

# Shrikanth S Narayanan

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/2899100/publications.pdf>

Version: 2024-02-01

408  
papers

12,437  
citations

66315

42  
h-index

46771

89  
g-index

422  
all docs

422  
docs citations

422  
times ranked

7060  
citing authors

#	ARTICLE	IF	CITATIONS
1	Itâ€™s not what you said, itâ€™s how you said it: An analysis of therapist vocal features during psychotherapy. <i>Counselling and Psychotherapy Research</i> , 2023, 23, 258-269.	1.7	0
2	Multi-Label Multi-Task Deep Learning for Behavioral Coding. <i>IEEE Transactions on Affective Computing</i> , 2022, 13, 508-518.	5.7	17
3	Joint Multi-Dimensional Model for Global and Time-Series Annotations. <i>IEEE Transactions on Affective Computing</i> , 2022, 13, 473-484.	5.7	1
4	Modeling Vocal Entrainment in Conversational Speech Using Deep Unsupervised Learning. <i>IEEE Transactions on Affective Computing</i> , 2022, 13, 1651-1663.	5.7	7
5	Robust Character Labeling in Movie Videos: Data Resources and Self-Supervised Feature Adaptation. <i>IEEE Transactions on Multimedia</i> , 2022, 24, 3355-3368.	5.2	3
6	Automated evaluation of psychotherapy skills using speech and language technologies. <i>Behavior Research Methods</i> , 2022, 54, 690-711.	2.3	28
7	End-to-end neural systems for automatic children speech recognition: An empirical study. <i>Computer Speech and Language</i> , 2022, 72, 101289.	2.9	7
8	Causal indicators for assessing the truthfulness of child speech in forensic interviews. <i>Computer Speech and Language</i> , 2022, 71, 101263.	2.9	2
9	A review of speaker diarization: Recent advances with deep learning. <i>Computer Speech and Language</i> , 2022, 72, 101317.	2.9	90
10	Intra-topic latency as an automated behavioral marker of treatment response in autism spectrum disorder. <i>Scientific Reports</i> , 2022, 12, 3255.	1.6	3
11	Automatic Analysis of Asymmetry in Facial Paralysis Patients Using Landmark-Based Measures. <i>Facial Plastic Surgery and Aesthetic Medicine</i> , 2022, 24, 491-493.	0.5	2
12	Confusion2Vec 2.0: Enriching ambiguous spoken language representations with subwords. <i>PLoS ONE</i> , 2022, 17, e0264488.	1.1	1
13	An automated quality evaluation framework of psychotherapy conversations with local quality estimates. <i>Computer Speech and Language</i> , 2022, 75, 101380.	2.9	4
14	Variation in compensatory strategies as a function of target constriction degree in post-glossectomy speech. <i>JASA Express Letters</i> , 2022, 2, 045205.	0.5	3
15	Representation of professions in entertainment media: Insights into frequency and sentiment trends through computational text analysis. <i>PLoS ONE</i> , 2022, 17, e0267812.	1.1	1
16	Leveraging Social Networks for the Assessment and Management of Neurological Patients. <i>Seminars in Neurology</i> , 2022, 42, 136-148.	0.5	9
17	Conversational correlates of rapid social judgments of children and adolescents with and without ASD. <i>Clinical Linguistics and Phonetics</i> , 2021, 35, 172-184.	0.5	1
18	Meta-Learning With Latent Space Clustering in Generative Adversarial Network for Speaker Diarization. <i>IEEE/ACM Transactions on Audio Speech and Language Processing</i> , 2021, 29, 1204-1219.	4.0	11

