

---

---

# Jan 2000 European Trip report: THISL & RESPITE

Dan Ellis  
International Computer Science Institute, Berkeley CA  
<dpwe@icsi.berkeley.edu>

## Outline

- 1 Thisl final project meeting (BBC Kingswood):**
  - final demonstrator
  - exotic data
  - SAVANT follow-on
  
- 2 Respite year 1 mtg (Euroforum Luxembourg):**
  - multistream for Aurora
  - other research, issues



---

---

1

## THISL

### (Thematic Indexing of Spoken Language)

- **Spoken document retrieval of BBC Broadcast News**
  - automatic off-air recording of 3-6 hrs daily news
  - ASR → IR index, RA encode → audio archive
  - web-based query & retrieval
- **Partners:**
  - Sheffield Univ (+ICSI), Softsound (ajr), BBC, IDIAP, FPMs, Thomson
- **Notable successes:**
  - live archive of ~3 yr (1000s of hours)
  - BBC sound archives very positive
  - continue & broaden operation
- **Final meeting; project ran 1997feb-2000jan**



# ICSI's SQI GUI

- Spoken queries promised in proposal  
.. although text-based web interface most used

File Options

## THISLgui IR demo

Status: **listening**

IR Status: recognizing...

Enter query:  History

Start date: 1997 June 25

End date: 1998 January 31

Programs: ABC World News Tonight, CNN Early Prime, CNN Headline News, CNN Primetime News, CNN The World Today

Results for:

Program	Date	Offset	Context
<input checked="" type="checkbox"/> CNN Headline News	1998jan11	08:00	british prime minister <b>tony</b>
<input type="checkbox"/> CNN The World Today	1997sep08	27:09	the royal family prime mini
<input type="checkbox"/> CNN The World Today	1997dec25	08:09	mascot for both prime mir
<input type="checkbox"/> PRI The World	1997oct17	43:58	monday <b>tony</b> account jou
<input type="checkbox"/> CNN The World Today	1997aug30	28:30	a quote very hard to <b>tell</b> e
<input type="checkbox"/> CNN Primetime News	1997dec30	10:48	the blast is national corres
<input type="checkbox"/> PRI The World	1997oct20	45:57	and our first installment of

Program: CNN Headline News Date: 1998jan11 File: ed980111 Stop playback

08:00 manslaughter for his part in the attack c. n. n.'s **tony** dark says four of the twelve jurors favored life in prison and to believe nichols had a minor role in the attack because of the deadlock judge richard may eighth will decide nichols played by law the judge cannot impose the death penalty

08:00 british prime minister **tony blair** says his country must not allow extremists to dictate the pace of peace talks with northern ireland two gunmen killed a bouncer in a drive by shooting in downtown helfast nightclub today investigators believe protestant gunmen

- + Thomson NLP ...

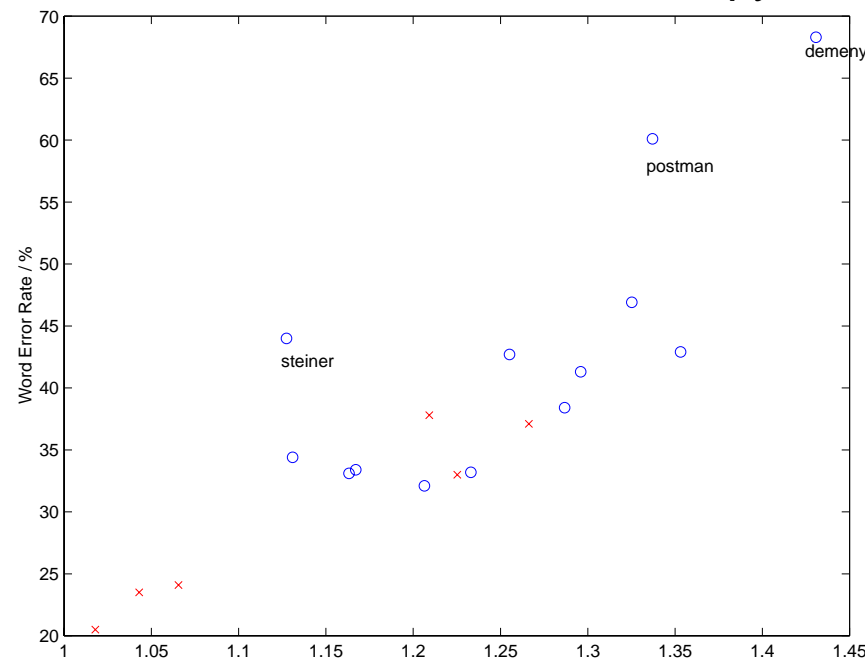


## Including out-of-domain ('exotic') data

- Suitability for non-news material?>
  - e.g. natural history, interviews, features:

