

Recording Situated Human Communications

John H.L. Hansen

John.Hansen@utdallas.edu



Center for Robust Speech Systems (CRSS)
Erik Jonsson School of Engineering & Computer Science
Department of Electrical Engineering; University of Texas at Dallas
Richardson, Texas 75083-0688, U.S.A.



John.Hansen@utdallas.edu

Slide 1

IEEE ASRU-07 PANEL

Dec 12, 2007



PANEL QUESTIONS to ADDRESS

- ◆ **Describe Data Collection**
recording devices, physical & social context, type of modes, amount & quality data
- ◆ **Goals**
Specific goals of data collection? Were they met?
- ◆ **Novel Aspects**
most innovative or unusual things about your data
- ◆ **Actual Uses / Distribution**
actual uses of data to date? used for other purposes?
Others using data? is it publicly available? whom should they contact?
- ◆ **People Involved & Cost**
people & cost involved in the data collection



John.Hansen@utdallas.edu


Slide 2

IEEE ASRU-07 PANEL

Dec 12, 2007



OVERVIEW OF CORPORA

-  RSPL – Robust Speech Processing Laboratory
-  RSPG – CSLR: Univ. of Colorado at Boulder
-  CRSS – Center for Robust Speech Systems

◆ SUSAS – Speech Under Simulated & Actual Stress:

speech under stress & emotion: speaking styles (fast, slow, soft, loud, clear), anger, computer tasks, Lombard effect, rollercoaster rides, etc. – Acoustic & Transcribed [LDC]

◆ ACCENT: 441 subjects; American English + 15 accents

RSPL-96: 240 Subjects, 4 core + sampling other accents American English; RSPG-01: 181 Subjects, 5 core + sampling (1-3 sessions)

◆ NGSW: National Gallery of the Spoken Word

60,000 Hrs; last 110 years; historical archives; US Presidents; etc.

www.ngsw.org <http://speechfind.utdallas.edu>



OVERVIEW OF CORPORA

-  RSPL – Robust Speech Processing Laboratory
-  RSPG – CSLR: Univ. of Colorado at Boulder
-  CRSS – Center for Robust Speech Systems

◆ CU-Move: Route Navigation & ASR in Cars

500 subjects; 5 speech domains; U.S. regions; acoustic car conditions; transcription issues


◆ UT-Scope: Speech Cognitive, Physical, Emotional

Stress; (in progress): 4 domains: Lombard Effect, Physical Stress, Cognitive Stress, Emotion; ~100 subjects per area; 1-4 sessions; multi-sensor


◆ UTDrive: Driver Behavior Modeling

(in progress): international collaboration; multi-sensor; multi-modal; 100-300 subjects









CU-MOVE




CU Participants


Infinite Speech Systems
(Visteon Corporation)




CU-Move Center Members



Motorola, Human Interface Lab
(Schaumburg, IL)






LABORATORIES
HRL (Malibu, CA)





