



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



IQLA-GIAT Summer School in  
*Quantitative Analysis of Textual Data*  
University of Padua, 21-25 September 2015  
*2nd edition*

## Final Report

The second edition of the IQLA-GIAT Summer School in *Quantitative analysis of textual data* took place at the University of Padua, Department of *Philosophy, Sociology, Education and Applied Psychology* (FISPPA) – Sociology buildings, from Monday 21st to Friday 25th September 2015. The Summer School was funded by the University of Padua, coordinated by prof. Arjuna Tuzzi (University of Padua, Dept. FISPPA) and organized by GIAT – *Interdisciplinary Text Analysis Group* ([www.giat.org](http://www.giat.org)) in collaboration with IQLA – *International Quantitative Linguistics Association* ([www.iqla.org](http://www.iqla.org)) and the Department of *Computational Linguistics* of the University of Trier ([www.uni-trier.de](http://www.uni-trier.de)).

The teaching staff included 10 lecturers from 7 European universities of 6 different countries: Toulouse II (France), Trier (Germany), Athens (Greece), Neuchâtel (Switzerland), Kraców (Poland), Padua and Trieste (Italy); and an IT expert from the private sector. In terms of the number and quality of participants, the Summer School achieved a good success. The participants belonged to five scientific branches - linguistics, computer sciences, economics, and psychology - and came from 7 different countries: Croatia, Germany, Greece, Slovakia, Switzerland, Ukraine, and Italy. The participants' CVs were excellent and well suited for the Summer School objectives: in terms of their specific research interests, experiences and interdisciplinary viewpoints, the participants formed a strongly motivated and adequately skilled group, and this proved to be a great advantage for both the participants and the teaching staff.

### Teaching activities

The quantitative analysis of textual data is an object of research in many sectors of linguistics, computer sciences, mathematics and statistics and it is used as a research tool within a number of disciplines such as psychology, sociology, sociolinguistics, education, history, political studies, communication and media studies. The recent evolution of information technologies (IT) and quantitative methods has led to a number of distinct but interrelated sectors (e.g. computational linguistics, information retrieval, natural language processing, text mining, text analytics, sentiment analysis, opinion mining, etc.) with interesting industrial applications and, today, *text analysis* is an umbrella term referring to a wide range of qualitative, quantitative and mixed-methods approaches. Recent studies have repeatedly stressed the need for developing interdisciplinary approaches and the IQLA-GIAT Summer School is the ideal environment for sharing innovative analytical tools by pooling together the research methods of both Quantitative Linguistics and Statistical Analysis of Textual Data.

The IQLA-GIAT Summer School is characterized by three main elements:

1. a general part devoted to quantitative linguistics;
2. a special issue addressed to the problem of “measuring style”  
(style and computational stylistics is the main topic of 2015 edition, in the first edition of 2013 the main topic was authorship attribution)
3. several lab-sessions dedicated to computer-aided analysis of textual data and data visualization.

Teaching activities raised questions that can be answered thanks to quantitative methods implemented within a text analysis framework and other procedures aimed to identify and compare the characteristics of texts.

Teaching activities included lectures and lab sessions, as well as workshops illustrating software and tools. The lab sessions took place in a computer lab of the Department FISPPA. Each participant was assigned a PC in which all the necessary software packages were available (the limit of 20 participants in the Summer School was determined by the number of PCs in the lab).

At the end of the classes all participants filled in a questionnaire to assess the main aspects of the Summer School (e.g. organization, teaching, materials, facilities and equipment, expectations, satisfaction rate, suggestions etc.) and four of them decided to take a non-compulsory written test in order to receive credits (3 CFU). The organizers were very pleased with the assessment because the participants' opinions were very good and those who took the test passed it with excellent results.

### **Enjoyable details**

The Summer School program also included social events, i.e. two lunch breaks in the beautiful garden of the Department buildings, a lunch in a restaurant in Ponte Corvo (district of Padua), a pizza-party and a social dinner in two restaurants of the city centre, coffee breaks at a café near the Department for light refreshments.

All participants and teachers received a conference kit including gadgets (bag, folder, maps, pen drive, brochures and flyers about the town, about botanical gardens, monuments and historical sites); moreover, Routledge, the publisher of the Journal of Quantitative Linguistics, the official journal of IQLA, gave a present for all participants of the Summer School: a 15 days free-access to their journals of the Linguistics field.



*And, of course...  
see you at the next edition!  
Prepare yourself to be challenged, excited and inspired...*

### **Many thanks to...**

- University of Padua and its International Relations Office for funding this first edition of the IQLA-GIAT Summer School;
- Department FISPPA for the great support provided by technical and administrative staff;
- collaborators Francesca Bacchin and Stefano Sbalchiero for their precious support and relentless dedication to work;
- members of the GIAT group, IQLA association, and Department of Computational linguistics of the University of Trier for their support (and friendship);
- colleagues, scholars, and IT experts for joining the Summer School and sharing with participants their expertise, research experiences and knowledge;
- candidates for their application and selected participants for joining the Summer School and showing great enthusiasm and willingness to learn.

# Programme

2nd edition 2015

## LECTURES (14h)

**Rocco Coronato** (University of Padova, Italy)

*(Under)mining Metaphors in Shakespeare with a little help from Quintilian*

**Reinhard Köhler** (University of Trier, Germany)

*General Methodology in Empirical Linguistics. Evaluation of Data and Hypothesis Testing*

*A Crash Course in the Central Terms and Concepts of Science*

**Sven Naumann** (University of Trier, Germany)

*Data-driven approaches to syntactic processing*

*Syntactic complexity: quantitative models and measures*

**Stefano Ondelli** (University of Trieste, Italy)

*Compiling Corpora: Representativeness, Balance and Sampling*

**Jacques Savoy** (University of Neuchâtel Switzerland)

*Application of Quantitative Linguistics methods to US political speeches*

## LAB SESSIONS (18h)

**Andrei Beliankou** (University of Trier, Germany)

*Introduction in Programming with R*

*Quantitative Models of Linguistic Data*

**Maciej Eder** (University of Kraków, Poland)

*Stylometry with the package ‘Stylo’: Explanatory methods*

*Stylometry with the package ‘Stylo’: Supervised methods*

**George Mikros** (National and Kapodistrian University of Athens, Greece)

*Cross-linguistic authorship attribution: Detecting stylomes in translated documents and texts written by bilingual authors*

*Machine learning methods in computational stylistics using the caret R package*

**Paolo Nadalutti** (IT expert)

*Visualizing content analysis information using web based tools*

**Pierre Ratinaud** (University of Toulouse II, France)

*IRaMuTeQ: corpus indexation, manipulation and simple description*

*The Reinert method in IRaMuTeQ*

## OTHER ACTIVITIES (6h)

**Francesca Bacchin and Stefano Sbalchiero** (University of Padova, Italy)

*Welcome*

*Quality Assessment of the School*

*Written exam*

*Free Lab time*

*Tutorship*