

## LECTURE 01: COURSE OVERVIEW AND OBJECTIVES

[Return to Main](#)

### Introduction:

[Overview](#)

[Syllabus](#)

[Technology](#)

### Software Resources:

[Internet Speech](#)

[Software](#)

[Toolkits](#)

### Educational Resources:

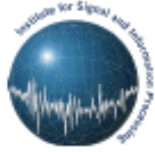
[SRSDR'02](#)

[Training](#)

[Short Course](#)

[Misc. Notes](#)

- Objectives:
  - Learn about basic technology
  - Understand theory at a fundamental level
  - Relate to other theory such pattern recognition, signal processing, computational linguistics, etc.
  - Develop perspective: Are the approaches we use specific to a speech signal?
  
- What we won't do:
  - Computer programming
  - Computer simulations
  - Matlab exercises
  - Teach you how to tune parameters
  - Train you to be speech technologists...
  
- Why?



**Introduction:**

- 01: Organization  
([html](#), [pdf](#))

**Speech Signals:**

- 02: Production  
([html](#), [pdf](#))
- 03: Digital Models  
([html](#), [pdf](#))
- 04: Perception  
([html](#), [pdf](#))
- 05: Masking  
([html](#), [pdf](#))
- 06: Phonetics and Phonology  
([html](#), [pdf](#))
- 07: Syntax and Semantics  
([html](#), [pdf](#))

**Signal Processing:**

- 08: Sampling  
([html](#), [pdf](#))
- 09: Resampling  
([html](#), [pdf](#))
- 10: Acoustic Transducers  
([html](#), [pdf](#))
- 11: Temporal Analysis  
([html](#), [pdf](#))
- 12: Frequency Domain Analysis  
([html](#), [pdf](#))
- 13: Cepstral Analysis  
([html](#), [pdf](#))
- 14: **Exam No. 1**  
([html](#), [pdf](#))
- 15: Linear Prediction  
([html](#), [pdf](#))
- 16: LP-Based Representations  
([html](#), [pdf](#))

**Parameterization:**

- 17: Differentiation  
([html](#), [pdf](#))
- 18: Principal Components  
([html](#), [pdf](#))

# ECE 8463: FUNDAMENTALS OF SPEECH RECOGNITION

Professor Joseph Picone  
Department of Electrical and Computer Engineering  
Mississippi State University

email: [picone@isip.msstate.edu](mailto:picone@isip.msstate.edu)  
phone/fax: 601-325-3149; office: 413 Simrall  
URL: [http://www.isip.msstate.edu/resources/courses/ece\\_8463](http://www.isip.msstate.edu/resources/courses/ece_8463)

Modern speech understanding systems merge interdisciplinary technologies from Signal Processing, Pattern Recognition, Natural Language, and Linguistics into a unified statistical framework. These systems, which have applications in a wide range of signal processing problems, represent a revolution in Digital Signal Processing (DSP). Once a field dominated by vector-oriented processors and linear algebra-based mathematics, the current generation of DSP-based systems rely on sophisticated statistical models implemented using a complex software paradigm. Such systems are now capable of understanding continuous speech input for vocabularies of hundreds of thousands of words in operational environments.

In this course, we will explore the core components of modern statistically-based speech recognition systems. We will view speech recognition problem in terms of three tasks: signal modeling, network searching, and language understanding. We will conclude our discussion with an overview of state-of-the-art systems, and a review of available resources to support further research and technology development.

Tar files containing a compilation of all the notes are available. However, these files are large and will require a substantial amount of time to download. A tar file of the html version of the notes is available [here](#). These were generated using wget:

```
wget -np -k -m http://www.isip.msstate.edu/publications/courses/ece_8463/lectures/current
```

A pdf file containing the entire set of lecture notes is available [here](#). These were generated using Adobe Acrobat.

Questions or comments about the material presented here can be directed to [help@isip.msstate.edu](mailto:help@isip.msstate.edu).



