

Linguistics 807 – Computational Linguistics

Course Syllabus

Department of Linguistics - Simon Fraser University – Spring Semester 2017

Instructor: **Dr. Maite Taboada**

Class: Tuesdays & Thursdays 10:30-12:20pm, AQ 3145

Office: RCB 8206 (may change soon)

Phone: 778-782-5585

E-mail: mtaboada@sfu.ca

Course Description

This course is an introduction to theoretical and applied issues in computational linguistics. Computational Linguistics (or Natural Language Processing) refers to the processing of human languages through computers, which is mostly an applied venture. It also covers theoretical aspects, such as the computational modeling of language processing phenomena, and it overlaps with research in psychology and neurolinguistics.

In this course, we will survey the field, and study a few specific applications. No previous knowledge of programming is required (although it would be an asset). Classes will consist of lectures, drawing on material from the main textbooks, but also from other sources (to be distributed throughout the semester). In each class, there will also be a practical part, where we will learn some basics of programming, and will experiment with existing natural language processing systems.

Required Textbooks

- Jurafsky, D. and J. H. Martin (2008). *Speech and Language Processing (2nd edition)*, Prentice Hall ISBN-13: 978-0131873216. Related web page [. \(Please make sure to get the second edition. There will be a third edition soon, but it's not available yet.\)](#).
- Bird, S., E. Klein and E. Loper (2009) *Natural Language Processing with Python*, O'Reilly Media. Print ISBN: 978-0-596-51649-9. Ebook ISBN: 978-0-596-80339-1. Related web page, and full online book:

Course expectations

1. Students are expected to **attend all classes and to arrive on time** so that classes may begin promptly. Announcements will be made at the beginning and end of classes regarding the assigned readings and the expectations for assignments.
2. Students are expected to **have read all assigned readings before class**. Because many students will be learning about a new field of study, some of the materials and concepts may seem fairly complex. You should probably read assigned readings and go over the lecture notes multiple times.
3. It is very important for students to **participate in class discussions and class activities**. Therefore, everyone must come to class prepared. Preparation includes carefully reading of the readings and assignments, noting any questions that you have.
4. Students will be responsible for **all materials** covered in the assigned readings and lectures. Assignments and exams will require students to refer to readings, lecture notes, and their own notes.
5. No late assignments will be accepted.

6. Following departmental policy, **make-up exams for in-class exams will not be given under any circumstances**. If you miss an exam, you will receive a grade of zero.
 - The only exception is if you are sick on the day of the exam. If you have to miss an in-class exam because of illness, you are required to contact me *prior* to the exam, if possible. When you return to class, I will need a note from your medical doctor specifying the date of your absence and the reason. There is a good chance that I will call the doctor to confirm the note.
 - If you have a doctor's note for the date of the exam, the weight of that exam will be re-allocated to the other course components.
 - Please check the following for further information on doctor's notes, and for a sample form that you can bring to your doctor: <https://www.sfu.ca/students/health/resources/faq/sick-notes.html>
7. Students will be **respectful** of other students and the instructor. In particular, students will not talk while the instructor or another student is talking.
8. Students should familiarize themselves with the **Linguistics Department's Standards** on Class Management and Student Responsibilities at:
<http://www.sfu.ca/linguistics/undergraduate/student-resources/department-standards.html>
9. **Cell phones**: please make sure that your cell phone (or any other noisy device, including alarms) is turned off, or at least silent, during class.
10. **Electronic devices**: You may use a laptop, tablet or similar device to take notes and work on class content. Please, make judicious use of electronic devices, as they can be distracting, to you and to other students. Do not use cell phones to text and do not use laptops and tablets to check online material that is unrelated to the class.
11. **Academic dishonesty** in all forms violates the basic principles of integrity and thus impedes learning. More specifically, academic dishonesty is a form of misconduct that is subject to disciplinary action and includes the following: cheating, fabrication, fraud, facilitating academic dishonesty, and plagiarism. For more information on academic honesty and student conduct, please visit the following websites:
<https://www.sfu.ca/students/calendar/2017/spring/fees-and-regulations/student-contract/academic-honesty.html>
<http://www.sfu.ca/policies/gazette/teaching.html>
A grade of "FD" may be assigned as a penalty for academic dishonesty.
12. Please note that students requiring accommodations as a result of a **disability** must contact the Center for Students with Disabilities (778-782-3112 or csdo@sfu.ca).
13. All student requests for accommodation for **religious practices** must be made in writing by the end of the first week of classes or no later than one week after a student adds a course.
14. This syllabus is **required reading**.

E-Mail policy – Use Canvas

Please, DO NOT send e-mail to my SFU address. Use the messaging system in Canvas instead.

1. I cannot always respond to messages late in the evening or on weekends. If you send a message late on Friday or during the weekend, I might not reply until the following week.

