

# BN-E Experiments in Cambridge

Do Yeong Kim, Mark Gales, and Phil Woodland

March 2005



Cambridge University Engineering Department

## Overview

- Increased Acoustic Model Training Data
- Increased Model Parameters
- SPron Model Building using Increased Data
- $10\times$ RT System Performance with the New Acoustic Models



## Increased Training Data

- Acoustic model training data

Data		size(hrs)	
bnac	Hub4	144	
tdt4	Oct 00 - Jan 01	231	
tdt4a	Mar 01 - Jul 01	377	bntr04-1350h (used in RT04)
bn03_1	Mar 03 - Nov 03	300	
bn03_2	Mar 03 - Nov 03	300	
bn03_3	Mar 03 - Nov 03	440	<i>newly added</i>
tdt2	Feb 98 - Jun 98	420	

- bn03\_3 was chosen from 13 sources using BBN's transcriptions.
- tdt2 transcriptions were generated using the P1+P2 system with a biased language model and the RT04 acoustic models.
- bntr05-1790h: bntr04-1350h + bn03\_3
- bntr05-2210h: bntr05-1790h + tdt2



## Increased Model Parameters

- The number of model parameters were increased using bntr04-1350h set.

	#Gaussians	#States	dev04	eval03	dev04f	eval04
MLE	302k	9k+	16.0	13.9	24.8	—
	415k	9k+	16.0	13.5	24.3	—
	398k	12k+	16.1	13.6	24.5	—
MPE	302k	9k+	13.2	11.2	19.6	—
	415k	9k+	12.8	11.1	19.4	—
	398k	12k+	13.1	11.1	19.5	—
P1+P2	302k	9k+	11.1	8.6	15.9	13.6
	415k	9k+	11.1	8.4	15.9	13.5
	398k	12k+	11.1	8.4	15.7	13.5

MLE & MPE: single pass decoding with RT03 trigram LM and RT03 CUED segment.

P1+P2: P2-fg-cn in the RT04 10×RT contrast system.

- Small but consistent gains in P1+P2 as well as single pass decoding.
- Similar gains from increasing Gaussians/state or the number states.



## Performance Improvement using Increased Data

	training set	dev04	eval03	dev04f	eval04
MLE	bntr04-1350h	16.0	13.9	24.8	—
	bntr05-1790h	16.1	13.7	24.4	—
	bntr05-2210h	16.0	13.6	24.5	—
MPE	bntr04-1350h	13.2	11.2	19.6	—
	bntr05-1790h	13.0	11.0	19.3	—
	bntr05-2210h	12.9	11.1	19.1	—
P1+P2	bntr04-1350h	11.1	8.6	15.9	13.6
	bntr05-1790h	11.1	8.5	15.7	13.5
	bntr05-2210h	11.0	8.4	15.6	13.2

MLE & MPE: single pass decoding with RT03 trigram LM and RT03 CUED segment.

P1+P2: P2-fg-cn in the RT04 10×RT contrast system (all models have 302k Gaussians).

- Gains of 0.1-0.5% abs in single pass decoding with MPE model.
- Consistently improved performance in P1+P2 framework, and a gain of 0.4% abs on eval04.



## SPron Model Building using Increased Data

- SPron acoustic models were trained on bntr05-2210h set.
  - the same SPron dictionary and model list as bntr04-1350h.

		dev04	eval03	dev04f
MLE	MPron	16.0	13.6	24.5
	SPron	15.6	13.5	24.2
MPE	MPron	12.9	11.1	19.1
	SPron	12.7	10.8	18.8