## LECTURE 01: COURSE OVERVIEW AND OBJECTIVES

- Objectives:
  - Learn about basic technology
  - Understand theory at a fundamental level
  - Relate to other theory such pattern recognition, signal processing, computational linguistics, etc.
  - Develop perspective: Are the approaches we use specific to a speech signal?
  
- What we won't do:
  - Computer programming
  - Computer simulations
  - Matlab exercises
  - Teach you how to tune parameters
  - Train you to be speech technologists...
  
- Why?

# SYLLABUS

## Contact Information:

Time	10 - 11 AM
Place	Lecture: 250 Simrall
Instructor	<b>Joseph Picone</b> Office: 413 Simrall Office Hours: 11-12 MWF (others by appt.) Email: <a href="mailto:picone@isip.msstate.edu">picone@isip.msstate.edu</a>
Class Alias	<a href="mailto:ece_8463@isip.msstate.edu">ece_8463@isip.msstate.edu</a>
URL	<a href="http://www.isip.msstate.edu/publications/courses/ece_8463">http://www.isip.msstate.edu/publications/courses/ece_8463</a>
Required Textbook(s)	X. Huang, A. Acero, and H.W. Hon, <i>Spoken Language Processing - A Guide to Theory, Algorithm, and System Development</i> , Prentice Hall, ISBN: 0-13-022616-5, 2001.
Prerequisite	S.J. Orfandis, <i>Introduction to Signal Processing</i> , Prentice-Hall, ISBN: 0-13-209172-0, 1996.
Reference Textbook(s)	F. Jelinek, <i>Statistical Methods for Speech Recognition</i> , MIT Press, ISBN: 0-262-10066-5, 1998.  J. Deller, et. al., <i>Discrete-Time Processing of Speech Signals</i> , MacMillan Publishing Co., ISBN: 0-7803-5386-2, 2000.  S. Pinker, <i>The Language Instinct: How the Mind Creates Language</i> , Harperperennial Library, ISBN: 0-0609-5833-2, 2000.  D. Jurafsky and J.H. Martin, <i>SPEECH and LANGUAGE PROCESSING: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition</i> , Prentice-Hall, ISBN: 0-13-095069-6, 2000.  S. Furui, <i>Digital Speech Processing, Synthesis, and Recognition</i> , Marcel Dekker, ISBN: 0-8247-0452-5, 2000.  D. O'Shaughnessy, <i>Speech Communications: Human and Machine</i> , IEEE Press, ISBN: 0-7803-3449-3, 2000.  L.R. Rabiner and B.W. Juang, <i>Fundamentals of Speech Recognition</i> , Prentice-Hall, ISBN: 0-13-015157-2, 1993.  L.R. Rabiner and R.W. Schafer, <i>Digital Processing of Speech Signals</i> , Prentice-Hall, ISBN: 0-13-213603-1, 1978.

## Grading Policies:

Exam No. 1	25%
Exam No. 2	25%
Exam No. 3	25%
Final Exam (Cumulative)	25%

Exams will be closed books and notes. You will be allowed to bring in one 8.5 x 11 in. (double-sided) sheet of notes. Calculators are allowed but should not be necessary. Homework will be assigned but not collected or graded.

**Schedule:**

The entries in the column labeled Section(s) refer to reading assignments in the course textbook.