#	ARTICLE	IF	CITATIONS
19	Generalized Multiview Shared Subspace Learning Using View Bootstrapping. IEEE Transactions on Signal Processing, 2021, 69, 4774-4786.	3.2	1
20	Multimodal Embeddings From Language Models for Emotion Recognition in the Wild. IEEE Signal Processing Letters, 2021, 28, 608-612.	2.1	15
21	Developing Neural Representations for Robust Child-Adult Diarization. , 2021, , .		2
22	Evidence of Task-Independent Person-Specific Signatures in EEG Using Subspace Techniques. IEEE Transactions on Information Forensics and Security, 2021, 16, 2856-2871.	4.5	8
23	Deep multiple instance learning for foreground speech localization in ambient audio from wearable devices. Eurasip Journal on Audio, Speech, and Music Processing, 2021, 2021, 7.	1.3	7
24	Aliasing artifact reduction in spiral real-time MRI. Magnetic Resonance in Medicine, 2021, 86, 916-925.	1.9	6
25	Health behaviour outcomes of a family based intervention for paediatric obesity in primary care: A randomized type <sc>II</sc> hybrid effectivenessâ€”implementation trial. Pediatric Obesity, 2021, 16, e12780.	1.4	12
26	Intrapersonal and interpersonal vocal affect dynamics during psychotherapy.. Journal of Consulting and Clinical Psychology, 2021, 89, 227-239.	1.6	18
27	A multimodal analysis of physical activity, sleep, and work shift in nurses with wearable sensor data. Scientific Reports, 2021, 11, 8693.	1.6	22
28	A computational lens into how music characterizes genre in film. PLoS ONE, 2021, 16, e0249957.	1.1	9
29	Romantic partner presence and physiological responses in daily life: Attachment style as a moderator. Biological Psychology, 2021, 161, 108082.	1.1	12
30	Computational Media Intelligence: Human-Centered Machine Analysis of Media. Proceedings of the IEEE, 2021, 109, 891-910.	16.4	18
31	Complexity of vocal tract shaping in glossectomy patients and typical speakers: A principal component analysis. Journal of the Acoustical Society of America, 2021, 149, 4437-4449.	0.5	6
32	Multi-Scale Speaker Diarization with Neural Affinity Score Fusion. , 2021, , .		3
33	A Computational Tool to Study Vocal Participation of Women in UN-ITU Meetings. , 2021, , .		0
34	Adversarial Defense for Deep Speaker Recognition Using Hybrid Adversarial Training. , 2021, , .		5
35	Dark tone quality and vocal tract shaping in soprano song production: Insights from real-time MRI. JASA Express Letters, 2021, 1, 075202.	0.5	1
36	Adversarial attack and defense strategies for deep speaker recognition systems. Computer Speech and Language, 2021, 68, 101199.	2.9	35

#	ARTICLE	IF	CITATIONS
37	A multispeaker dataset of raw and reconstructed speech production real-time MRI video and 3D volumetric images. Scientific Data, 2021, 8, 187.	2.4	16
38	Knowledge and Attitudes Toward an Artificial Intelligence-Based Fidelity Measurement in Community Cognitive Behavioral Therapy Supervision. Administration and Policy in Mental Health and Mental Health Services Research, 2021, , 1.	1.2	6
39	Temporal Dynamics of Workplace Acoustic Scenes: Egocentric Analysis and Prediction. IEEE/ACM Transactions on Audio Speech and Language Processing, 2021, 29, 756-769.	4.0	3
40	Improved 3D real-time MRI of speech production. Magnetic Resonance in Medicine, 2021, 85, 3182-3195.	1.9	11
41	Automated quality assessment of cognitive behavioral therapy sessions through highly contextualized language representations. PLoS ONE, 2021, 16, e0258639.	1.1	13
42	Mitigating the Bias of Heterogeneous Human Behavior in Affective Computing. , 2021, , .		3
43	Privacy and Utility Preserving Data Transformation for Speech Emotion Recognition. , 2021, , .		2
44	Coregulation of therapist and client emotion during psychotherapy. Psychotherapy Research, 2020, 30, 591-603.	1.1	33
45	Improved real-time tagged MRI using REALTAG. Magnetic Resonance in Medicine, 2020, 84, 838-846.	1.9	2
46	Speaker Diarization Using Latent Space Clustering in Generative Adversarial Network. , 2020, , .		8
47	Deblurring for spiral real-time MRI using convolutional neural networks. Magnetic Resonance in Medicine, 2020, 84, 3438-3452.	1.9	24
48	TILES-2018, a longitudinal physiologic and behavioral data set of hospital workers. Scientific Data, 2020, 7, 354.	2.4	36
49	Neural Speech Decoding During Audition, Imagination and Production. IEEE Access, 2020, 8, 149714-149729.	2.6	17
50	Modeling Behavioral Consistency in Large-Scale Wearable Recordings of Human Bio-Behavioral Signals. , 2020, , .		5
51	Identifying Truthful Language in Child Interviews. , 2020, , .		2
52	Speaker-Invariant Affective Representation Learning via Adversarial Training. , 2020, , .		24
53	Learning Domain Invariant Representations for Child-Adult Classification from Speech. , 2020, , .		1
54	Bringing in the Outliers: A Sparse Subspace Clustering Approach to Learn a Dictionary of Mouse Ultrasonic Vocalizations. , 2020, , .		1

