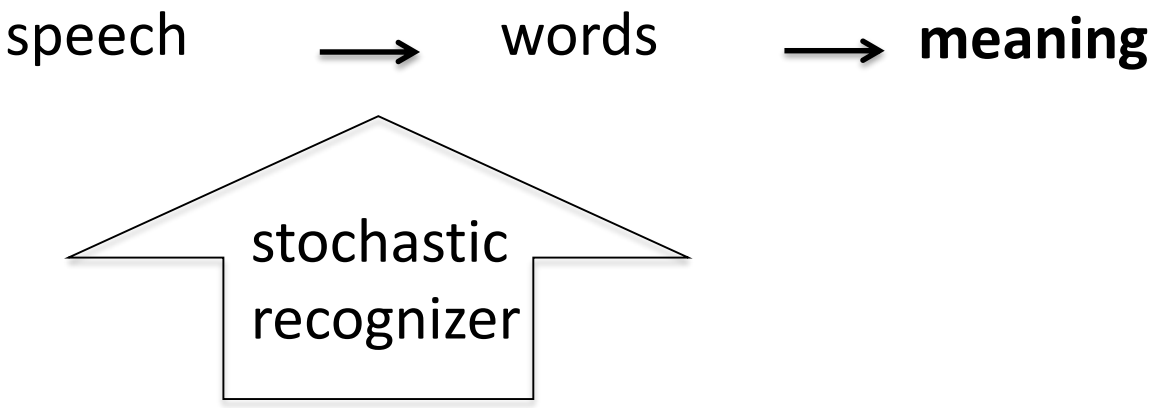
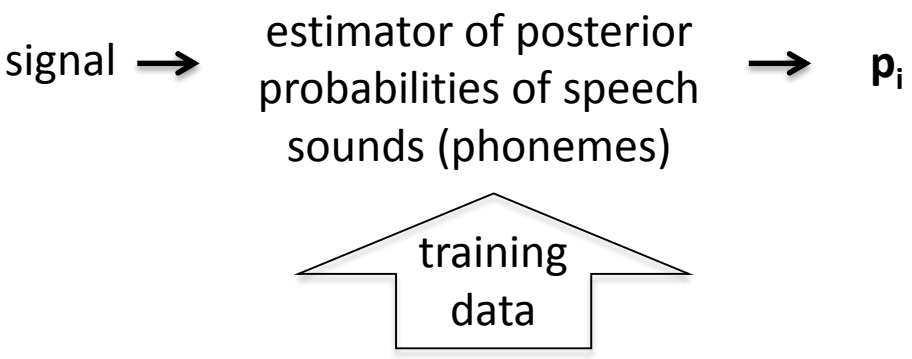
Towards machines, which know when they know and when they do not know

Hynek Hermansky

Johns Hopkins University
and
Brno University of Technology


speech
words
meaning
stochastic
recognizer

The Problem



Q: How good is the estimate?

- can we believe the estimate ?
- can we improve the estimate (unsupervised adaptation)

signal
(speech spectrum)

