

# 4D03 Computers&Linguistic Analysis (C01)

Academic Year: 2022

Term: Fall

Day/Evening: D

Instructor: Dr. Gemma Repiso Puigdelliura

Email: [repisopg@mcmaster.ca](mailto:repisopg@mcmaster.ca)

---

## Course Objectives

The objective of this course is to become familiar with the fundamentals of corpus linguistics and analysis of corpora using the R language. Students will be trained in conceptual and practical aspects of corpus analysis using R and the package `quanteda`. By the end of the course, students will be able to compile and annotate textual corpora, use a variety of corpus tools to perform searches and return lexical frequency lists, concordances, etc., visualize linguistic patterns in corpora, and carry out a small-scale-corpus-based research project.

---

## Textbooks, Materials & Fees

Lecture slides and additional resources will be made available on Avenue to Learn.

---

## Method of Assessment

30% - **Homework**. Each of these assignments will include practice questions to compile, analyze, and visualize corpus data. Students can work in pairs and have to write the name of their partner in the assignment.

15% **Labs assignments**. The class has a lab component in which students will have to solve directed questions on corpus analysis using R. Each week, students will have to submit the completed lab assignments. Reasonable attempts will be graded with 4 points, incomplete attempts will be graded with 2 points, and non-reasonable attempts (i.e., not submitted or attempts missing more than half of the answers) will be graded with a 0. The lowest-graded lab assignment will be dropped.

25% - **Midterm exam**. The exam will be cumulative and cover the topics until the week of the exam. The exam will consist of programming exercises similar to the exercises resolved in class.

30% **Final project**. The final paper/project for this course will be a small-scale-corpus-based research project. This project will allow students to use the tools learned in class to examine a research question of their interest in the field of linguistics. Students will have to pose research questions that can be answered with corpus linguistics research, compile a textual corpus, retrieve lexical frequencies, concordances, and visualize patterns in the corpus to answer their research questions.

5% **Final project presentation**. Students will present their final project in the class. This presentation will allow students to summarize their final project, get feedback on their analysis, and practice formal oral presentation skills.

---

## Policy on Missed Work, Extensions, and Late Penalties

Assignments submitted within 72 hours of the due date are penalized 15% of the value of the assignment. Being late more than 72 hours on an assignment results in a grade of zero for that assignment. If you submit an MSAF or other approved document (via the office of the Dean of your Faculty), contact Professor Repiso-Puigdelliura to arrange an alternate due date. Work missed for legitimate reasons is expected to be passed in at a later date as agreed with the instructor. Students can miss one lab assignment without penalty.

---

## Please Note the Following Policies and Statements:

### Academic Integrity

You are expected to exhibit honesty and use ethical behaviour in all aspects of the learning process. Academic credentials you earn are rooted in principles of honesty and academic integrity. It is your responsibility to understand what constitutes academic dishonesty.

Academic dishonesty is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage. This behaviour can result in serious consequences, e.g. the grade of zero on an assignment, loss of credit with a notation on the transcript (notation reads: "Grade of F assigned for academic dishonesty"), and/or suspension or expulsion from the university. For information on the various types of academic dishonesty please refer to the Academic Integrity Policy, located at <https://secretariat.mcmaster.ca/university-policies-procedures-guidelines/>

The following illustrates only three forms of academic dishonesty:

- plagiarism, e.g. the submission of work that is not one's own or for which other credit has been obtained.
- improper collaboration in group work.
- copying or using unauthorized aids in tests and examinations.

**Authenticity / Plagiarism Detection**

Some courses may use a web-based service (Turnitin.com) to reveal authenticity and ownership of student submitted work. For courses using such software, students will be expected to submit their work electronically either directly to Turnitin.com or via Avenue to Learn (A2L) plagiarism detection (a service supported by Turnitin.com) so it can be checked for academic dishonesty.

Students who do not wish to submit their work through A2L and/or Turnitin.com must still submit an electronic and/or hardcopy to the instructor. No penalty will be assigned to a student who does not submit work to Turnitin.com or A2L. All submitted work is subject to normal verification that standards of academic integrity have been upheld (e.g., on-line search, other software, etc.). To see the Turnitin.com Policy, please go to [www.mcmaster.ca/academicintegrity](http://www.mcmaster.ca/academicintegrity).