#	ARTICLE	IF	CITATIONS
55	Meta-Learning for Robust Child-Adult Classification from Speech. , 2020, , .		7
56	The Role of Annotation Fusion Methods in the Study of Human-Reported Emotion Experience During Music Listening. , 2020, , .		1
57	Affect Estimation with Wearable Sensors. Journal of Healthcare Informatics Research, 2020, 4, 261-294.	5.3	6
58	Improving speaker diarization for naturalistic child-adult conversational interactions using contextual information. Journal of the Acoustical Society of America, 2020, 147, EL196-EL200.	0.5	6
59	Modeling Behavior as Mutual Dependency between Physiological Signals and Indoor Location in Large-Scale Wearable Sensor Study. , 2020, , .		5
60	Trapezoidal Segment Sequencing: A Novel Approach for Fusion of Human-Produced Continuous Annotations. , 2020, , .		0
61	Variability in individual constriction contributions to third formant values in American English /É¹/. Journal of the Acoustical Society of America, 2020, 147, 3905-3916.	0.5	3
62	Automatic Prediction of Suicidal Risk in Military Couples Using Multimodal Interaction Cues from Couples Conversations. , 2020, , .		11
63	Auto-Tuning Spectral Clustering for Speaker Diarization Using Normalized Maximum Eigengap. IEEE Signal Processing Letters, 2020, 27, 381-385.	2.1	48
64	Clinical state tracking in serious mental illness through computational analysis of speech. PLoS ONE, 2020, 15, e0225695.	1.1	44
65	Robust Speaker Recognition Using Unsupervised Adversarial Invariance. , 2020, , .		14
66	Vocal tract shaping of emotional speech. Computer Speech and Language, 2020, 64, 101100.	2.9	15
67	Leveraging Linguistic Context in Dyadic Interactions to Improve Automatic Speech Recognition for Children. Computer Speech and Language, 2020, 63, 101101.	2.9	7
68	Vocal Tract Articulatory Contour Detection in Real-Time Magnetic Resonance Images Using Spatio-Temporal Context. , 2020, , .		5
69	Automated Empathy Detection for Oncology Encounters. , 2020, , .		2
70	How an aglossic speaker produces an alveolar-like percept without a functional tongue tip. Journal of the Acoustical Society of America, 2020, 147, EL460-EL464.	0.5	1
71	User-Based Collaborative Filtering Mobile Health System. , 2020, 4, 1-17.		4
72	Dynamic off-axis resonance correction for spiral real-time MRI of speech. Magnetic Resonance in Medicine, 2019, 81, 234-246.	1.9	24

#	ARTICLE	IF	CITATIONS
73	Efficient estimation and model generalization for the total variability model. Computer Speech and Language, 2019, 53, 43-64.	2.9	2
74	Role Specific Lattice Rescoring for Speaker Role Recognition from Speech Recognition Outputs. , 2019, , .		1
75	On Role and Location of Normalization before Model-based Data Augmentation in Residual Blocks for Classification Tasks. , 2019, , .		0
76	On Evaluating CNN Representations for Low Resource Medical Image Classification. , 2019, , .		12
77	Generating labels for regression of subjective constructs using triplet embeddings. Pattern Recognition Letters, 2019, 128, 385-392.	2.6	3
78	A Multimodal View into Music's Effect on Human Neural, Physiological, and Emotional Experience. , 2019, , .		10
79	Discovering Optimal Variable-length Time Series Motifs in Large-scale Wearable Recordings of Human Bio-behavioral Signals. , 2019, , .		7
80	Speaker Agnostic Foreground Speech Detection from Audio Recordings in Workplace Settings from Wearable Recorders. , 2019, , .		9
81	Robust Speech Activity Detection in Movie Audio: Data Resources and Experimental Evaluation. , 2019, , .		10
82	Bluetooth Based Indoor Localization Using Triplet Embeddings. , 2019, , .		5
83	Learning Shared Vector Representations of Lyrics and Chords in Music. , 2019, , .		2
84	Toward Robust Interpretable Human Movement Pattern Analysis in a Workplace Setting. , 2019, , .		8
85	Improving the Prediction of Therapist Behaviors in Addiction Counseling by Exploiting Class Confusions. , 2019, , .		6
86	An Empirical Study of Speech Processing in the Brain by Analyzing the Temporal Syllable Structure in Speech-input Induced EEG. , 2019, , .		7
87	Total Variability Layer in Deep Neural Network Embeddings for Speaker Verification. IEEE Signal Processing Letters, 2019, 26, 893-897.	2.1	4
88	Reinforcing Self-expressive Representation with Constraint Propagation for Face Clustering in Movies. , 2019, , .		7
89	Family-of-origin aggression, dating aggression, and physiological stress reactivity in daily life. Physiology and Behavior, 2019, 206, 85-92.	1.0	13
90	Intermittently tagged real-time MRI reveals internal tongue motion during speech production. Magnetic Resonance in Medicine, 2019, 82, 600-613.	1.9	9