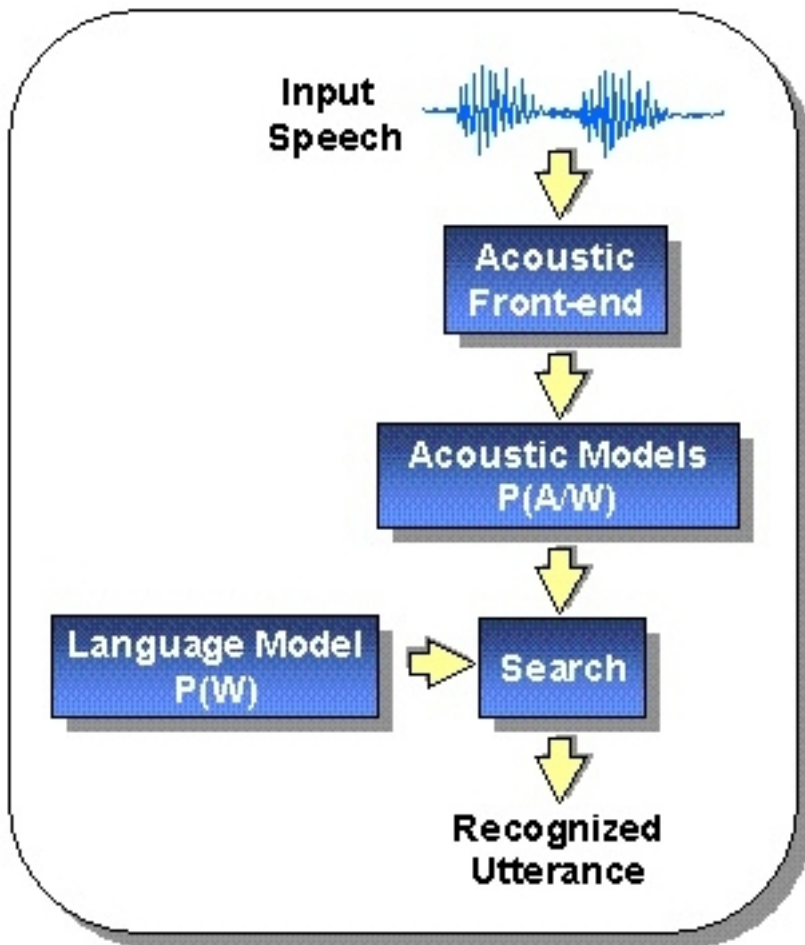
Class	Date	Section(s)	Topic(s)
01	01/07	<a href="#">1.1 - 1.5</a>	Course Overview; Introduction
02	01/09	<a href="#">2.1.2</a>	Speech Physiology
03	01/11	<a href="#">6.2</a>	Speech Production Models
04	01/14	<a href="#">2.1.3, 2.1.4</a>	Hearing Physiology
05	01/16	<a href="#">2.1.3.4</a>	Perception and Masking
06	01/18	<a href="#">2.2</a>	Phonetics and Phonology
07	01/23	<a href="#">2.3 - 2.5</a>	Syntax and Semantics
08	01/25	<a href="#">5.5, 9.3</a>	Sampling
09	01/28	<a href="#">5.6, 5.7</a>	Resampling
10	01/30	<a href="#">10.1 - 10.4</a>	Acoustic Transducers
11	02/01	<a href="#">5.4</a>	Temporal Analysis
12	02/04	<a href="#">5.1 - 5.3</a>	Frequency Domain Analysis
13	02/06	<a href="#">6.4 - 6.5</a>	Cepstral Analysis
<b>14</b>	<b>02/08</b>	<b>Lectures 1-11</b>	<b>Exam No. 1</b>
15	02/11	<a href="#">6.1 - 6.3</a>	Linear Prediction
16	02/13	<a href="#">6.5.3</a>	LP-Based Representations
17	02/15	<a href="#">9.3.3</a>	Differentiation
18	02/18	<a href="#">9.3.4</a>	Principal Components
19	02/20	<a href="#">9.3.4</a>	Linear Discriminant Analysis
20	02/22		
21	02/25		
22	02/27		
23	03/01		

24	03/04		
25	03/06		
26	03/08		
27	03/18		
28	03/20		
29	03/22		
30	03/25		
<b>31</b>	<b>03/27</b>	<b>Lectures 12 - 28</b>	<b>Exam No. 2</b>
32	04/01		
33	04/03		
34	04/05		
35	04/08		
36	04/10		
37	04/12		
38	04/15		
39	04/17		
<b>40</b>	<b>04/19</b>	<b>Lectures 29 - 37</b>	<b>Exam No. 3</b>
41	04/19		
42	04/22		
43	04/24		
44	04/26		
45	04/29		
<b>46</b>	<b>05/06</b>	<b>Cumulative</b>	<b>Final Exam (8 - 11 AM)</b>