Toyota Central R&D Labs


CU-Move Corpus License Members

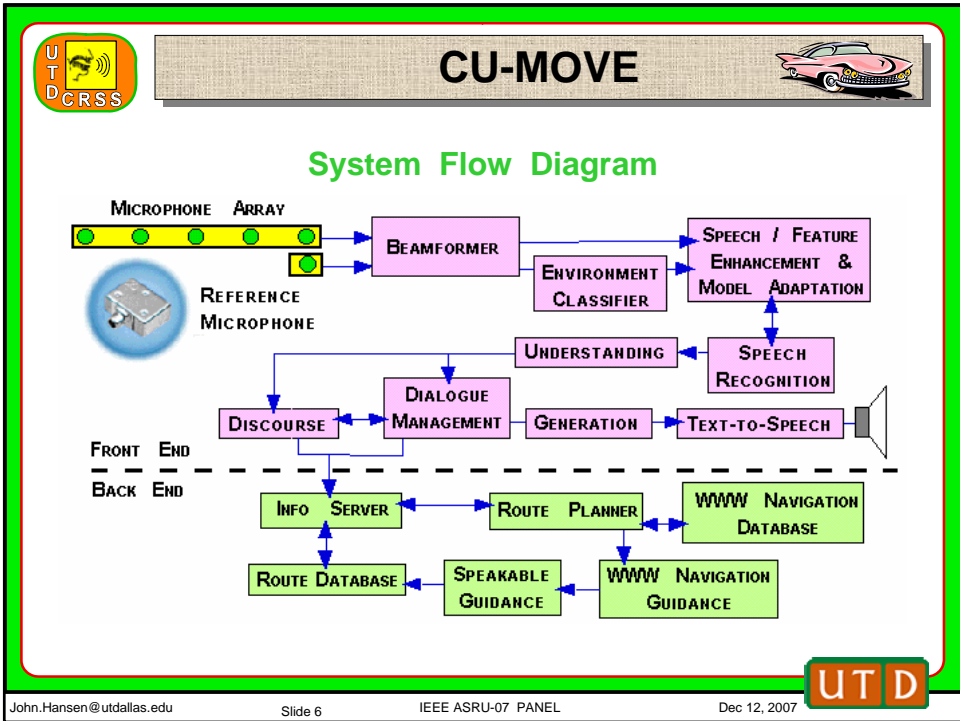
Voice Signal Technologies
(Woburn, MA)

Panasonic Speech Technology Lab
(Santa Barbara, CA)



John.Hansen@utdallas.edu
Slide 5
IEEE ASRU-07 PANEL
Dec 12, 2007





CU-MOVE



- ◆ Designed & Constructed Microphone Array & Reference
- ◆ 5 Channel Microphone Array, 1 Reference (Knowles)
- ◆ Motorola Cell Phone Microphone Channel
- ◆ AKG / OnStar Microphone (front Dashboard)



John.Hansen@utdallas.edu

Slide 7

IEEE ASRU-07 PANEL

Dec 12, 2007



CU-MOVE: Corpus Development



Web Address: <http://cumove.utdallas.edu/>

- ◆ 6 U.S. Cities: various regional dialects; +600GB Data



John.Hansen@utdallas.edu

Slide 8

IEEE ASRU-07 PANEL

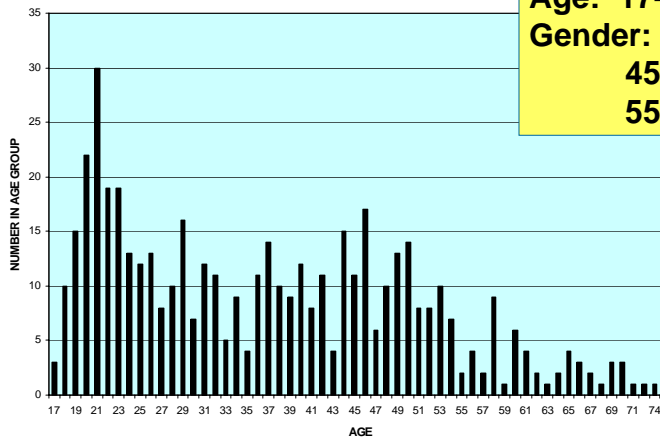
Dec 12, 2007



CU-MOVE



AGE DISTRIBUTION



Age: 17-75 yrs
Gender:
45% Male
55% Female



CU-MOVE



◆ Goals

Collect In-Vehicle Speech Data; Route Navigation;
Transcription; US Geographical Distribution

◆ Novel Aspects

Large Corpus; In-Car; Noise; Real-Dialog Conditions

◆ Actual Uses / Distribution

Released to 10 participants; Building ASR engines in
car environments; not publicly available;
contact John.Hansen@utdallas.edu

