

Odyssey 2010: The Speaker and Language Recognition Workshop

**Brno, Czech Republic
28 June – 1 July 2010**

ISBN: 978-1-61782-122-6

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2010) by the International Speech Communications Association
All rights reserved.

Printed by Curran Associates, Inc. (2011)

For permission requests, please contact the International Speech Communications Association
at the address below.

International Speech Communications Association
c/o Emmanuelle Foxonet
Lous Tourils
F-66390 Baixas, France

Phone: 33 468 385 827
Fax: 49 228 735 639

secretariat@isca-speech.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

SESSION 1: SPEAKER RECOGNITION – LVCSR AND HIGH LEVEL FEATURES

Constrained Subword Units for Speaker Recognition	1
<i>Doris Baum, Daniel Schneider, Timo Mertens, Joachim Köhler</i>	
Computationally Efficient Speaker Identification for Large Population Tasks using MLLR and Sufficient Statistics	7
<i>Achintya Kumar Sarkar, S. Umesh, Shakti Prasad Rath</i>	
Modeling Prosody for Speaker Recognition: Why Estimating Pitch May Be a Red Herring	12
<i>Kornel Laskowski, Qin Jin</i>	

SESSION 2: FEATURES FOR SPEAKER RECOGNITION

Connectionist Transformation Network Features for Speaker Recognition	20
<i>Alberto Abad, Jordi Luque</i>	
An i-vector Extractor Suitable for Speaker Recognition with both Microphone and Telephone Speech	28
<i>Mohammed Senoussaoui, Patrick Kenny, Najim Dehak, Pierre Dumouchel</i>	
Investigation of Spectral Centroid Magnitude and Frequency for Speaker Recognition	34
<i>Jia Min Karen Kua, Tharmarajah Thiruvaran, Mohaddeseh Nosratioghods, Eliathamby Ambikairajah, Julien Epps</i>	
Temporally Weighted Linear Prediction Features for Speaker Verification in Additive Noise	40
<i>Rahim Saeidi, Jouni Pohjalainen, Tomi Kinnunen, Paavo Alku</i>	

SESSION 3: BACKGROUND MODELING IN SPEAKER RECOGNITION, FORENSICS

Multiple Background Models for Speaker Verification	47
<i>Wei-Qiang Zhang, Yuxiang Shan, Jia Liu</i>	
Training Universal Background Models for Speaker Recognition	52
<i>Mohamed Omar, Jason Pelecanos</i>	
SPES: The BKA Forensic Automatic Voice Comparison System	58
<i>Timo Becker, Michael Jessen, Sebastian Alsbach, Franz Groß, Torsten Meier</i>	
Estimating the Precision of the Likelihood-Ratio Output of a Forensic-Voice-Comparison System	63
<i>Geoffrey Stewart Morrison, Tharmarajah Thiruvaran, Julien Epps</i>	
Investigation of Speaker-Clustered UBMs based on Vocal Tract Lengths and MLLR Matrices for Speaker Verification	71
<i>Achintya Kumar Sarkar, S. Umesh</i>	

SESSION 4: SPEAKER AND LANGUAGE RECOGNITION – SCORING, CONFIDENCES AND CALIBRATION

Cosine Similarity Scoring without Score Normalization Techniques	79
<i>Najim Dehak, Reda Dehak, James Glass, Douglas Reynolds, Patrick Kenny</i>	
Unsupervised Speaker Adaptation based on the Cosine Similarity for Text-Independent Speaker Verification	84
<i>Stephen Shum, Najim Dehak, Reda Dehak, James Glass</i>	
Experiments in SVM-Based Speaker Verification Using Short Utterances	91
<i>Mitchell McLaren, Robbie Vogt, Brendan Baker, Sridha Sridharan</i>	
Detection Target Dependent Score Calibration for Language Recognition	99
<i>Raymond W. M. Ng, Cheung Chi Leung, Tan Lee, Bin Ma, Haizhou Li</i>	

SESSION 5: SPEAKER RECOGNITION – INTER-SESSION VARIABILITY

Weighted Nuisance Attribute Projection	105
<i>William Campbell</i>	
On the Use of Factor Analysis with Restricted Target Data in Speaker Verification	111
<i>Javier Gonzalez-Dominguez, Brendan Baker, Robbie Vogt, Joaquin Gonzalez-Rodriguez, Sridha Sridharan</i>	

Intra-speaker Variability Effects on Speaker Verification Performance	117
<i>Juliette Kahn, Nicolas Audibert, Solange Rossato, Jean-François Bonastre</i>	
Joint Factor Analysis for Speaker Recognition Reinterpreted as Signal Coding Using Overcomplete Dictionaries	125
<i>Daniel Garcia-Romero, Carol Y. Espy-Wilson</i>	