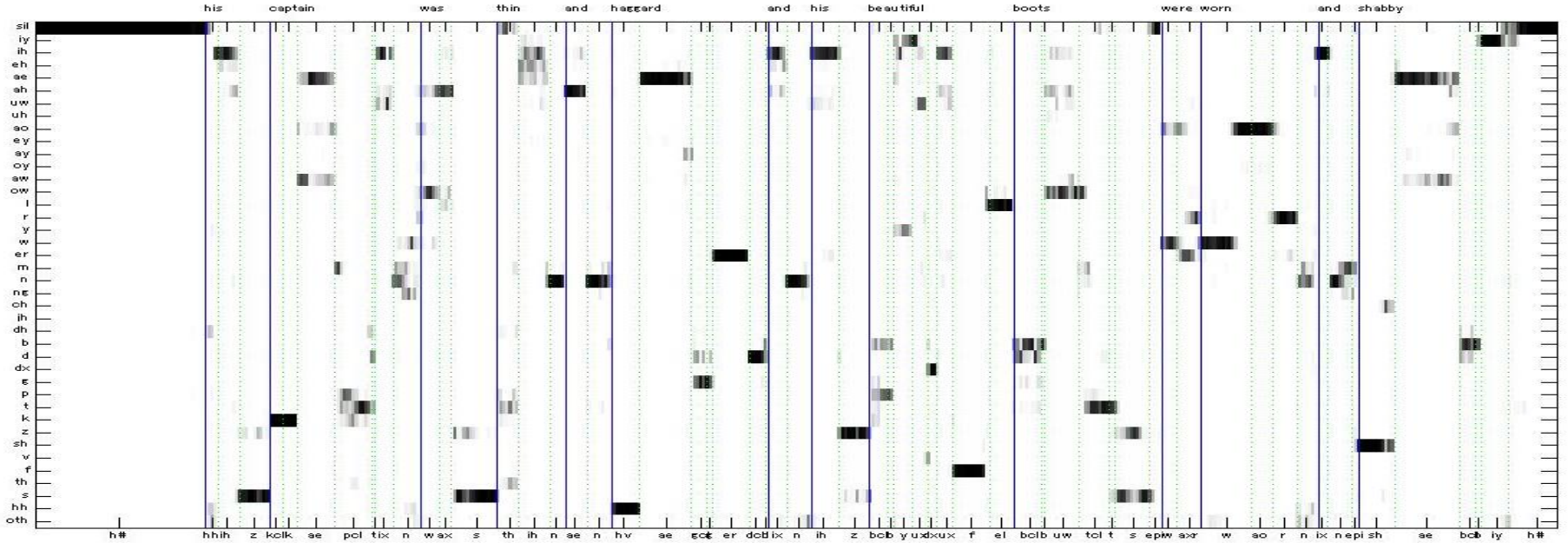
→

hierarchical deep
neural network
estimating posteriors
of speech sounds

$\mathbf{p}_i$

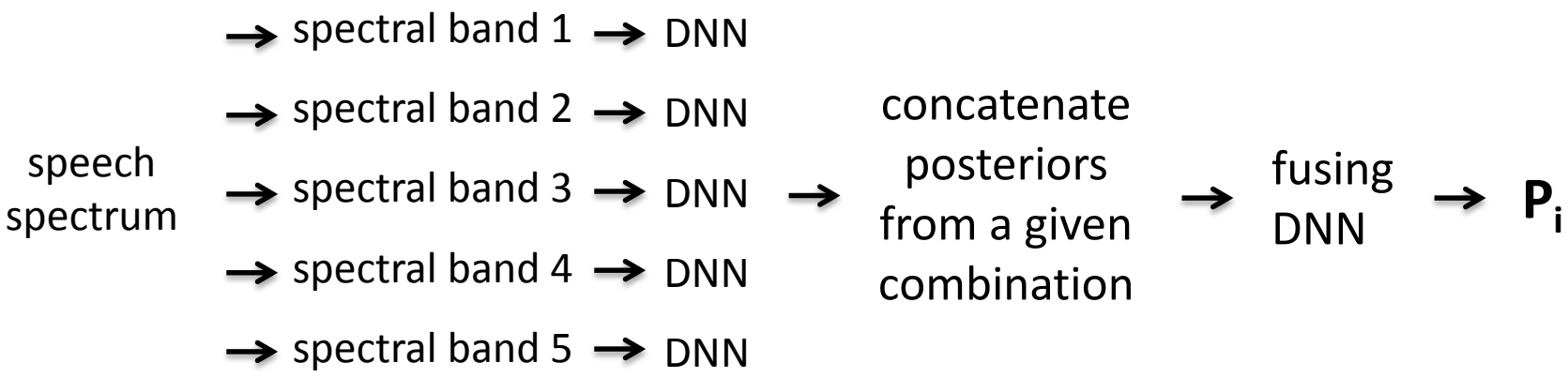
→

posteriogram (vectors of $\mathbf{p}_i$s in time)



- Do not know the correct answer (ground truth) but want to know
  - How good or bad could be the result ?

hierarchical deep neural network



- DNN – 4 hidden layer, estimating posteriors of context dependent states oh HMM and merging them to estimates of context independent phonemes
- fusing DNN – 4 hidden layer, input are concatenated posteriors from the first stage DNNs, outputs are posteriors of context independent phonemes

training:
full combination (all frequency bands)
2 training conditions (clean speech and speech corrupted by broadband noise with varying SNRs)

test:
2^5-1 = 31 possible combinations per training condition (all together 62 estimators)

Where could we get the necessary information?

1. from the output of the hierarchical DNN estimator
2. from data on hidden layers of the estimator
3. overall likelihood of the best sound sequence

evaluation criteria:

- how well our measures correlate with
  - errors in recognition of phoneme sequences ?
  - comparisons with ideal posteriograms representing hand-labeled speech?
  - comparisons with posteriograms derived from training-like data
- ability to pick up the least corrupted stream in adaptive multi-stream system

although some sort of the **computer** can either way

hopefully cin-cin o-bi **computer** connected with

Machines which know when they do not know !

- **He who knows not and knows not he knows not, he is a fool—shun him;**
- He who knows not and knows he knows not, he is simple—teach him;
- He who knows and knows not he knows, he is asleep—wake him;
- He who knows and knows he knows, he is wise—follow him!

Machines, which know when they know, and know when they do not know