◆ People Involved & Cost

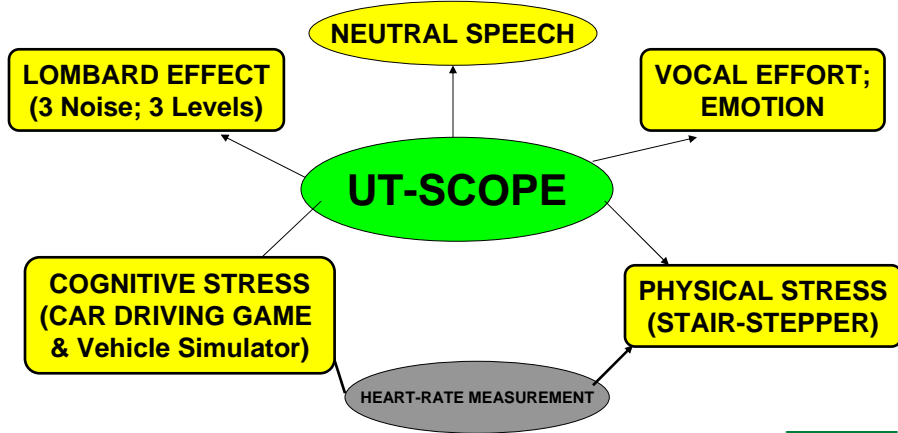
15 people (3 yr effort); cost involved: \$350k





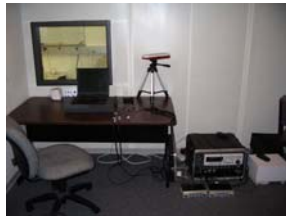
UT-SCOPE: Speech Under Cognitive & Physical Stress, Lombard Effect, Emotion

UT-SCOPE Database



UT-SCOPE: Corpus Development

◆ Data collection:



Sound booth



Instruments



Sensors



FOSTEX 8-Channel

Stair Stepper



Car Simulator





UT-SCOPE: Stress, Lombard, Emotion

(in progress)

◆ Goals

Collect Speech under Actual Cognitive, Physical, Lombard, Emotion Stress Conditions

◆ Novel Aspects

First corpus of it's kind; Calibration Test tones; multiple sessions; multiple sensors

◆ Actual Uses / Distribution

2 Phases collected; 2 being completed (not released yet – have contacted LDC)

contact John.Hansen@utdallas.edu

◆ People Involved & Cost

5 people (2 yr effort); cost involved: \$300k

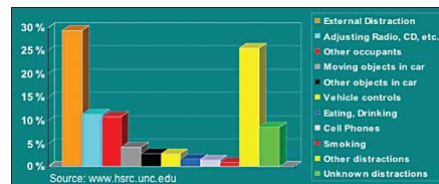
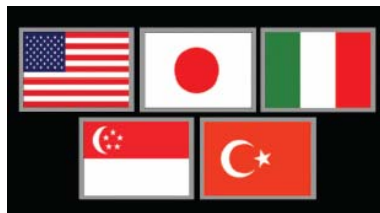


UTDrive

www.utd.edu/research/utdrive/

NEDO Funded Project “Driving Behavior”

K. Takeda, J. Hansen, H. Abut,...





Data – Collection Vehicle

Two Cameras

Microphone Array

Close-Talk Mic.

Hands-free

GPS

Distance Sensor

Heart-rate & Blood Pressure

Brake & Gas Pedal Sensors

OBD-2 (CAN-Bus)

Data Acquisition Unit



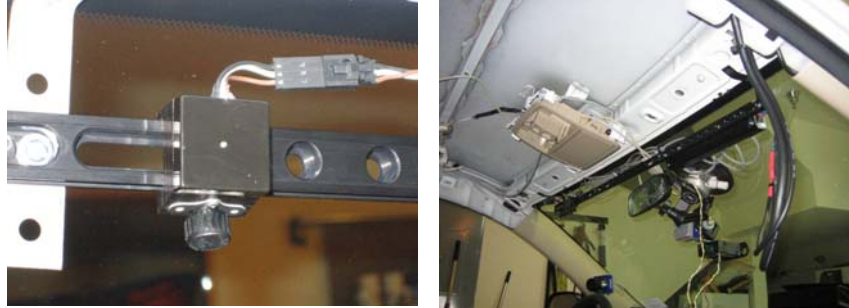
Data – Collection: Screen Output

AI 0	ACT	AI 1	ACT	AI 2	ACT	AI 3	ACT	AI 4	ACT	AI 5	ACT	AI 6	ACT	AI 7	ACT
-599	(dB)	9	(dB)	488	(dB)	-1153	(dB)	420	(dB)	1952	(V)	0393	(V)		





