Linguistics H384
Language and Computers

Course goals In the past decade, the widening use of computers has had a profound influence on the way ordinary people communicate, search and store information. For the overwhelming majority of people and situations, the natural vehicle for such information is natural language. Text and to a lesser extent speech are crucial encoding formats for the information revolution.

In this course, you will be given insight into the fundamentals of how computers are used to represent, process and organize textual and spoken information, as well as tips on how to effectively integrate this knowledge into working practice. We will cover the theory and practice of human language technology. Topics include text encoding, search technology,tools for writing support, machine translation, dialog systems, computer aided language learning and the social context of language technology.

Instructor: Chris Brew

Office: 200 Oxley Hall

Phone: 292-5420

E-mail: cbrew@ling.ohio-state.edu

Office hours: (at least for the first week)

  Tu 4:30-5:00
  Th 4:30-5:00pm
      or by appointment

Meeting time: MW 1:30-3:18pm

Classroom: 29 Derby Hall (DB)

Course website: http://www.ling.ohio-state.edu/~cbrew/H384/

Credits: 5

Course prerequisites: none

Readings: There is no textbook for this course, but there will be readings (mostly online) assigned periodically throughout the course.

Course requirements: There will be various reading selections throughout the quarter, but most of the material will be introduced solely in the classroom. There will be roughly one online **quiz** per week, to ensure the material covered in class is mastered. And there will be approximately one **homework** (exercise sheet) per week. These assignments give the opportunity to explore new aspects of the topics discussed in class.

Grading   Grades will be based on classroom discussion (10%), quizzes (20%), homeworks (30%), a midterm exam (20%), and a final examination (20%).

PARTICIPATION AND READINGS	20%	
QUIZZES	20% (8@2.5% each)	
HOMEWORKS	20% (8@2% each)	
MITERM	20%	WEDNESDAY, MAY 5th, 1:30pm
FINAL	20%	WEDNESDAY, JUNE 9, 1:30pm

- Homework assignments are due by the beginning of each class (3:30pm)—you may hand them in or e-mail them to me. You can work together on the homework assignments but write out your own answers. Your homework grade will be based on both quality and effort. LATE HOMEWORK POLICY: You will lose one (1) letter grade (10 percentage points) for every 24 hour period after the turn-in time.

- Quizzes are due by midnight of the day indicated. They will be shut off once the deadline is reached. You will have ample time to complete the quiz, so DO NOT PUT IT OFF TO THE LAST MINUTE!

- Because homeworks, quizzes, and exams reflect material covered in class and because there will be much in-class discussion, attendance is essential for doing well in the class.

- If you feel that I have graded anything incorrectly or improperly, please contact me outside of class. I will be happy to address your concerns.

Grading scale   (Scores in percentages)

A	93-100	B+	87-89	C+	77-79	D+	67-69	E	0-59
A-	90-92	B	83-86	C	73-76	D	63-66		
		B-	80-82	C-	70-72	D-	60-62		

Make-up Policy: As you generally will have a week to take them, there are no make-ups for the quizzes.

If you plan on missing either the midterm or final, you will have to provide extensive documentation for your excuse. See me immediately if this is the case.

Academic Misconduct: To state the obvious, academic dishonesty is not allowed. Cheating on tests or on other assignments will be reported to the University Committee on Academic Misconduct. The most common form of misconduct is plagiarism. Remember that any time you use the ideas or the materials of another person or persons, you must acknowledge that you have done so in a citation. This includes material that you have found on the Web. The University provides guidelines for research on the Web at http://gateway.lib.ohio-state/tutor/.

Class etiquette: I expect you to respect one another, to respect me, and to respect yourself. To that end, I expect you to obey the following rules:

- Do not read newspapers in class. Do not criticize others' questions. Do not pack up early. Shut off your cell phone. If for some reason, you must leave early or you have an important call coming in, notify me before class.

- Participate: share experiences, ask questions, express your opinions. Ask me to provide more information and see me during office hours for help or clarification or recommendations for further research.

- If you ever have a problem with the course or with me, I prefer you see me during office hours (or schedule an appointment). If, for some reason, you feel you cannot discuss an issue with me, please take your concern to the TA Coordinator (see below).

