

## What you won't find in the annotated corpus

As pointed out in the Guidelines, the Latin literary texts included in the TBL corpus document words **of Greek origin written in Greek or in Latin script**. As regards the latter category, i.e., Graecisms written in Latin characters, we have decided to exclude them from our analysis **only if there is evidence that they are integrated into the Latin language**. As they are part of the Latin lexicon, from the point of view of our analysis they do not provide relevant information as to the degree of bilingualism of a text: rather, they have to be considered as exogenous elements that have permanently entered the target language. To give an example, the fact that Macrobius uses Greek words such as *stomachus* (< στόμαχος; Sat. 7, 15, 8), *philosophia* (< φιλοσοφία; Sat. 1, 7, 3), *epistula* (< ἐπιστολή; Sat. 2, 1, 14), or *purpura* (< πορφύρα; Sat. 3, 13, 12), which are (more or less) widely attested in the Latin literature, does not constitute evidence for the author's command of Greek, since these lexemes, although of Greek origin, were presumably no longer perceived as Greek by Latin speakers.

The methodology we have used to identify (and therefore to exclude from annotation) the integrated Graecisms which became part of the Latin lexicon has been to check whether a word of Greek origin written in Latin characters has been included in a reference dictionary, in this case specifically, Charlton T. Lewis and Charles Short's *A Latin Dictionary* (Oxford, Clarendon Press, 1879), available online on the Perseus website ([www.perseus.tufts.edu/hopper/](http://www.perseus.tufts.edu/hopper/)). If a word is included in this dictionary, we have systematically excluded it from our annotation process. Consequently, **of all the graphically Latinized words of Greek origin, only those that are not lemmatized in the dictionary have been tagged in the TBL corpus**.

Lastly, we have excluded from annotation the rare case of long portions of text written exclusively in Greek and found in particular sections of a given work (e.g., in Appendix), since they do not constitute code-switching phenomena in the proper sense of the term.