

Syllabus: Patterns in Language

New York University, Spring 2018

Up to date as of 2018-01-21

Course

LING-UA 6.001

Patterns in Language

Mondays and Wednesdays, 11am–12:15pm

Room 104, Department of Linguistics, 10 Washington Place

Course website

<https://www.nyu.edu/projects/champollion/patterns/>

Instructor

Professor Lucas Champollion

Homepage: <https://www.nyu.edu/projects/champollion/>

champollion@nyu.edu

Office hours: by appointment (room 404, 10 Washington Place)

Recitation

LING-UA 6.002

Thursdays, 3:30pm–4:45pm

Room 104, 10 Washington Place

Paul Feitzinger

pf42@nyu.edu

Office hours — Monday 9:30am-10:30am, Friday 12noon-1pm or by appointment (10 Washington Pl, room TBD)

Required textbook

Language and Computers

Markus Dickinson, Chris Brew, Detmar Meurers

ISBN: 978-1-4051-8305-5

Available at the NYU Bookstore, and online (about \$40)

Also available as an ebook, to buy or rent

Course description

If a computer is smart enough to beat humans at Jeopardy, does it follow that machines can think? Is it possible to predict the spread of the flu based on patterns in Google searches? Did Shakespeare really write that sonnet? Scientists use patterns in language to answer these questions, using the same concepts that underlie everyday applications like search engines, automatic translators, speech recognition, spell-checkers, and auto-correction tools. This course takes you on a tour of these applications, focusing on the technological and linguistic ideas behind them. You will gain practical hands-on experience and insight into how they work. No programming experience is required. The only background you need is curiosity about language and some everyday experience with computers.

Prerequisites

None; this course assumes no prior knowledge of linguistics or computer science, and only high-school level math background. Don't worry if you are not a native speaker of English.

Workload and grade calculation

- Homework assignments (30%)
- Midterm exam (30%)
- Final exam (30%)
- Recitation (10%)

Your letter grade for the class will be assigned on the following scale.

A	95–100%	B	83–86%	C	73–76%	D	63–66%
A–	90–94%	B–	80–82%	C–	70–72%	F	0–62%
B+	87–89%	C+	77–79%	D+	67–69%		

Fractional grades will be truncated down to the nearest integer (e.g. 94.7 is truncated to 94).

Data privacy disclaimer

We are using the commercial software CoCalc to run many of the homework assignments. The use of this software is for educational purposes only. This is a 3rd party software, which means that it is not an NYU-supported service. The data privacy, FERPA, and security protections in place are not at the same level as NYU-supported services (like your NYU Gmail, NYU Classes, etc.). We will structure assignments so that no highly sensitive information is needed to use the tool, but please note that we are subject to the terms of service set by the platform's developer (<https://cocalc.com/policies/terms.html>). If you have any concerns about the platform, please let us know as soon as possible.

Policies and other remarks

Attendance policy. We know that you are serious students but that you are also adults who budget time between coursework, jobs, personal/family commitments, and unforeseen circumstances. While we do not take attendance in class, as in any course, you will do best if you come to class every time and on time.

What to do if you miss class. Go to the NYU Classes site, <https://newclasses.nyu.edu>. All the materials will be posted to the course site. Review these and read the textbook chapter extra attentively. It's also a good idea to make a friend in the class so that you can trade notes in case one of you misses a lecture.

Extra credit. Midterm and final may or may not include some extra credit questions, worth up to an additional 5% of each test grade. Homework assignments may or may not occasionally have extra credit questions. There may or may not be an entire homework towards the end of the course that is more challenging than the others and that counts as extra credit in its entirety. Extra credit is an equal opportunity for everyone in the class to improve their grades, so it will not be issued to individual students upon request. If you want some grade security, we suggest doing all the homework assignments and not missing class.

How to contact us. The best way is by e-mail (see above for our e-mail addresses), or talk to us right before or right after class. Take advantage of office hours if you can. We will also be happy to meet you by appointment.

Personal situations. If anything arises during the semester that may affect your classroom performance, please come talk to the professor or your TA. If you wait until the end of the semester, we won't be able to help you. We're more likely to be able to address the situation if you speak to us when it happens. If something catastrophic happens in your life that prevents you from continuing your studies for an extended period of time, you should talk to the Dean of Students (currently Dean Kalb, richard.kalb@nyu.edu).