**Homework:**

No.	Due Date	Description
1	01/21	<a href="#">Speech Production</a>
2	01/28	<a href="#">Speech Perception</a>
3	02/04	<a href="#">Linguistics</a>

## HUMAN LANGUAGE TECHNOLOGY: SPEECH RECOGNITION IS MULTIDISCIPLINARY



- Acoustic Front-End: Signal Processing
- Acoustic Models: Pattern Recognition, Linguistics
- Language Model: Natural Language Processing
- Search: Computational Linguistics, Cognitive Science



## Automatic Speech Recognition

### About our Software

The primary goal of our Internet-Accessible Speech Recognition Technology project is to create a [freely available](#), modular, state-of-the-art speech recognition system that can be easily modified to suit your research needs. The system is built on top of a vast hierarchy of general purpose [C++ classes](#) that implement generic math, data structure, and signal processing concepts. Click here for an official version of our [public domain license](#) statement.

In addition to the core recognition software, we distribute other tools that minimize the effort required to use our system to recreate important benchmarks:

- [Software](#)
- [Documentation](#)
- [Tutorials](#)
- [Demonstrations](#)

### Featured Photo



[Click to enlarge.](#)

Aravind Ganapathiraju successfully defends his thesis titled [Support Vector Machines](#) and becomes ISIP's first Ph.D. graduate. Aravind was one of the original ISIP members and was a key contributor to most of the software you find on this web site.

### This Month's Guided Tutorial

with Joe Picone

Hi, my name is Joe Picone, and I'm the ISIP Director. (Aside from getting to know the root password, I am not sure there are many benefits to this job.) Welcome to the first guided tutorial of our new speech site. An ISIP member will take you on a new tutorial every month.



This month, I will introduce you to two annual workshops we host at [Mississippi State University](#) that provide an opportunity for hands-on learning about our system. These workshops are attended by a diverse audience ranging from entry-level graduate students to senior professionals. Let's [begin](#) with a simple introduction to speech recognition.

[\(Visit the guided tutorial archive...\)](#)

### Speech Recognition Resources

- [Data, Models, Lexicons, and Other Language Resources](#)
- [Performance and Computing Resource Guidelines](#)
- [Recent Research and Evaluation Results](#)
- [Courses, Seminars, and Presentations](#)
- [Relevant Internet-Based Resources](#)
- [Famous Moments in Our History](#)
- [What's New On This Web Site](#)
- [Search Our Web Site](#)

## Software

### About our Software

Our [vision](#) stems from the fact that research commonly suffers from a creative backlog due to rewriting of common functions, and the time spent in debugging such things as file I/O. The ISIP Foundation Classes (IFCs) and software environment are designed to meet this need, providing everything from complex data structures to an abstract file I/O interface.

Our Prototype System is supported across a wide range of platforms including [Sun Solaris](#), [Linux](#), and [Cygwin](#) on [Windows](#) computers, as long as the minimum [software](#) and [hardware requirements](#) are met. The latest version of our Prototype System can be downloaded by following our [CVS instructions](#). Then follow the simple [quick start](#) guide and you will be on your way.

### Download Our Software

- (02/15/02) [Production System \(v0.0\)](#): A research environment that includes a generalized hierarchical Viterbi search-based decoder. Recommended for serious speech and signal processing researchers.
- (09/14/01) [Prototype System \(v5.11\)](#): A cross-word context-dependent LVCSR system. Recommended for speech technologists and application developers.
- (11/29/00) [TIDIGITS Toolkit \(v5.7\)](#): An easy-to-use toolkit that demonstrates the essential steps in building a state-of-the-art speech recognition system. Recommended for novices.

Visit our software release [archive](#) for previous release information.



Brain: The same thing we do every night, Pinky.  
Try to take over the world!

