



## Two major themes: modeling and evaluation

- **Modeling:**

- **Modeling of speaker variability is fundamental (all forms of variability)**
- **Do our models generalize?**
- **Are statistical models learning the proper things?**

- **Evaluation:**

- **Better evaluation paradigms**
- **Better metrics**



# BREAKOUT: MODELING

- **Auditory modeling: new research; has it been assimilated**
- **Human performance (quantify using such diagnostic measures as articulation index)**
- **Imaging of auditory cortex; modulation transfer functions of auditory cortex; cochlear implant analysis; timing.**
- **Real-time MRI; microbeam on articulation;**



## SUMMARY: MODELING

- **Features: area functions ok for synthesizers...**
- **Reaction times: people make very quick decisions about words; do we need to carry all this information to all levels.**
- **Feedback in hearing systems**
- **Human listening experiments (e.g. Miller/Nicely data; do we know how to run human listening experiments?)**
- **Kids vs. adults vs. elderly? Male/female? Smokers?**



## SUMMARY: EVALUATION

- **Performance as a function of signal to noise ratio; value of knowing when it goes to chance**
- **Performance as a function of training data**
- **Word error rate is not all that bad; single error rate might not be good; error analysis important**
- **more work to characterize amount of data needed for training**



## SUMMARY: EVALUATION

- **Training duty cycle: Differential in performance of a speaker in the training set vs. out of the training set as a function of the amount of data; with 2000 hours of data does this relationship change; generalization?**
- **Do current evaluations accurately assess generalization? (2000 hours of data?)**
- **Evaluation without language model**



# SUMMARY: STATISTICS

- **Discriminative training**
- **Is what we doing relevant to large-scale tasks? Will they scale?**
- **What mathematical tools are the most promising? Can we transfer this information to other areas of the community (e.g., those interested in auditory modeling)**
- **Promote more of a connection between advanced machine learning techniques and physiological models;**
- **At what places do we apply machine learning?**
- **Are we really generalizing? Test speaker similar to a training speaker or true generalization?**
- **Recognizers tend towards the average**