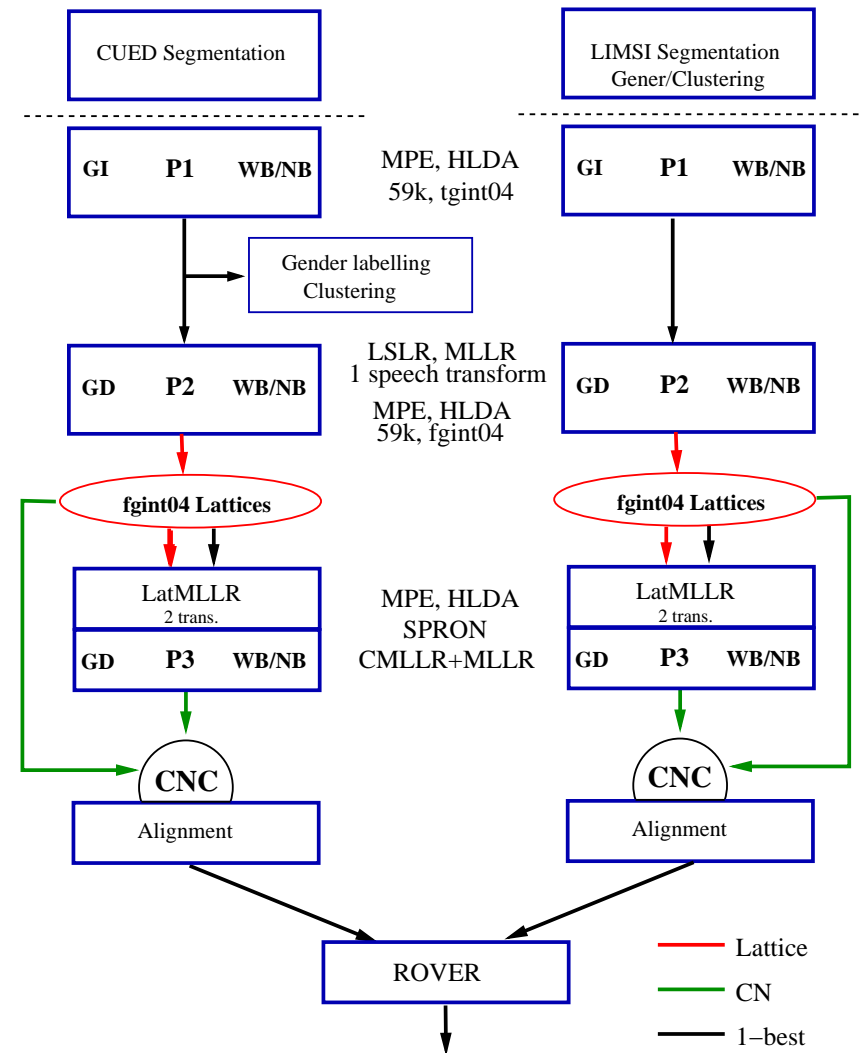
MLE & MPE: single pass decoding with RT03 trigram LM and RT03 CUED segment.  
Both MPron and SPron model were trained on bntr05-2210h set and have 302k Gaussians.

- MPE trained SPron model showed 0.2-0.3% abs better performance than MPron in single pass decoding.



# RT04 10×RT Primary System : Dual Segmentations

- Two separate sub-systems:
  - sub-system 1: CUED segmenter
  - sub-system 2: LIMSI segmenter
- Each sub-system:
  - P1: GI MPron (no fg expansion)
  - P2: GD MPron
  - P3: GD SPron
  - CNC using P2 and P3
- Combining outputs using ROVER
- Ran in 9.9×RT on eval04



## System Performance

Seg	bntr04-1350h		bntr05-2210h	
	dev04f	eval04	dev04f	eval04
CUED	15.5	13.3	15.3	12.8
LIMSI	14.9	12.8	14.8	12.5
ROVER	14.7	12.6	14.5	12.3

- Each sub-systems showed 0.1-0.5% abs gains.
- Combination of dual segmentations still work.
- 0.2-0.3% abs gains after ROVER.





## Conclusion

- New acoustic models were built with additional training data. The amount training data is now 2210hrs in total.
- The new models gave consistent gains up to 0.3% abs in MLE and 0.1-0.5% abs in MPE training with the same number of model parameters.
- The new models also showed 0.4% abs better performance on eval104 in P1+P2 framework with adaptation.
- SPron models were built using 2210hrs training set and evaluated in the RT04 10×RT (dual segments) systems framework and showed 0.3% abs gain on eval104.

