Math & CS Seminar
Author vs. Translator.
Who reigns the text? Applying computational stylistics methods to translated literary texts

George K. Mikros
Professor of Computational and Quantitative Linguistics
National and Kapodistrian University of Athens & University of Massachusetts, Boston

Authorship identification techniques have been extensively used for the attribution of texts in specific authors as long as these texts are produced originally from one of them in his/her mother-tongue. However, there is little experience in testing authorship identification methods in cases where the translator is the aim of the identification. Stylometric theory assumes that each author possess a distinct, unique “writeprint” which is expressed quantitatively through the idiosyncratic occurrence variation of its most frequent linguistic structures and various indices of unconscious linguistic behavior such as lexical “richness” formulas, word and sentence lengths etc. If such a “writeprint” exists, then it should be text topic and genre neutral. Translations test the theory of “writeprint” in its extreme. If the identity of the author survives through the process of translation and can be traced in a text that originally was written in another language then stylometric authorship attribution would increase its methodological robustness and reliability.

This talk will present a brief overview of the evolution of stylometry the last decades and will explain its basic methodological tools for extracting “writeprints” from texts (stylometric features, machine learning algorithms). Then it will turn to translations of classic Russian literature as an application area and will investigate whether computational stylistic methods can be successfully applied to uncover not only the original author but also the translator. The reported research results will be discussed under the general theoretical framework of quantitative linguistics.

Friday February 6th
3:00 – 3:50 pm in CH 446