

Computing and Linguistics

Director of undergraduate studies: Simon Charlow (simon.charlow@yale.edu) (Linguistics); Computing and Linguistics website

The Computing and Linguistics major provides multidisciplinary training in the computational study of human language, the development of systems for natural language processing, and the automated analysis of textual data in applications in the humanities, social sciences, and sciences. Students learn the foundational tools and methods that underlie this work, including areas of computer science, statistics and data science, and linguistics, and apply them to some empirical domain, through coursework and an independent research project in the senior year.

The B.A. in Computing and Linguistics exposes students to the fundamental ideas and foundational techniques of the field, while the B.S. provides more extensive training and engagement in research, preparing students for graduate work in the area.

PREREQUISITES

There are three prerequisites for this major and they fall in three areas of study: (1) *statistics*, satisfied through S&DS 1000, or 1230, or 2200, or comparable background in statistics (e.g., through a score of 5 on the AP Statistics exam) as approved by the director of undergraduate studies (DUS); (2) *programming*, satisfied through an appropriate introductory programming course or comparable experience as approved by the DUS; and (3) *linguistics*, satisfied through one 1000-level Linguistics course. It is also advisable that students have some background in single-variable calculus, before beginning this major.

REQUIREMENTS OF THE MAJOR

See Link to the *YC CPLI Elective in CP & Ling* attribute indicating courses approved for the major requirements.

B.A. degree program The B.A. degree program requires 11 term credits beyond the prerequisites and not including the senior requirement. Core courses, as listed below, are required from the following categories: 2 math core courses; 1 statistics core course (S&DS 2380); 2 linguistics core courses; 2 computation core courses; 3 advanced courses; 1 elective, and 1 senior requirement course.

B.S. degree program The B.S. degree program requires 14 term credits beyond the prerequisites and not including the senior requirement. Core courses, as listed below, are required from the following categories: 2 math core courses; 2 statistics core courses; 3 linguistics core courses; 2 computation core courses; 3 advanced courses; 2 electives, and 2 senior requirement courses.

Math core courses Both B.A. and B.S. degree students must take one course in proof-based discrete mathematics (one of MATH 2440, LING 2249, or CPSC 2020) and one course in linear algebra (either MATH 2220 or MATH 2250).

Statistics core courses These provide foundations in probability and statistical theory. B.A. degree students satisfy this requirement by taking S&DS 2380; B.S. degree students choose between two options (1) one of S&DS 2400 or S&DS 2410, together

with S&DS 2420; (2) S&DS 2380 and either S&DS 2300 or any S&DS course numbered 2420 or above.

Linguistics core courses These courses, LING 2320, 2530, and 2630, expose students to the nature of linguistic structure and its variability across languages, at the level of sound (phonology), form (syntax) and meaning (semantics). B.A. degree students must take 2 out of these 3 courses, while B.S. degree students must take all 3.

Computation core courses Computational studies of language rest crucially on the foundations of computer science and programming. To this end, both B.A. and B.S. degree students must take CPSC 2010 and 2230.

Advanced courses Both B.A. and B.S. degree students must take 1 advanced course in linguistic structure, either LING 3350, 3540, or 3640; 1 course in natural language processing, either CPSC 4770 or LING 2270; and 1 course in machine learning, either S&DS 2650, 3650, or CPSC 3810.

Electives Elective courses may be used to explore the application of the techniques of computational linguistics across a range of disciplines or to deepen expertise in these techniques. Courses that are pre-approved to satisfy the elective requirement are listed on the Computing and Linguistics major website, but other relevant courses may satisfy this requirement with DUS approval. B.A. degree students take 1 elective course; B.S. degree students take 2 electives.

Credit/D/Fail No course taken Credit/D/Fail may be applied toward the requirements of the major (other than as prerequisites).

Outside credit Courses taken at another institution or during an approved summer or term-time study abroad program may count toward the major requirements with DUS approval.

SENIOR REQUIREMENT

All Computing and Linguistics majors enroll in the capstone seminar CSLI 4900 in the fall of the senior year. This seminar includes discussion of student research, as well as presentations by researchers in the field from both inside and outside of Yale. B.A. degree students complete a senior project as part of CSLI 4900, working either on an independent project supervised by a Yale faculty member with relevant expertise or as part of a group effort of capstone seminar participants. B.S. degree students enroll in the capstone seminar in the fall and continue work on their senior project in the spring. The senior project of B.S. degree students must involve independent research.

ADVISING

Students interested in the Computing and Linguistics major are encouraged to consult with the DUS. Further information about the major and answers to FAQs are available on the Computing and Linguistics website. The entire selection of courses by students in the major must be approved by the DUS.

SUMMARY OF MAJOR REQUIREMENTS

Prerequisites *Statistics*: one of S&DS 1000, 1230, or 2200 or comparable experience; *Programming*: introductory programming course or comparable experience; *Linguistics*: one 1000-level LING course

Number of courses *B.A. degree*—11 term credits beyond prereqs and not incl senior req; *B.S. degree*—14 term credits beyond prereqs and not including senior req

Specific courses required *For both degrees*—2 computational core courses: CPSC 2010 and CPSC 2230; for *B.A. degree*—S&DS 2380

Distribution of courses *Both degrees*—2 math core courses, 1 adv linguistics structure course, 1 adv natural language processing course, 1 adv course in machine learning; *B.A. degree*—2 linguistics core courses, 1 elective; *B.S. degree*—2 statistics core courses, 3 linguistics core courses, 2 electives

Substitution permitted Elective courses in computational linguistics, machine-learning and applications of computational linguistics, as approved by DUS

Senior requirement *Both degrees*—Capstone seminar CSLI 4900; *B.S. degree*—one additional semester of senior project