<i>Data set</i>	<i>Words</i>	<i>WER</i>	<i>OOV</i>	<i>Av.Fr.Ent'py</i>
6 TV & radio news	31k	29.2 ±7.6 %	0.84%	1.14 ± 0.11
Exotic (13 varied files)	44k	38.9 ±8.4 %	0.70%	1.25 ± 0.09

- correlation of WER & av. entropy



---

---

## **SAVANT (formerly ThisI-2)**

- **Proposal to EU for follow-on to ThisI**
  - BBC very keen
  - ThisI seen as success
- **Same team (almost)**
  - Sheffield, ICSI, Cambridge, BBC, IDIAP  
+ ITC-IRST, Intrasoft, Tecmath
- **New emphases:**
  - video (database, keyframes, cut detection)
  - nonspeech audio ('actualities')
  - information structuring  
(speaker turns, program structure)
  - summarization
  - filtering & retrieval
- **Proposal submitted Jan 17th**



---

---

## 2

# RESPITE

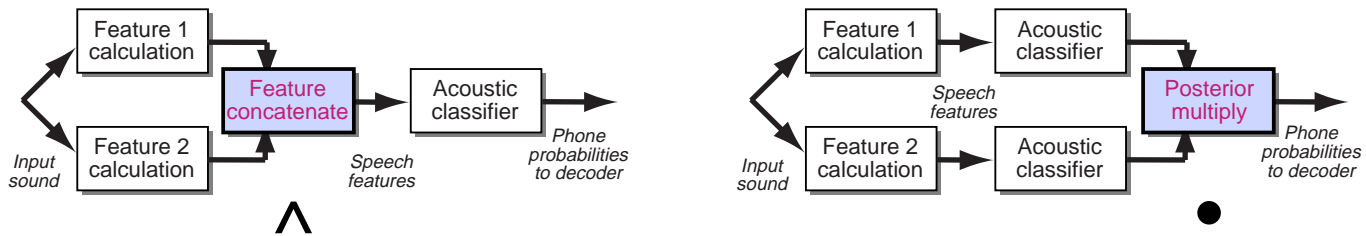
## (Recognizing Speech by Partial Info. Techs.)

- **Multistream & missing data informed by CASA, SNR estimation, confidence**
  - plus putting it all together
  - target application: in-car voice dialling
- **Partners:**
  - Sheffield (Phil Green), ICSI, IDIAP, FPMs, ICP-Grenoble, Matra-Nortel, DaimlerChrysler
- **Duration: Jan 1999 - Dec 2001**
  - first year-end meeting
  - held at European Commission in Luxembourg
  - informally met new project officer



# Combining feature streams

- **How to allocate feature dimensions to models?**
  - lower-dimension models train more quickly
  - higher-dimension models find more interactions



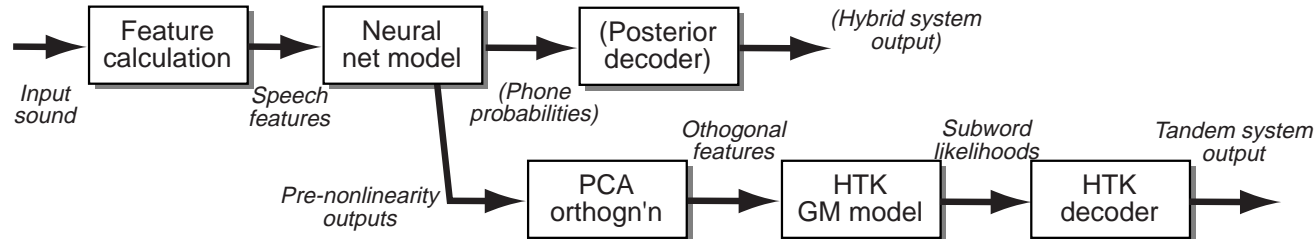
- **Variations of PLP & MSG for Aurora:**