Students with Disabilities: Students who need an accommodation based on the impact of a disability should contact me to arrange an appointment as soon as possible to discuss the course format, to anticipate needs, and to explore potential accommodations. The instructor relies on the Office of Disability Services for assistance in verifying the need for accommodations and developing accommodation strategies. Students who have not previously contacted the Office for Disability Services are encouraged to do so (614-292-3307; www.ods.ohio-state.edu).

Schedule:

Month	Week	Day	Date	Topic	Assignments
Mar	1	M	29	Intro to class	
Mar		W	31	Computing Machinery and Intelligence	
Apr	2	M	5	Text Encoding	
		W	7	Speech Encoding	HW1/Quiz1 due (M)
	3	M	12	Internet Searching	
		W	14	Spam filtering	HW2/Quiz2 due (W)
	4	M	19	Spelling & Grammar correctors	
		W	21	Spelling & Grammar correctors	HW3/Quiz3 due (W)
	5	M	26	Machine Translation	
		W	28	Machine Translation	HW4/Quiz4 due (W)
May	6	M	3	Machine Translation	
		W	5	MIDTERM	
	7	M	10	Dialogue systems	
		W	12	Dialogue systems	HW5/Quiz5 due (W)
	8	M	17	Dialogue systems	
		W	19	Dialogue systems	HW6/Quiz6 due (W)
	9	M	24	Computer-aided language learning	
		W	26	Computer-aided language learning	HW7/Quiz7 due (W)
	10	M	31	Memorial Day, no class	
June		W	2	Social context of language technology use	HW8/Quiz8 due (W)
	11	W	9	FINAL EXAM	3:30pm

Computational Linguistics research If you find yourself loving this material, I encourage you to come see me. We have a computational linguistics discussion group (mostly graduates, but all welcome) and several opportunities for research projects.

Tentative schedule

- Week 1: Introduction: Computational intelligence.

  Introduction to the general area. What makes us decide that a language user is acting like a human?

  Readings for Wednesday (short response piece due in class) The Turing Test". Hofstadter's Coffeehouse conversation

- Week 2: Language and physics: Speech encoding.

  A look at digital representations of speech, including spectrograms, so-called "pictures" of speech. Lab session: making a speaking clock

- Week 3: Language and the internet: Searching

  What facilities exist for searching for information on the internet. What you need to know about the technology in order to use it well. Differences between specific and general queries. How to evaluate query results. Lab session: query-by-example

- Week 4: Language and the internet: Language identification and spam filtering

  Techniques for classifying documents, including what language(s) they are written in. Are they junk mail? Are statistical techniques better than rule-based ones, or not? When will the techniques fail? Lab session: spoofing spam filters

- Week 5: Spelling and grammar correction

  What do so-called "grammar checkers" and "spelling correctors" do? What do such programs base their advice on? When does it make sense to use such tools and what kind of errors are to be expected? Lab session: grammar checkers

- Week 6: Machine translation

  What do the free internet-based translation services manage to do, and where do they fail? For what purposes can automatic machine translation work reliably? What translation support functions can a computer provide? Lab session: internet translation systems

- Week 7: Machine understanding

  A closer look at what a machine would have to know in order to "understand": is it the grammar, the meaning, the culture, all three, or something else? Lab session: working with computational grammars

- Week 8: Dialog systems I

  Eliza and its surprising success in engaging people in conversation. When are dialog systems used and for what purpose? A closer look at the components of a dialog system. What kind of knowledge needed to make it work? And where do we put that knowledge? Lab session: build your own Eliza

- Week 9: Computer-Aided Language Learning

  What is involved in learning a foreign language? What role in language learning can computers play? A look at everything from vocabulary training, via presentation of learning material to providing feedback on learner errors and progress. Lab session: working with CALL

- Week 10: Social context of language technology use

  How do we react to computers that make use of language? What does it mean for the way we see ourselves? What assumptions do we make about every user of language, be it a human or a machine? Are there ethical issues to consider in using computers to emulate language?

Disclaimer This syllabus is subject to change. All important changes will be made in writing, with ample time for adjustment. (Midterm and final dates, however, will not change.)