**Courses with an On-Line Element**

Some courses use on-line elements (e.g. e-mail, Avenue to Learn (A2L), LearnLink, web pages, capa, Moodle, ThinkingCap, etc.). Students should be aware that, when they access the electronic components of a course using these elements, private information such as first and last names, user names for the McMaster e-mail accounts, and program affiliation may become apparent to all other students in the same course. The available information is dependent on the technology used. Continuation in a course that uses on-line elements will be deemed consent to this disclosure. If you have any questions or concerns about such disclosure please discuss this with the course instructor.

**Online Proctoring**

Some courses may use online proctoring software for tests and exams. This software may require students to turn on their video camera, present identification, monitor and record their computer activities, and/or lockdown their browser during tests or exams. This software may be required to be installed before the exam begins.

**Conduct Expectations**

As a McMaster student, you have the right to experience, and the responsibility to demonstrate, respectful and dignified interactions within all of our living, learning and working communities. These expectations are described in the Code of Student Rights & Responsibilities (the "Code"). All students share the responsibility of maintaining a positive environment for the academic and personal growth of all McMaster community members, whether in person or online.

It is essential that students be mindful of their interactions online, as the Code remains in effect in virtual learning environments. The Code applies to any interactions that adversely affect, disrupt, or interfere with reasonable participation in University activities. Student disruptions or behaviours that interfere with university functions on online platforms (e.g. use of Avenue 2 Learn, WebEx or Zoom for delivery), will be taken very seriously and will be investigated. Outcomes may include restriction or removal of the involved students' access to these platforms.

**Academic Accommodation of Students with Disabilities**

Students with disabilities who require academic accommodation must contact Student Accessibility Services (SAS) at 905-525-9140 ext. 28652 or [sas@mcmaster.ca](mailto:sas@mcmaster.ca) e-mail to make arrangements with a Program Coordinator. For further information, consult McMaster University's Academic Accommodation of Students with Disabilities policy.

**Email correspondence policy**

It is the policy of the Faculty of Humanities that all email communication sent from students to instructors (including TAs), and from students to staff, must originate from each student's own McMaster University email account. This policy protects confidentiality and confirms the identity of the student. Instructors will delete emails that do not originate from a McMaster email account.

**Modification of course outlines**

The University reserves the right to change dates and/or deadlines etc. for any or all courses in the case of an emergency situation or labour disruption or civil unrest/disobedience, etc. If a modification becomes necessary, reasonable notice and communication with the students will be given with an explanation and the opportunity to comment on changes. Any significant changes should be made in consultation with the Department Chair.

**Request for Relief for Missed Academic Term Work****McMaster Student Absence Form (MSAF)**

In the event of an absence for medical or other reasons, students should review and follow the Academic Regulation in the Undergraduate Calendar "Requests for Relief for Missed Academic Term Work".

**Academic Accommodation for Religious, Indigenous and Spiritual Observances (RISO)**

Students requiring academic accommodation based on religious, indigenous or spiritual observances should follow the procedures set out in the RISO policy. Students should submit their request to their Faculty Office normally within 10 working days of the beginning of term in which they anticipate a need for accommodation or to the Registrar's Office prior to their examinations. Students should also contact their instructors as soon as possible to make alternative arrangements for classes, assignments, and tests.

**Copyright and Recording**

Students are advised that lectures, demonstrations, performances, and any other course material provided by an instructor include copyright protected works. The Copyright Act and copyright law protect every original literary, dramatic, musical and artistic work, including lectures by University instructors. The recording of lectures, tutorials, or other methods of instruction may occur during a course. Recording may be done by either the instructor for the purpose of authorized distribution, or by a student for the purpose of personal study. Students should be aware that their voice and/or image may be recorded by others during the class. Please speak with the instructor if this is a concern for you.

**Extreme Circumstances**

The University reserves the right to change the dates and deadlines for any or all courses in extreme circumstances (e.g., severe weather, labour disruptions, etc.). Changes will be communicated through regular McMaster communication channels, such as McMaster Daily News, A2L and/or McMaster email.

---

**Topics & Readings**

Introduction to corpus linguistics

R: Vectors and data frames

R: ifelse and for loops

R: Piping and plotting

Creating a corpus

Plotting with quanteda

R: Regular expressions

R: Regular expressions and corpus searches

Corpus analysis: concordances

Corpus analysis: frequencies

POS-tagging

---