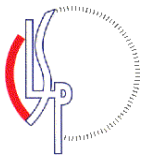
Yasuhiro Minami

WS97 Presentation

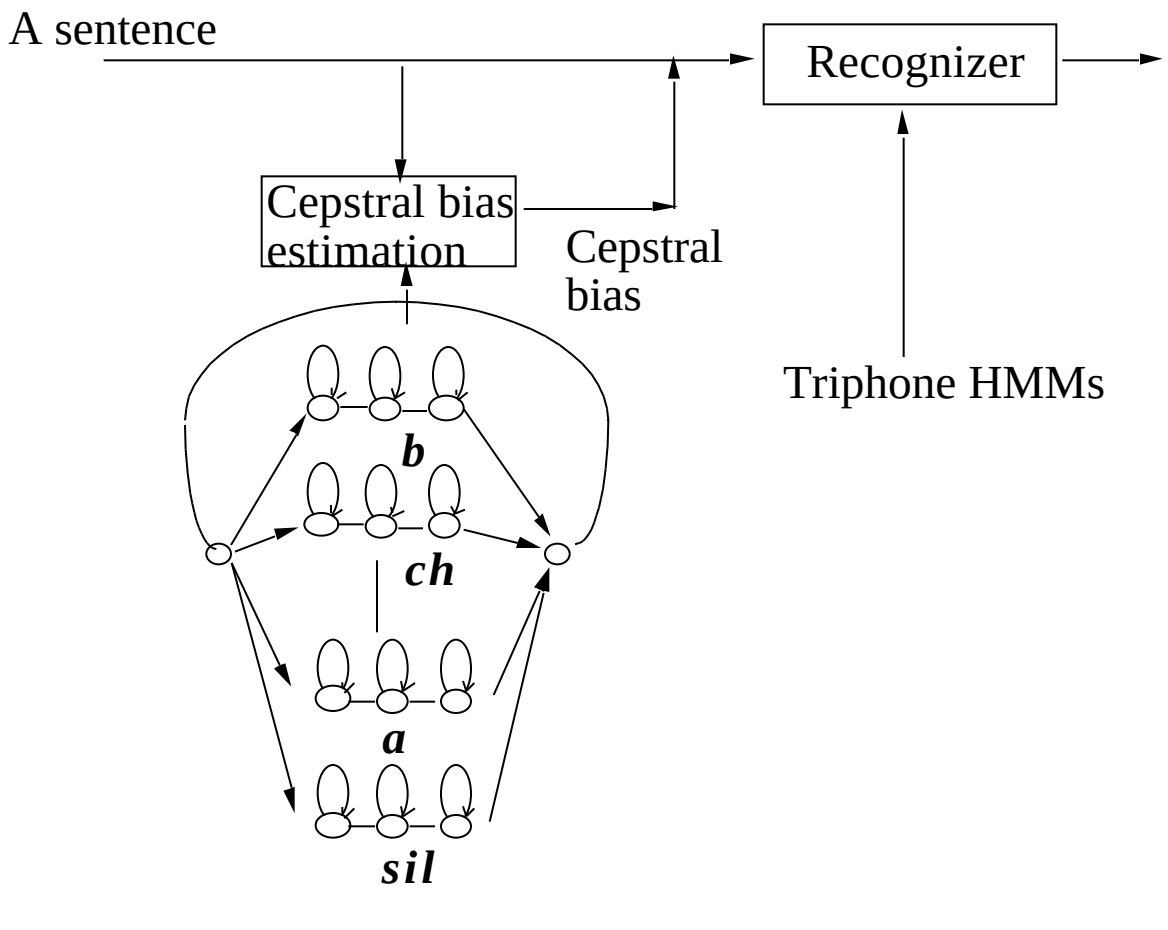


Cepstral bias adaptation

Objective: Compensating channel distortion.

