CRSS
Center for Robust Speech Systems
UTDallas: Univ. of Texas at Dallas

John Hansen

John H.L. Hansen

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

John.Hansen@utdallas.edu

ARTS
ANONYMOUS REAL-TIME SPEECH

RSTL: Robust Speech Technologies Lab

CILab: Cochlear Implant Processing Lab

UTDrive: Driver Modeling & Smart Vehicle Lab

SML: Speech & ML Lab

MSP: Multimodal Signal Processing Lab

Background

John Hansen

◈ **RSTL: Speech Processing, Speaker Analysis, Modeling, Recognition, Robustness, Diarization, ML, Forensics**

◈ **+885 Journal & Conf. Papers Published, 14 Books,…**

◈ **+$28M in external research funding**

◈ **101 PhDs/MS thesis students graduated: (9 areas) EE, CE, CS, BE,TE, Speech Science, Hearing Science, Linguistics, Cognitive Sciences**

◈ **OSAC-SPEAKER (Vice Chair); Security Clearance (+15yrs);**

◈ **Funded projects: AFRL, DARPA, CIA/TSWG, FBI, NSA, SPAWAR Systems, US-Army, NSF, NIH, NASA, CTTSO, etc.**

◈ **Hansen: faculty (35yrs) – Univ. Texas–Dallas: ECE, BE, Speech/Hearing, Univ. Colorado – Boulder: SLHS, ECE   Duke Univ.: ECE, BME, Duke Med. Cntr**

Georgia Institute of Technology · Duke University · University of Colorado Boulder · UT Dallas
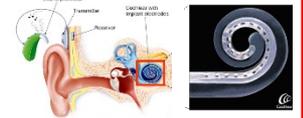
Current & Recent Research Projects

John Hansen

NIH: CCi-Mobile: ML/SP Advancements Cochlear Implants Naturalistic
$5.9M

NSF CCRI: Speech Tech & Naturalistic Audio (150Khrs)
$1.2M
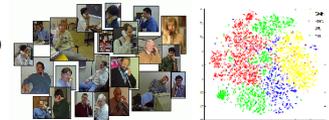
FBI/BCOE: Forensic SID / Assessment of Voice Disguise Use & Impact

CTTSO: Intrinsic Voice Variability Factors and SID

**Normalization for Robust Speech Systems
In-Set / Out-of-Set Speaker Recognition
Dialect, Language, Accent ID**
$8M
18yrs

Arabic

RATS: Robust SAD, SID, LID, KWS
$900k

ROAR-2: Speech Under Stress

TSWG/CIA: Speaker Variability Analysis: Human & Automatic SID

FBI/BCOE: Speaker Variability: Automatic Audio Stream Processing

John Hansen

ARTS — ANONYMOUS REAL-TIME SPEECH

"Are you all, like, going out tonight?" — Alice

"Um, are you all going out tonight?" — Pseudo Alice

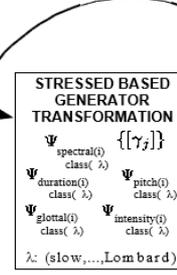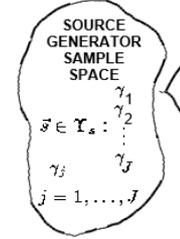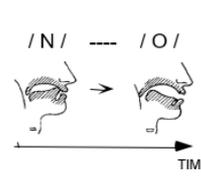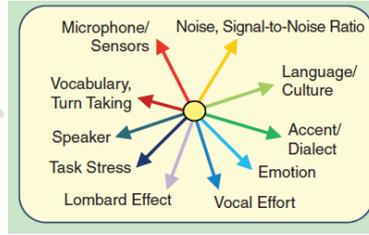◈ **+25yrs research focused on Normalization / Modification of Speaker, Speech Under Stress, Lombard Effect, Vocal Effort, Accent/Dialect, emotion, mismatch**
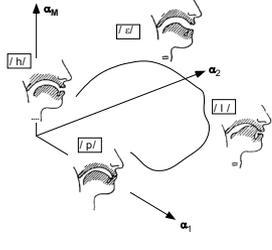
ARTICULATORY PRODUCTION SPACE

SPEECH FEATURE SPACE

"H-E-L-P" ● NEUTRAL ● STRESSED

Microphone/Sensors — Noise, Signal-to-Noise Ratio — Language/Culture — Accent/Dialect — Emotion — Vocal Effort — Lombard Effect — Task Stress — Speaker — Vocabulary, Turn Taking

SOURCE GENERATOR SAMPLE SPACE

STRESSED BASED GENERATOR TRANSFORMATION

NEUTRAL

STRESSED

Variability

within speaker

across speakers
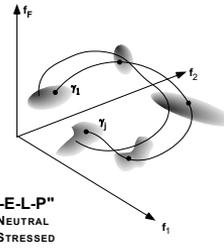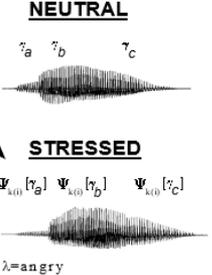
J.H.L. Hansen, "Analysis and Compensation of Speech under Stress and Noise for Environmental Robustness in Speech Recognition," *Speech Communications*, Special Issue on Speech Under Stress, 1996
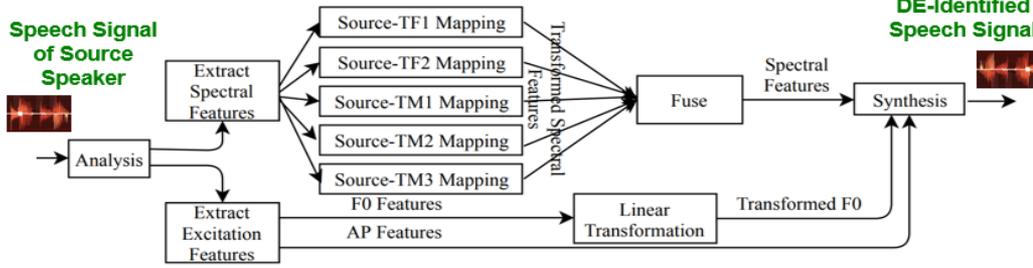G. Zhou, **J.H.L. Hansen**, and J.F. Kaiser, "Nonlinear Feature Based Classification of Speech under Stress," *IEEE Transactions on Speech & Audio Processing*, vol. 9, no. 2, pp. 201-216, 2001.
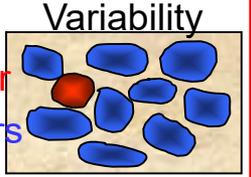
◈ **Recent Forensic Voice Morphing:** protect ID victim/witness

Speaker DE-Identification framework: privacy protection

Speech Signal of Source Speaker → Analysis → Extract Spectral Features → Source-TF1 Mapping / Source-TF2 Mapping / Source-TM1 Mapping / Source-TM2 Mapping / Source-TM3 Mapping → Transformed Spectral Features → Fuse → Spectral Features → Synthesis → **DE-Identified Speech Signal**

Extract Excitation Features → F0 Features / AP Features → Linear Transformation → Transformed F0

Source: Sxx   Target: Txx
New Transformed Source: NSxx

F.Bahmaninezhad, C.Zhang, **J.H.L.Hansen,** "An Investigation of Domain Adaptation in Speaker Embedding Space for Speaker Recog.," Speech Communication, vol. 129, 2021
F. Bahmaninezhad, C. Zhang, **J.H.L. H**ansen, "'Convolutional Neural Network based Voice Conversion for Speaker De-Identification," ISCA Speaker Odyssey-2018, June 26-29, 2018