SESSION 6: DIARIZATION

Online Diarization of Telephone Conversations	133
<i>Oshry Ben-Harush, Itshak Lapidot, Hugo Guterman</i>	
Factor Analysis-based Approaches Applied to the Speaker Diarization Task of Meetings: A Preliminary Study	139
<i>Pavel Tomasek, Corinne Fredouille, Driss Matrouf</i>	
Unsupervised Compensation of Intra-Session Intra-Speaker Variability for Speaker Diarization	146
<i>Hagai Aronowitz</i>	
On the Use of GSV-SVM for Speaker Diarization and Tracking	154
<i>Viet Bac Le, Claude Barras, Marc Ferras</i>	

SESSION 7: SPEAKER AND LANGUAGE RECOGNITION - EVALUATIONS AND PERFORMANCE TESTING

Evaluation of the Vulnerability of Speaker Verification to Synthetic Speech	159
<i>Phillip DeLeon, Michael Pucher, Junichi Yamagishi</i>	
Hunting for Wolves in Speaker Recognition	167
<i>Lara Stoll, George Doddington</i>	
The 2009 NIST Language Recognition Evaluation	173
<i>Alvin Martin, Craig Greenberg</i>	
The Albayzin 2008 Language Recognition Evaluation	180
<i>Luis Javier Rodriguez-Fuentes, Mikel Penagarikano, German Bordel, Amparo Varona</i>	

SESSION 8: HUMAN PERFORMANCES IN SPEAKER RECOGNITION, SPEAKER CLUSTERING AND PARTITIONING

Human Assisted Speaker Recognition In NIST SRE10	188
<i>Craig Greenberg, Alvin Martin, Linda Brandschain, Joseph Campbell, Christopher Cieri, George Doddington, John Godfrey</i>	
Speaker Clustering Via the Mean Shift Algorithm	194
<i>Themos Stafylakis, Vassilis Katsouros, George Carayannis</i>	
The Speaker Partitioning Problem	202
<i>Niko Brümmer, Edward de Villiers</i>	
Speaker Linking in Large Data Sets	210
<i>David van Leeuwen</i>	

SESSION 9: LANGUAGE RECOGNITION – GENERAL AND DATA

Estimating and Exploiting Language Distributions of Unlabeled Data	217
<i>Alan McCree</i>	
Data Selection and Calibration Issues in Automatic Language Recognition - Investigation with BUT-AGNITIO NIST LRE 2009 System	223
<i>Zdenek Jancik, Oldrich Plchot, Niko Brümmer, Lukas Burget, Ondrej Glembek, Valiantsina Hubeika, Martin Karafiat, Pavel Matejka, Tomas Mikolov, Albert Strasheim, Jan "Honza" Cernocky</i>	
Comparison of Large-Scale SVM Training Algorithms for Language Recognition	230
<i>Sandro Cumani, Fabio Castaldo, Pietro Laface, Daniele Colibro, Claudio Vair</i>	
Coping with Two Different Transmission Channels in Language Recognition	238
<i>Florian Verdet, Driss Matrouf, Jean-François Bonastre, Jean Hennebert</i>	

SESSION 10: LANGUAGE RECOGNITION – PHONOTACTICS

Improved Modeling of Cross-Decoder Phone Co-occurrences in SVM-Based Phonotactic Language Recognition	246
<i>Mikel Penagarikano, Amparo Varona, Luis Javier Rodriguez-Fuentes, German Bordel</i>	
Parallel Acoustic Model Adaptation for Improving Phonotactic Language Recognition	254
<i>Cheung Chi Leung, Bin Ma, Haizhou Li</i>	
PCA-based Feature Extraction for Phonotactic Language Recognition	259
<i>Tomáš Mikolov, Oldrich Plchot, Ondrej Glembek, Lukáš Burget, Jan “Honza” Cernocký</i>	
Improving Language Recognition with Multilingual Phone Recognition and Speaker Adaptation Transforms	264
<i>Andreas Stolcke, Murat Akbacak, Luciana Ferrer, Sachin Kajarekar, Colleen Richey, Nicolas Scheffer, Elizabeth Shriberg</i>	

SESSION 11: DIALECT RECOGNITION

Discriminative Phonotactics for Dialect Recognition Using Context-Dependent Phone Classifiers	271
<i>Fadi Biadisy, Hagen Soltau, Lidia Mangu, Jiri Navratil, Julia Hirschberg</i>	
Suprasegmental Acoustic Cues of Foreignness in Czech English	279
<i>Jan Volín, Radek Skamitzl</i>	
Exploiting Variety-Dependent Phones in Portuguese Variety Identification	287
<i>Oscar Koller, Alberto Abad, Isabel Trancoso</i>	
Author Index	