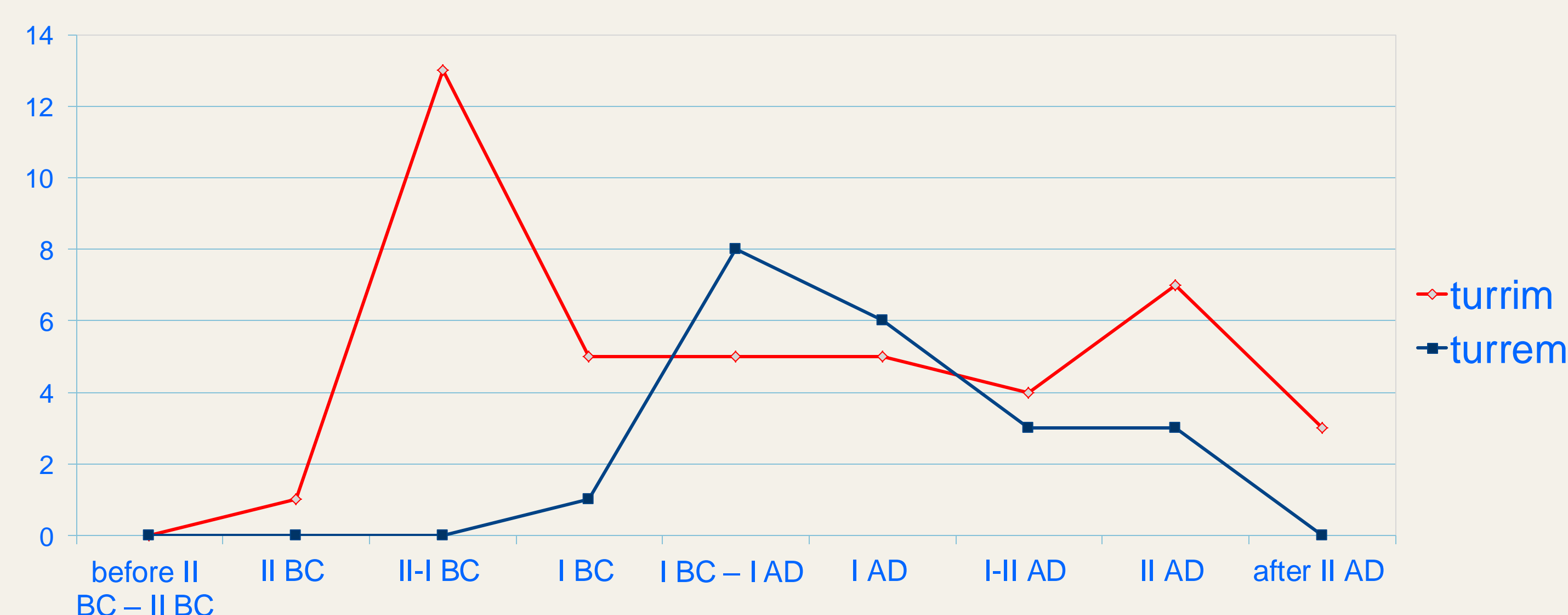
Russian selection: Sobolevsky 1948; Zaytsev et al. 1974; Borovsky, Boldyrev 1975; Yarkho, Loboda 1998; Tronsky 2001; Miroshenkova, Fyodorov 2003.

English selection: Bennett 1895; Greenough, Kittredge, Howard, D’ooge 1903; Morwood 1999; Wheelock, Lafleur 2005;.

Further improvements are hardly possible: no significant differences between other popular grammars.

4. How the study works: 3rd declension example

Lexeme	Frequency with -im	Frequency with -em
turris	49 (45)	28 (32)
puppis	57	60
tussis	106	1
febris	17	61
sitis	166	0



5. Results for 3rd declension nouns

➤ “Golden latin”: the **primary** reason for the discrepancies

➤ Different goals while writing the grammars: the **secondary** reason

➤ Frequencies: **-im** forms **not marginal**, should be listed in the grammars

➤ Frequencies: do *sitis* and *tussis* have the standard **-em** ending?

➤ Diachrony: most **-em** forms appear **after** “Golden latin”

➤ Diachrony: since **-im** could be oscan influence (Tronsky 2001: 169), could the **-em** form signify the **formation of Latin language integrity**?

References

- Tronsky 2001 – I.M. Tronsky, A.V. Grosheva, M.L. Kisilier, E.R. Kryuchkova, N.N. Nadel', A.I. Solopov, N.N. Kazansky. *Istoričeskaya grammatika latinskogo yazyka. Obšč'eindoeuropejskoe yazykovoe sostoyanie (voprosy rekonstrukcii)*. Moscow: “Indrik”.
- Bennett 1895 – C.E. Bennett. *A Latin Grammar*. Norwood: Norwood Press, 1895.