<i>Features</i>	<i>Parameters</i>	<i>baseline WER ratio</i>
plp12•dplp12	136k	97.6%
plp12^dplp12	124k	89.6%
msg3a•msg3b	145k	101.1%
msg3a^msg3b	133k	85.8%
plp12•dplp12•msg3a•msg3b	281k	76.5%
plp12^dplp12^msg3a^msg3b	245k	74.1%
plp12^dplp12•msg3a^msg3b	257k	63.0%



# Tandem connectionist models

- **Posterior combination for HTK systems?**
- **Answer: use posteriors as HTK input features**



- (GMM system does not know they are phones)
- **Result: better performance than either alone!**
  - neural net has trained discriminatively
  - GMM HMMs learn context-dependent structure
  - extract complementary info from training data

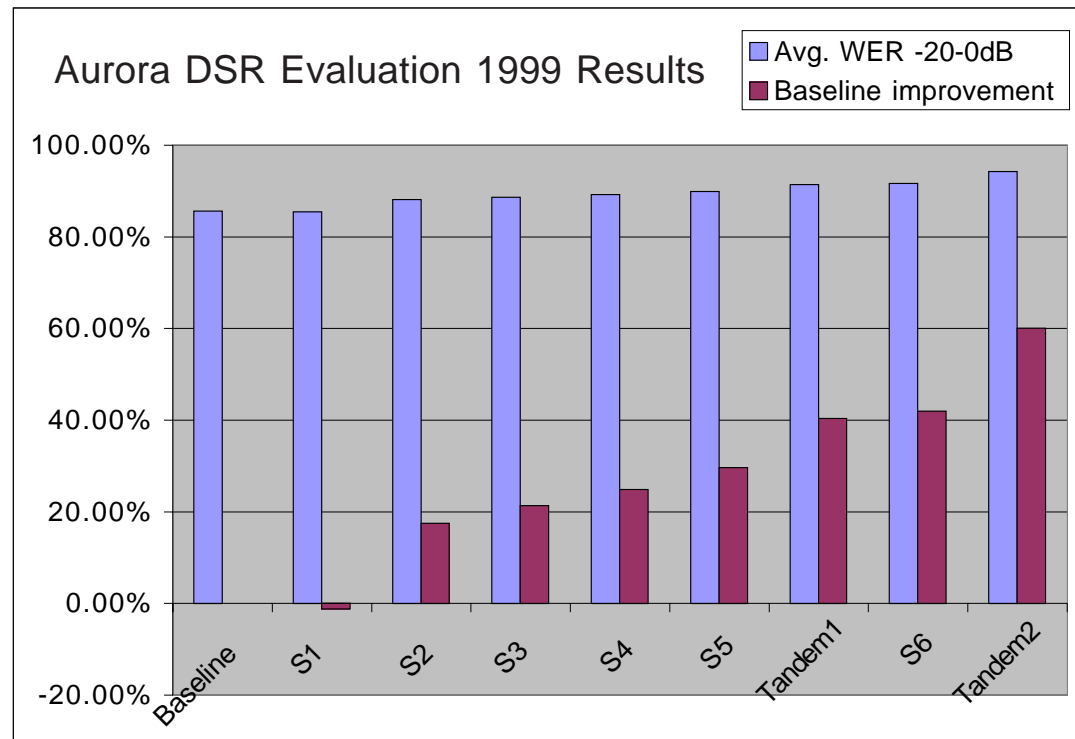
<i>System-features</i>	<i>baseline WER ratio</i>
HTK-mfcc	100.0%
Hybrid-mfcc	84.6%
Tandem-mfcc	64.5%
Tandem-plp+msg	47.2%





# Aurora “Distributed SR” evaluation

- 7 telecoms company submissions:



- Tandem systems from OGI-ICSI-Qualcomm
- **Best features for transmission?**
  - (filtered) subband energies may be sufficient



# Other RESPITE issues

- **Demonstrator: integrate w/ commercial sys?**
- **Research presentations:**
  - Herve Glotin, ICP-Grenoble: CASA labeling
  - Andy Morris, IDIAP: Full-comb mu-band weights
  - Herve Boulard: HMM-squared
  - Christophe Ris, FPMs: SNR est. for missing data
  - Sheffield: missing data with deltas
  - Jon Barker, Sheffield: CASA toolkit

