

ASRU'2007, Kyoto, December 10, 2007

Handling unexpected acoustic data

Hervé Bourlard

IDIAP Research Institute, Martigny, Switzerland





I DISAGREE !!!

Unexpected acoustic events (1)



- Big difference between
 - OOV (or « out-of-vocabulary acoustic data »)
 - Rare/unexpected word
- I don't know what a rare word is...
- In Bayesian classification (minimizing Bayes' risk, i.e., word error rate):
 - Words with low prior probability
 - But prior probability should be conditional on context (which we don't know how to do)

Unexpected acoustic events (2)



- In IIR:
 - Most (70%) words are rare: Zipf's law with very long tail!!!
 - Focus on *information* instead of
 - Optimizing Recall and Precision (or F measure)
 - We don't know how to train current ASR systems to optimize those cost functions
 - But perhaps not impossible....

Possible research trends



- Investigation of « unexpected » acoustic data as a pure acoustic « outsider » detector
 - Telephone vs microphone speech
 - Normal speaking style vs shouting, etc
 - Talking to a computer or to other human beings
- Context awareness
 - Environmental
 - How to automatically identify (adapt to) « context »???
 - Tasks:
 - Dictation, command and control, etc
 - Human-human or human-computer dialogue systems
 - Distillation/information extraction