Syllabus: Linguistics 478/578 (also Speech and Hearing Science 478/578)

Speech Technology, Fall 2006

Time: Tuesday/Thursday 11:00-12:15  
Place: Social Sciences 224 (ICL)  
Professor: Natasha Warner  
Office hours: T/Th 9:30-10:30, Douglass 320  
Phone: 626-5591  
Email: nwarner@u.arizona.edu (do not use d2l email to reach me, I don't read it)  
Course webpage: d2l.arizona.edu

Course description:  
The main focus of this course will be on speech synthesis (making computers talk) and speech recognition (making computers take dictation), with some time allotted to other speech and language technologies. Speech technology is an active industry, and there is great potential both for people with knowledge of speech and people with knowledge of computer science (or ideally both) to work in that industry. The purpose of this course is to give you background that would be useful if you pursue work in the speech technology industry.

Prerequisites
Either a background in phonetics (such as Linguistics 314 or 515 or Speech and Hearing 267) or strong programming skills are required for this course. (A background in both is explicitly not expected or required.) The course will be accessible for students with knowledge of speech but no programming background. It will also be useful for students with strong programming background but no knowledge of speech. There will be some readings on basic acoustic phonetics, and students with no previous experience with phonetics should read these soon.

Readings


Ladefoged, P. 2001. A Course in Phonetics. 4th ed. Heinle and Heinle. Ch. 8 only. (on d2l site)


Course requirements: 478
5 homework assignments* 50% (10% each)  
Paper topic proposal 5%  
Partial annotated bibliography 5%  
Paper progress report 5%  
Presentation on final paper 5%  
Final paper 25%  
Class participation 5%

* Some of the homework assignments will be carried out as small group projects.

Course requirements: 578
Same as above, except that all things worth 5% above are worth 4% each for 578, and a presentation on a research article (not one for your term paper) is also required, and is worth 5%.

The homework assignments, intermediate reports on the paper, presentation, and final paper are required of
all students. For students in 578, the paper must involve collection of original data related to some topic of the course, and the paper is likely to be about 15-20 pages long. For students in 478, the paper may either be a literature review on some topic in speech technology, or may incorporate collection of original data, and it is likely to be about 10-15 pages long. Some homework assignments may also include additional questions which are required only of the students in 578, but which can be done by students in 478 for extra credit.

All students must read all of the required readings and be prepared to discuss them in class. Questions on the readings will be included in the homework assignments.

All students should attend class every day except in cases of dire emergency or serious illness. Attendance will not be taken, but you cannot get a good grade for participation without being here to participate. If attendance becomes a problem, I reserve the right to give short pop quizzes and add these to the grading system, adjusting the percentages above as necessary. If I feel a need to do this, the change will be announced in advance.

All assignments must be turned in by 2 PM on the day due, except the final paper, which must be turned in by 4 PM on the day due. Late assignments will be docked 10% of the possible grade per day late, unless you have a very good documented reason for the lateness. Please do not submit homework by email or fax without asking ahead of time. If you do not turn in your homework during class, you should turn it in to my mailbox.

Approximate course schedule (subject to change)

<table>
<thead>
<tr>
<th>Dates</th>
<th>Topic</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/22-24</td>
<td>Introduction, IPA transcription</td>
<td></td>
</tr>
<tr>
<td>8/29</td>
<td>Acoustic phonetics</td>
<td>read Rodman Ch. 1, Holmes and Holmes Ch. 1, Ladefoged Ch. 8</td>
</tr>
<tr>
<td>8/31-9/5</td>
<td>Digital signal processing, articulatory synthesis</td>
<td>read Rodman Ch. 2</td>
</tr>
<tr>
<td>9/7-12</td>
<td>LPC synthesis, parametric synthesis</td>
<td>read Rodman Ch. 4, Holmes and Holmes Ch. 6, HW 1 due 9/12</td>
</tr>
<tr>
<td>9/14-19</td>
<td>Concatenative synthesis, waveform coding</td>
<td>read Holmes and Holmes Ch. 5</td>
</tr>
<tr>
<td>9/21-26</td>
<td>Prosody, PSOLA, Text-to-phonemes</td>
<td>read Holmes and Holmes ch. 7</td>
</tr>
<tr>
<td>9/28-10/3</td>
<td>Text-to-phonemes, other topics in synthesis</td>
<td>HW 2 due 10/3, Paper topic proposal due 9/28</td>
</tr>
<tr>
<td>10/5-10</td>
<td>Applications of synthesis, evaluation of synthesis</td>
<td>read Rodman Ch. 7, one half of 578 article presentations</td>
</tr>
<tr>
<td>10/12-17</td>
<td>Main issues of ASR, rule-based recognition, template matching, front-end analysis</td>
<td>read Rodman Ch. 3, Holmes and Holmes Ch. 8, 10, HW 3 due 10/12</td>
</tr>
<tr>
<td>10/19-24</td>
<td>HMMs, training HMMs</td>
<td>read Holmes &amp; Holmes Ch. 9, Partial annotated bibliography due 10/19</td>
</tr>
<tr>
<td>10/26-10/31</td>
<td>Continuous speech, language modeling</td>
<td>read Holmes and Holmes Ch. 12, HW 4 due 10/26</td>
</tr>
<tr>
<td>11/2-7</td>
<td>Artificial neural networks, evaluating ASR</td>
<td>read Holmes and Holmes ch. 13</td>
</tr>
</tbody>
</table>
We will have two guest speakers from speech technology industries during the semester. The dates of their visits will be announced later, and the schedule adjusted accordingly.

**Note**
Appropriate academic behavior is expected, e.g. cheating and plagiarism are unacceptable, disruptive behavior in class is unacceptable, and the student code of conduct ([http://info-center.ccit.arizona.edu/~studpubs/policies/studcofc.htm](http://info-center.ccit.arizona.edu/~studpubs/policies/studcofc.htm)) should be followed. It is also expected that students will treat others in the classroom with respect.