### Software-Related Resources

- [Documentation](#): html-based documentation that includes links to the actual source code.
- [Tutorials](#): step-by-step instructions for building a state-of-the-art LVCSR system.
- [Demos](#): conduct an experiment using our remote job submission facility; explore our Java applets.

Consult our [legacy software archive](#) for some of our oldies but goodies.

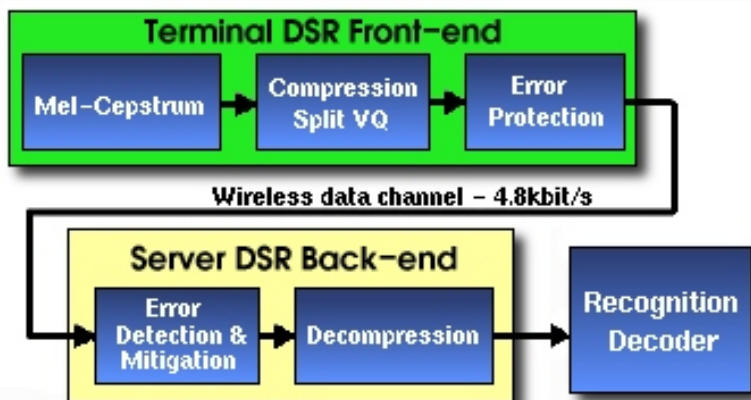
# Aurora Evaluations



[overview](#) | [downloads](#) | [publications](#) | [performance](#) | [collaborations](#)

## Mission:

Our goal is to evaluate and compare the robustness of feature extraction algorithms in the presence of background noise. These evaluations are being conducted under the [Aurora Distributed Speech Recognition](#) working group of [The European Telecommunications Standards Institute](#) (ETSI).



## What's New:

- (11/12/01) [Multiple-CPU Eval Package \(v1.4.2\)](#): Added a utility to remove silence. Fixed a minor bug in the command line interface for `lm_scale`.
- (10/31/01) [Short Training Set Definition \(v1.4.1\)](#): We have included file lists that define the two 7,138 utterance training sets.

**Overview:** Review the status of the Aurora evaluations, and learn more about the history of this project.

**Downloads:** Download the latest versions of our recognition software and evaluation scripts.

**Publications:** Access background information related to this project. View our [mailing list archive](#).

**Performance:** A quick look at the overall performance of the baseline system using a variety of robust front ends.

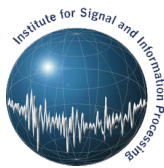
---

[Up](#) | [Home](#) | [Site Map](#) | [What's New](#) | [Projects](#) | [Publications](#)  
[Speech](#) | [Administration](#) | [About Us](#) | [Search](#) | [Contact](#)

Please direct questions or comments to [help@isip.msstate.edu](mailto:help@isip.msstate.edu)

## Technical Program:

Jon Hamaker  
[hamaker@isip](mailto:hamaker@isip)



# SPEECH RECOGNITION SYSTEM DESIGN REVIEW MISSISSIPPI STATE UNIVERSITY JANUARY 10-11, 2002



Joe Picone  
[picone@isip](mailto:picone@isip)

## Local Arrangements:

Jordan Blaize  
[blaize@isip](mailto:blaize@isip)

## Registration:

Jim Sesser  
[sesser@isip](mailto:sesser@isip)

## Program Committee:

Kaihua Huang  
[huang@isip](mailto:huang@isip)

Feng Zheng  
[zheng@isip](mailto:zheng@isip)

Naveen Parihar  
[parihar@isip](mailto:parihar@isip)



We promise that if you attend this year's Speech Recognition System Design Review that you will definitely have more fun than this guy. You might even have a better chance of catching a fish or two. This is the third in a series of design reviews intended to be chance for our users to provide input into the development of the system.