Corrupted speech

Block diagram for Cepstral bias adaptation



Formulation of cepstral bias estimation

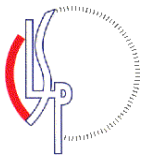
$$x_t = y_t + b$$

$$b = \arg\max_b P(X|b,M)$$

X = (x_1,x_2,…, x_T): Input data

M: HMMs

$$b'_i = \frac{\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{m=1}^{M}\gamma_t(n,m)\frac{y_{t,i}-\mu_{n,m,i})}{\sigma^2_{n,m,i}}}{\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{\gamma_t(n,m)}{\sigma^2_{n,m,i}}}$$

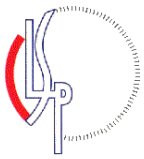


Recognition experiments

Task: WS96 Dev Set from Switch Board
Baseline:
HMM: WS96 Word Internal triphone models
(from syllable team)
Decoder: HVite word internal decoding

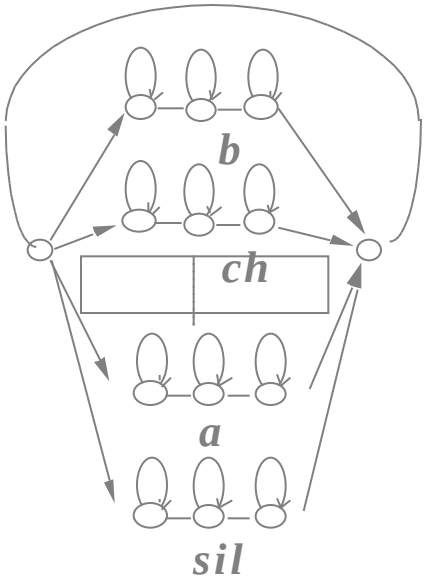
Word Error rate

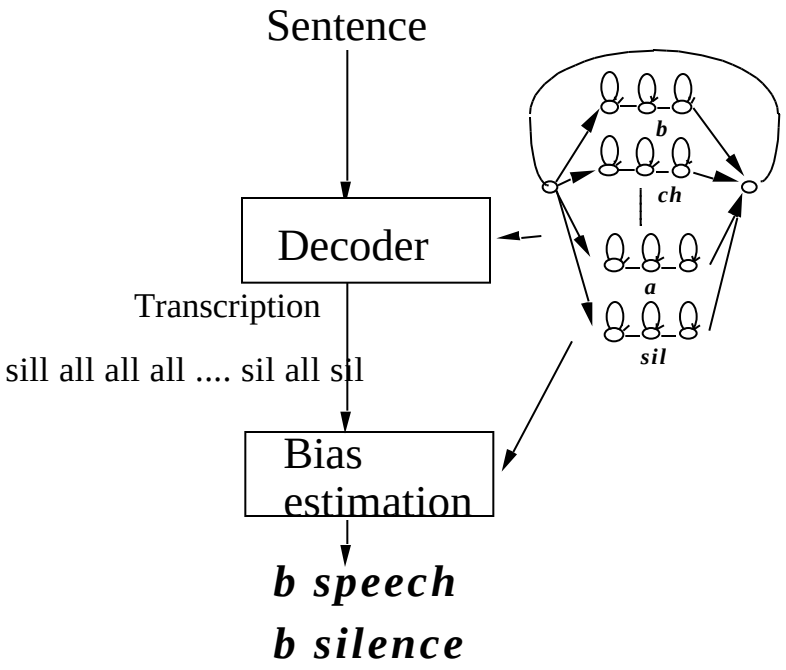
	Short Test Set	Complete Test Set
Baseline	52.2%	52.3%
Cepstral bias adaptation	50.5%	52.1%



SWITCHBOARD SPECIFIC ISSUES

Because we need **robust** and **fast** estimation of models we use:

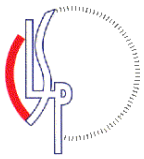




SHOULD BE BETTER

Sentence Duration Statistics

Length in seconds	# of sentences
0 ≤ sentence length ≤ 0.5	141
0.5 ≤ sentence length ≤ 1.0	502
1.0 ≤ sentence length ≤ 2.0	512
2.0 ≤ sentence length ≤ 3.0	362
3.0 ≤ sentence length	602



Future Work

- Delta Cepstral adaptation, Delta-Delta Cepstrum adaptation
- Noise adaptation