UTDrive – Mic Array Design



- ◆ Newport optical rail provides adjustability
- ◆ Microphones are mounted in small project boxes attached to the rail
- ◆ Shielded multi-conductor cable with drain wire connects the microphone to the preamp



John.Hansen@utdallas.edu

Slide 17

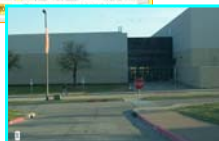
IEEE ASRU-07 PANEL

Dec 12, 2007



Data – Collection: Drive Route

Distance:
5.7miles
Time: 15 min



Distance: 7 miles
Time: 15 min



John.Hansen@utdallas.edu

Slide 18

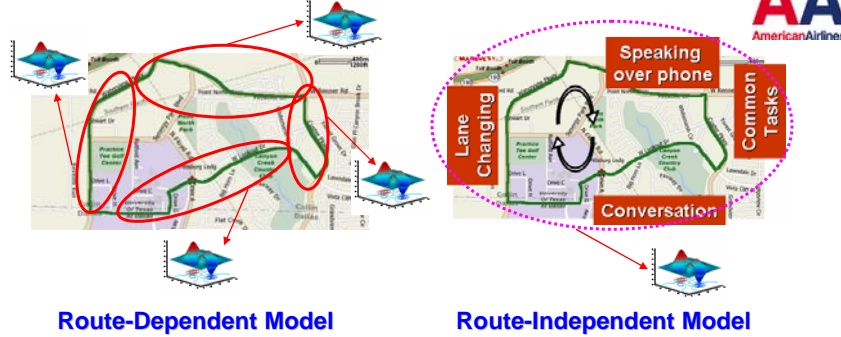
IEEE ASRU-07 PANEL

Dec 12, 2007



UTDrive: Distraction Detection

- ◆ Data: 8 Drivers
- ◆ Two GMMs: Neutral vs Distraction models
- ◆ Two modes:
 - ◆ Route-Dependent: Train & Test on the same leg of the route
 - ◆ Route-Independent: Train & Test with the whole route
- ◆ 5 seconds worth of data/token



UTDrive: Multi-Modal Transcription

Category	Event	Start Time	End Time	Transcription
Speech	Speech	00:00:00	00:00:05	...
Driving Behavior	Change Lane	00:00:10	00:00:15	...
Route Info	Change Bush Access East	00:00:20	00:00:25	...
Distraction Tasks	Speaking over phone	00:00:30	00:00:35	...

- ◆ Speech –voice dialog in car, information access
- ◆ Driver –actions (head, hands, eyes, etc)
- ◆ Car –exterior (context of road conditions, weather, etc)
- ◆ Car –CAN-bus (steering angle, vehicle speed, brake, acceleration,..)





UT-Drive: Driver Behavior

(in progress)

◆ **Goals**

Collect multiple sensor data to develop international model for driving behavior (Japan, US, Turkey, etc)

◆ **Novel Aspects**

International Collaboration; diverse sensor data; multiple sessions; transcription protocols



◆ **Actual Uses / Distribution**

Alliance of Universities working on In-Car systems; not distributed yet (NEDO supported); contact John.Hansen@utdallas.edu

◆ **People Involved & Cost**

6 people (3 yr effort); cost involved: \$+250k + growing!



OVERVIEW OF CORPORA



RSPL – Robust Speech Processing Laboratory



RSPG – CSLR: Univ. of Colorado at Boulder



CRSS – Center for Robust Speech Systems

◆ **CU-Move: Route Navigation & ASR in Cars**

500 subjects; 5 speech domains; U.S. regions; acoustic car conditions; transcription issues

◆ **UT-Scope: Speech Cognitive, Physical, Emotional**

Stress; (in progress): 4 domains: Lombard Effect, Physical Stress, Cognitive Stress, Emotion; ~100 subjects per area; 1-4 sessions; multi-sensor



◆ **UTDrive: Driver Behavior Modeling**

(in progress): international collaboration; multi-sensor; multi-modal; 100-300 subjects