[Workshop  
Registration](#)

[Workshop  
Program](#)

[Travel  
Information](#)

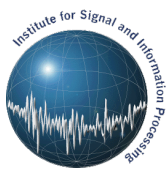
[Project  
Overview](#)

[Photo  
Album](#)

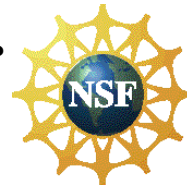
Please direct questions or comments to [help@isip.msstate.edu](mailto:help@isip.msstate.edu)

## Technical Program:

Joe Picone  
[picone@isip](mailto:picone@isip)



## SPEECH RECOGNITION SYSTEM TRAINING WORKSHOP MISSISSIPPI STATE UNIVERSITY MAY 12-18, 2002



## Local

### Arrangements:

Bill Chapman  
[chapman@isip](mailto:chapman@isip)

### Registration:

Joseph Langley  
[langley@isip](mailto:langley@isip)



## Program

### Committee:

Aravind  
Ganapathiraju  
[ganapath@isip](mailto:ganapath@isip)

Summer at Mississippi State means two things: baseball and a speech recognition short course. We invite you to attend a workshop at which we will provide hands-on training on the ISIP public domain speech recognition system. This workshop is geared towards entry-level students and professionals interested in doing research with the ISIP system.

Travel expenses for graduate students attending the workshop will be subsidized by ISIP. Seating is limited, so we encourage you to register early.

Jie Zhao  
[zhao@isip](mailto:zhao@isip)

Shivali Srivastava  
[srivasta@isip](mailto:srivasta@isip)

Ram Sundaram  
[sundaram@isip](mailto:sundaram@isip)

<a href="#">Registration</a>	<a href="#">Workshop</a>	<a href="#">Workshop Program</a>	<a href="#">Travel Information</a>	<a href="#">Project Overview</a>	<a href="#">Photo Album</a>
------------------------------	--------------------------	--------------------------------------	--	--------------------------------------	---------------------------------

Please direct questions or comments to [help@isip.msstate.edu](mailto:help@isip.msstate.edu)



# COURSES






what's new - projects - publications - speech - search - up

- [ASE 6713](#) - Introduction to Acoustics
  - [ECE 2991](#) - Unix Software Tools
  - [ECE 3111](#) - Digital Devices Design Laboratory
  - [ECE 3183](#) - Electrical Engineering Systems
  - [ECE 3283](#) - Electronics
  - [ECE 3713](#) - Digital Devices and Logic Design
  - [ECE 3813](#) - Signals and Systems
  - [ECE 4000](#) - Special Topics for Undergraduates
  - [ECE 4012](#) - Senior Design (Old)
  - [ECE 4512](#) - Senior Design I
  - [ECE 4522](#) - Senior Design II
  - [ECE 4773](#) - Introduction to Digital Signal Processing (DSP)
  - [ECE 7000](#) - Directed Individual Study
  - [ECE 7000](#) - Special Topics in Speech Recognition
  - [ECE 8463](#) - Fundamentals of Speech Recognition
  - [ECE 8990](#) - Information Theory
  - [ECE 8990](#) - Pattern Recognition
  - [ISIP 0000](#) - Fundamentals of Speech Recognition (short course)
  - [ISIP 0001](#) - Unix Software Engineering
  - [SA XXXX](#) - Starkville Academy Young Explorers Club
-

[Up](#) | [Home](#) | [Site Map](#) | [What's New](#) | [Projects](#) | [Publications](#)  
[Speech](#) | [Administration](#) | [About Us](#) | [Search](#) | [Contact](#)

Please direct questions or comments to [help@isip.msstate.edu](mailto:help@isip.msstate.edu)

# Index of /publications/courses/ece\_8463/lectures/archive

<a href="#">Name</a>	<a href="#">Last modified</a>	<a href="#">Size</a>	<a href="#">Description</a>
 <a href="#">Parent Directory</a>	28-Jan-2002 09:23	-	
 <a href="#">lectures/</a>	09-Feb-2002 09:46	-	
 <a href="#">resources/</a>	07-Jan-2002 09:34	-	

Apache/1.3.9 Server at www.isip.msstate.edu Port 80