2. I can only respond to questions that can be answered in a sentence or two. Questions requiring longer replies should be asked in class or during office hours.
3. Please proof-read your message to make sure that your question is clear. In addition, I would appreciate questions that are expressed in an appropriately polite manner.
4. Because of the large number of messages that I receive, it may be several days before I am able to reply to your message. It is often easier and faster to ask a question in class, or during office hours. You can also call me on my office phone (but note that number cannot receive text messages). I may also provide answers in class, or post on Canvas to the whole class, instead of replying to individual messages.
5. If your question is about course content or general course administration, you should consider asking within your study group first, or posting on Canvas, rather than sending a message to me individually.

Information on copyright

In Canada, copyright law automatically protects written and creative works (e.g., text, art, music or performance), in all formats, as soon as they are created and until 50 years after the creator's death. The creator of the work (i.e., author, composer or artist) often owns copyright, though for published works the publisher may own copyright. A work does not need to be marked with © or a copyright statement to be protected. A copyright protected work cannot be copied (i.e., scanned, printed, downloaded, emailed or photocopied) without the copyright owner's permission, except in certain situations outlined in the *Copyright Act* (see SFU's Fair Dealing Policy for details).

Course materials such as PowerPoint slides, lecture notes, the lecture itself, assignments and exams are all protected by copyright. You are NOT allowed to reproduce them, mail them or upload them to any site for a purpose other than your own study.

SFU students, faculty and staff are required to abide by Canada's *Copyright Act* as well as SFU's copyright policies. Find information about what this means for students at SFU's copyright website, copyright.sfu.ca, or by contacting the Copyright Office at copy@sfu.ca. If you are copying works, using them in presentations and projects, or sharing them with other students, you should familiarize yourself with the guidelines and limits contained in SFU's Fair Dealing Policy.

Course Grade

The final grade will be calculated according to the percentages below. The final grade will take into account class attendance and participation (especially for students who are close to the next letter grade). More detail on the each of the grade components will be provided throughout the semester.

COMPONENT	WEIGHT
Assignments	40%
Participation and class presentation	20%
Project	20%
Term paper	20%

Percentage scores on assignments and other course components will be based on objective criteria. Final letter course grades will be computed from percentage scores on all the course components. The following table provides a rough estimate of grade breakdowns for the final grade. Due attention will be

given to the verbal descriptions listed below. There is no university-wide standard scale. An instructor adopts a grade scale appropriate to the level and content of the course.

A+	100%	to 96%	Extraordinary performance
A	< 96%	to 91%	Excellent performance
A-	< 91%	to 86%	
B+	< 86%	to 81%	Good performance
B	< 81%	to 76%	
B-	< 76%	to 71%	
C+	< 71%	to 66%	Satisfactory performance
C	< 66%	to 61%	
C-	< 61%	to 56%	Marginal performance
D	< 56%	to 50%	
F	< 50%	to 0%	Unsatisfactory performance (fail)

Grade Appeals

If a student wishes to contest the marking of an exam, assignment or paper, the instructor can agree to remark his/her entire exam at the instructor's convenience and not in front of the student. A grade reconsideration may raise the grade, lower the grade, or leave the grade unchanged, as stated in Policy T20.01, clause IV.2.

The only reason a grade change will be made is if there is an arithmetic error or if it has been determined that the exam, assignment or paper deserves a lower grade or a higher grade after it has been remarked.

The following are not reasons for reconsideration of a grade:

- The student is on probation.
- The student wants to get into Business or any other program.
- The student worked hard and thinks this should be a factor.
- The student does not like the grade scale.
- The student's score is x% below the next grade and would like the instructor to ignore the difference.

Topics and Readings

The following list is subject to change, depending on students' interests. These are suggestions, and students are encouraged to consult other readings, especially when preparing for assignments, presentations and papers. Note also that readings may be added.

Note: JM stands for "Jurafsky and Martin"; BKL for "Bird, Klein and Loper". Additional readings will be added in Canvas.

Week	Date	Topic	Read (before class)
1	Jan 5	Introduction	
2	Jan 10 Jan 12	Introduction, cont. Regular expressions	JM ch. 1, BKL ch. 1 & 2 JM ch. 2, BKL ch. 3

3	Jan 17 Jan 19	Morphology	JM ch. 3, BKL ch. 4
4	Jan 24 Jan 26	Language models	JM ch. 4
5	Jan 31 Feb 2	Part-of-speech tagging	JM ch. 5, BKL ch. 5
6	Feb 7 Feb 9	Parsing	JM ch. 12, 13, BKL ch. 8, 9
7	Feb 14 Feb 16	<i>Reading week. No classes</i> <i>Reading week. No classes</i>	
8	Feb 21 Feb 23	Semantics	JM ch. 17, 18, BKL ch. 10
9	Feb 28 Mar 2	Discourse	JM ch. 21
10	Mar 7 Mar 9	Data and learning from data Managing linguistic data	BKL ch. 6 BKL ch. 11
11	Mar 14 Mar 16	Information extraction Sentiment analysis	JM ch. 22, BKL ch. 7
12	Mar 21 Mar 23	Question answering and summarization Dialogue and conversational agents	JM ch. 23 JM ch. 24
13	Mar 28 Mar 30	Machine translation	JM ch. 25
14	Apr 4 Apr 6	Project presentations	