Linguistics 384: Language and Computers (Winter 05)

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: Detmar Meurers

- *Office:* 201a Oxley Hall (enter through 201 computer lab; if locked, knock loudly)
- *Phone:* 292-0461 (usually email works better)
- *Email:* dm@ling.osu.edu
- *Office hours:* Tuesday 11:00-12:00, and by appointment

Course meets: Tuesday and Thursday 12:30-2:18pm in 245 Central Classrooms

Course website: `http://ling.osu.edu/~dm/05/winter/384/`

The updated syllabus, assignments, slides, etc. will be posted there; check it regularly!

Course email: `meurers384@ling.osu.edu`

Mail sent to this address is forwarded to the official email addresses (`Name.Number@osu.edu`) of all students enrolled in the class and the instructor. Note that you should read email sent to your official osu account on a daily basis—it'll also help you avoid high library fines!

Anonymous feedback: If you have comments, complaints, or ideas you'd like to send me anonymously, you can use the web form at `http://ling.osu.edu/~dm/feedback/` to do so.

Please send me ordinary email for anything that you'd like to receive a reply to—there really is no way for me to find out who sent me something via the anonymous feedback form!

Readings: There is no textbook for this course (the topic is quite new, at OSU and elsewhere). There will be some readings assigned periodically throughout the course.

I will distribute slides in class for each unit. These will also be available on the web after the class in which they are first distributed. **These slides are only a skeleton of the material covered; they cannot replace actually being in class.** In my experience, students who actively participate in class enjoy the course more and get much better grades than those who don't—very surprising, isn't it? ;-)

Course requirements: The basic requirement is **regular attendance in class** and **active participation**. There will be roughly one online **quiz** per topic, to ensure the material covered in class is mastered. And there will be one **homework** (exercise sheets) per topic these assignments give the opportunity to explore new aspects of the topics discussed in class. The **midterm** will consist of the material covered in the first half of the class, and the **final** will cover the contents covered in the second half of the class.

Grading: Grades will be based on classroom discussion, quizzes, homeworks, a midterm exam, and the final examination, using the following scheme:

Participation	10%	
Quizzes	20%	
Homeworks	30%	
Midterm	20%	(on Thursday, Feb 10, 12:30–2:18pm)
Final	20%	(on Tuesday, Mar 15, **11:30 (!)**–1:18pm)

- Given that homeworks, quizzes, and the exams reflect the material covered in class, attendance is essential for doing well in this class.

- Homework assignments are due by the beginning of class, on a piece of paper. **No late homeworks are accepted.**

  I encourage group work on the homework assignments, but each of you should write out your own answers. Note that group work means that everyone in the group contributes and fully understands what you turn in—where in doubt, I may ask you individually to explain to me in person what you turned in.

- Quizzes are administered through WebCT (`http://class.osu.edu`) and are due by midnight of the day indicated. They will be shut off automatically once the deadline is reached. You will generally have a week to complete a quiz, so **do not put it off to the last minute!** The quizzes naturally are open book, so you should view them as an opportunity for reviewing the material covered, based on the handouts.

  Note: I do not promise to remind you when you have a quiz due. It is your responsibility to keep up with the syllabus.

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

- Grading scale (scores in percentages):

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

Make-up Policy: If you know you won't be able to make a deadline or exam, please see me **before** you miss the deadline or exam. If you miss the midterm or final, you will have to provide extensive written documentation for your excuse.

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

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, 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.edu/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:

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

- Do not read newspapers, materials from other classes, etc. in class. Do not pack up early. Switch off your cell phone. If for some reason, you must leave early or you have an important call coming in, notify me before class.

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. I rely 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 (292-3307; `http://www.ods.ohio-state.edu`).

Topics:

1. Storing language on the computer: Text and speech encoding.

   Writing systems used for language. Representing text on the computer. Digital representations of speech.

2. Searching: web, library catalogs, and other language-based databases

   What facilities exist for searching for language-based information? Different query languages and what they allow you to do. Differences between specific and general queries. How to evaluate the results of a search.

3. Classifying documents: Language identification and spam filtering

   Techniques for classifying documents. What language(s) are they written in? Are they junk mail? Are statistical techniques better than rule-based ones, or not? When will the techniques fail?

4. Writer's aids: 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?

5. 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? A closer look at what makes machine translation such a hard task. Is it the grammar, the meaning, the culture, all three, or something else?

6. Dialog systems

   Eliza and its surprising success in engaging people in conversation. When are dialog systems used, for what purpose? A closer look at the components of a dialog system. Where is what kind of knowledge needed to make it work?

7. Computer-Aided Language Learning

   What is involved in learning a foreign language? What role in language learning can computers play: from vocabulary training, via presentation of learning material, to providing feedback on learner errors and progress.

8. 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.