Accommodations. If you have a diagnosed disability that requires special accommodations, please bring your paperwork to class and talk to the professor and to your TA. For further information, see the Moses Center website at <https://www.nyu.edu/life/safety-health-wellness/students-with-disabilities.html>

Religious holidays. Whenever feasible, we will avoid scheduling exams and deadlines on religious holidays (please let us know if we have overlooked one). If you anticipate being absent because of any religious observance, you should tell us in advance whenever possible. If you are unable to attend class because of your documented religious beliefs, you will not be penalized for any class, recitation, exam, or assignment deadline missed on that day or days, and you will be given a make-up exam or an assignment deadline extension.

Homework Assignments

Assignments are due at the beginning of class, before the start of lecture. Assignments will be graded check-plus (full credit), check (90% credit towards the homework grade), or check-minus (80% credit). Late assignments may be turned in no later than the next class meeting, but cannot receive a grade higher than a check. If you have documentation of a medical situation or other emergency that prevents you from turning in an assignment on time, please email the TA and the instructor as soon as possible.

Collaboration. We encourage you to work on homework with other students, but you must work on each part of the homework assignment yourself, and you must write up and turn in your own answers. If you work with others, please add to your homework the names of the people that you worked with. This will not affect your grade or their grade.

Academic Honesty Policy. Cheating will not be tolerated and may cost you your grade and have repercussions in your career (for example, having a cheating record will all but negate your chances of getting into a decent law school). **You must solve all homework assignments alone and without help from other students.** If you use any resource other than the textbook and lecture slides as you prepare your homework assignment, you must provide a scholarly citation for the resource and explain the extent to which you have used it. The following is a non-exhaustive list of examples of what counts as cheating in this course: (i) copying a homework or exam from another student, with or without the student's knowledge; (ii) plagiarism, that is: reproducing data or ideas from a scholarly

source, website, or other resource (other than the textbook and lecture slides) without attributing authorship where it is due; (iii) including solutions obtained from answer keys or similar sources into one's homework submission. Please read the University policies on Academic Integrity, which are described on <http://cas.nyu.edu/academic-integrity.html> for CAS and on <https://tinyurl.com/nyu-policy> for NYU (also via <https://www.nyu.edu> > About NYU > Policies and Guidelines > University Policies and Guidelines > Academic Affairs & Faculty > Academic Integrity for Students at NYU).

You will fail any homework on which you have cheated, even if you have only cheated on a part of it. You will fail the course if you cheat repeatedly or if you cheat on a test. If we have reasonable grounds to suspect a case of plagiarism or cheating, we reserve the right to refer the case directly (i.e., without first talking it over with the students) to the Dean of Students. This is because the dean is in a better position than ourselves to distinguish first-time from repeat offenders and to apply fair and consistent standards across NYU.

How to succeed in this course. This course requires a certain degree of involvement on your part if you wish to earn more than a passing grade. If you want to do well, you should:

- come to all the classes and take notes (you can download or print out the slides before class);
- check your email often for announcements;
- take full advantage of recitations and office hours — the TA is there to answer your questions;
- do all the reading before the day for which it is assigned;
- set aside at least two or three hours each week to do the homework, and don't wait to do it at the last minute;
- not be afraid to ask your TA/professor for help if you have trouble with an assignment.

Schedule

Wk	Date	Topics	Notes
1	Jan 22 Jan 24	Organization, Introduction Jupyter notebook crash course in Python	
2	Jan 29 Jan 31	Ch 1 (w/o speech recognition): Encodings Probability and language models	Feb 4: add/drop period ends
3	Feb 5 Feb 7	Probability and language models Ch 2: Writers' Aids	
4	Feb 12 Feb 14	Ch 3: Tutoring Systems Ch 3: Tutoring Systems	
5	Feb 19 Feb 21	(no class: President's Day) Ch 4: Search	
6	Feb 26 Feb 28	Ch 4: Search (buffer)	
7	Mar 5 Mar 7	midterm review midterm	
8	Mar 12 Mar 14	(no class: Spring break) (no class: Spring break)	
9	Mar 19 Mar 21	Ch 5: Classifiers and machine learning Ch 5: Classifiers and machine learning	
10	Mar 26 Mar 28	Ch 6: Dialog Systems and AI Ch 6: Dialog Systems and AI	
11	Apr 2 Apr 4	Ch 7: Machine Translation Ch 7: Machine Translation	
12	Apr 9 Apr 11	Forensic Linguistics (buffer)	
13	Apr 16 Apr 18	Speech synthesis Speech recognition	
14	Apr 23 Apr 25	Ch 8: Impact of LT / Google Flu Trends (buffer)	
15	Apr 30 May 2	(buffer) (buffer)	
16	May 7	Final review	last class
	May 14	final exam, 10am-11:50am	