# CS 441 – Automatic Speech Recognition
# Course Syllabus and Calendar – Spring 2003

*Professor Don Colton*

Brigham Young University Hawaii

## This Document Is Preliminary

## 1  Overview

Automatic Speech Recognition (ASR) is a computerized process. It receives as its input a speech recording. It produces as its output a transcription. In the computing world, ASR is an unsolved problem, which means that nobody has yet demonstrated the ability to do it as well as humans can.

This course is designed for two types of student. First, it is designed for Computer Science students who have already achieved some skill in programming, generally CS 301 (Algorithms), and who want to explore the rich task domain of speech input problems, including spoken language translation. Second, it is designed for Linguistics students who have already achieved some skill using phonetic transcription, and who want to explore the current realities and future promise of linguistic computerization as a labor-saving device. The goal of the course is to take these two types of students and any others interested in ASR and to move them forward in understanding and ability, bringing them closer to their own personal goals.

This course is experimental. It exists because the instructor, Brother Colton, did his PhD work in Computer Science and Engineering, in the area of spoken language systems, and this course gives him the chance to share his interests and excitement with students. The content will change somewhat from year to year as the instructor discovers things that work well, tries out things that seem promising, and discards things that did not work as well as planned. This means two things to you the student: the course will probably never be the same twice in a row, and if you have particular interests you would like to see addressed, speak up. It is possible that the calendar can be rearranged to fit in your area of interest. Some topics are best addressed in response to a question. Some topics that are incredibly interesting will never come up unless someone asks. There is simply too much material available in the field of spoken language systems to attend to all of it this semester. So please feel free to suggest topics that are of interest to you.

Lab assignments will be a part of this class. Generally each lab will have a research paper option and a programming option. For this two-credit class, I am targeting two major lab assignments during the semester.

### 1.1  The Course

- **Course Number:** CS 441
- **Title:** Automatic Speech Recognition
- **Course Description:** Introduction to automatic speech recognition by computers, including digital sampling, Fourier transformation, phonemic classification by neural network, and Viterbi search. (Prerequisite: CS 440.)
- **Textbook:** none at this time.
- Class Time: TTh 11:00–12:50 PM
  Final Exam: Thu 18 Jun, 11:00–12:50 PM
  Classroom: GCB 153

### 1.2  The Instructor

- **Instructor (me):** Don Colton
- **My email:** don@colton.byuh.edu
- **My Office:** GCB 130 B
- **Office Hours:** MWF 11 AM to 1 PM
- **Office Hours:** TTh 1 PM to 3 PM

## 2  Subject to Change

It is possible that I will revise some aspects of the course as we go along. Any changes I make are likely to be to your advantage. If any of my changes seems unfair to you, let me know. I will try to correct it.

## 3  Grading Overview

Your final grade will be based on 1000 points assigned in various categories. Some extra credit points may be available. The total-points grading will be as follows:

| 930+ | A | 900–929 | A- | 870–899 | B+ |
|------|---|---------|----|---------|----|
| 830–869 | B | 800–829 | B- | 770–799 | C+ |
| 730–769 | C | 700–729 | C- | 670–699 | D+ |
| 630–669 | D | 600–629 | D- | 0–599 | F |

# 4   Course Calendar

Here is the day-by-day plan.

```
Day  1a Overview, Resonance
Day  1b Phonetics, Mondegreens
Day  2 Files and Quantization
Day  3 Learning the IPA
Day  4 Using the IPA
Day  5 More IPA
Day  6 Spectrograms
Day  7 Spectrograms
Day  8
Day  9
Day 10
Day 11
Day 12
Day 13
Day 14
Day 15 Final Exam (2 hr)
```

# 5   Topics

I intend to address the following topics, depending on student interest and the availability of appropriate tools.

## 5.1   Wave Files

Students will learn manner in which sounds are recorded digitally for use on computers. They will explain quantization. They will become familiar with the Cambridge "sox" program for sound file conversion. Students will explain why so many sound formats exist and how they are used. Students will identify the most popular formats for recording sound on computers, and will explain their advantages and disadvantages.

## 5.2   Phonetics

Students will become familiar with the international phonetic alphabet (IPA). They will use it to transcribe recorded spoken utterances into IPA. They will translate written English into IPA. They will translate IPA into written English. Students will become familiar with the Moby dictionary.

(tipa font for latex)

## 5.3   Spectrograms

Students will learn how to read a spectrogram and will demonstrate their ability by reading spectrograms in homework and on tests.

## 5.4   Neural Networks

Students will be introduced to neural networks as a way of estimating the likelihood that a given frame of speech comes from an attempt to articulate a specified phoneme. Students will explain how a neural network is trained and will discuss issues affecting recognition accuracy.

## 5.5   Project

Students will complete a project. CS students will write programs to translate text into IPA. Advanced students may write programs that translate IPA into text.

## 5.6   Ideas

Here are some ideas I am not sure how to use just yet.
   * (s|sh)(cons)(r|l|w)(vowel)(cons) = syllable

# 6   Special Needs

Brigham Young University Hawaii is committed to providing a working and learning atmosphere, which reasonably accommodates qualified persons with disabilities. If you have any disability that may impair your ability to complete this course successfully, please contact the students with Special Need Coordinator, Leilani A'una at 293-3518. Reasonable academic accommodations are reviewed for all students who have qualified documented disabilities. If you need assistance or if you feel you have been unlawfully discriminated against on the basis of disability, you may seek resolution through established grievance policy and procedures. You should contact the Human Resource Services at 780-8875.

# 7   Preventing Sexual Harassment

Title IX of the education amendments of 1972 prohibits sex discrimination against any participant in an educational program or activity that receives federal funds, including Federal loans and grants. Title IX also covers student-to-student sexual harassment. If you encounter unlawful sexual harassment or gender-based discrimination, please contact the Human Resource Services at 780-8875 (24 hours).