#	ARTICLE	IF	CITATIONS
91	Task-dependence of articulator synergies. Journal of the Acoustical Society of America, 2019, 145, 1504-1520.	0.5	12
92	Estimating Individualized Daily Self-Reported Affect with Wearable Sensors. , 2019, , .		9
93	Trapezoidal Segmented Regression: A Novel Continuous-scale Real-time Annotation Approximation Algorithm. , 2019, , .		1
94	Prediction of Psychological Flexibility with multi-scale Heart Rate Variability and Breathing Features in an "in-the-wild" Setting. , 2019, , .		3
95	A system for the 2019 Sentiment, Emotion and Cognitive State Task of DARPA's LORELEI project. , 2019, , .		1
96	Stress and Anxiety Measurement "In-the-Wild" Using Quality-aware Multi-scale HRV Features. , 2019, 2019, 7056-7059.		5
97	Predicting Human-Reported Enjoyment Responses in Happy and Sad Music. , 2019, , .		4
98	Toward Visual Voice Activity Detection for Unconstrained Videos. , 2019, , .		11
99	Subspace techniques for task-independent EEG person identification. , 2019, 2019, 4545-4548.		3
100	Violence Rating Prediction from Movie Scripts. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 671-678.	3.6	14
101	A modular architecture for articulatory synthesis from gestural specification. Journal of the Acoustical Society of America, 2019, 146, 4458-4471.	0.5	11
102	Cross-Modal Coordination of Face-Directed Gaze and Emotional Speech Production in School-Aged Children and Adolescents with ASD. Scientific Reports, 2019, 9, 18301.	1.6	2
103	Imputing Missing Data In Large-Scale Multivariate Biomedical Wearable Recordings Using Bidirectional Recurrent Neural Networks With Temporal Activation Regularization. , 2019, 2019, 2529-2534.		14
104	Breathing Rate Complexity Features for "In-the-Wild" Stress and Anxiety Measurement. , 2019, , .		6
105	3D dynamic MRI of the vocal tract during natural speech. Magnetic Resonance in Medicine, 2019, 81, 1511-1520.	1.9	26
106	Motion-Capture Patterns of Voluntarily Mimicked Dynamic Facial Expressions in Children and Adolescents With and Without ASD. Journal of Autism and Developmental Disorders, 2019, 49, 1062-1079.	1.7	18
107	Design feasibility of an automated, machine-learning based feedback system for motivational interviewing.. Psychotherapy, 2019, 56, 318-328.	0.7	47
108	Engineering Innovation in Speech Science: Data and Technologies. Perspectives of the ASHA Special Interest Groups, 2019, 4, 411-420.	0.4	6

#	ARTICLE	IF	CITATIONS
109	Multimodal Human and Environmental Sensing for Longitudinal Behavioral Studies in Naturalistic Settings: Framework for Sensor Selection, Deployment, and Management. Journal of Medical Internet Research, 2019, 21, e12832.	2.1	25
110	Lessons Learned: Recommendations For Implementing a Longitudinal Study Using Wearable and Environmental Sensors in a Health Care Organization. JMIR MHealth and UHealth, 2019, 7, e13305.	1.8	21
111	Advances in vocal tract imaging and analysis. , 2019, , 34-50.		7
112	Modeling Multiple Time Series Annotations as Noisy Distortions of the Ground Truth: An Expectation-Maximization Approach. IEEE Transactions on Affective Computing, 2018, 9, 76-89.	5.7	14
113	Acoustic Denoising Using Dictionary Learning With Spectral and Temporal Regularization. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 967-980.	4.0	9
114	Analysis of speech production real-time MRI. Computer Speech and Language, 2018, 52, 1-22.	2.9	36
115	Explaining Coronal Reduction: Prosodic Structure and Articulatory Posture. Phonetica, 2018, 75, 151-181.	0.3	15
116	Unsupervised Discovery of Character Dictionaries in Animation Movies. IEEE Transactions on Multimedia, 2018, 20, 539-551.	5.2	13
117	A Computational Study of Expressive Facial Dynamics in Children with Autism. IEEE Transactions on Affective Computing, 2018, 9, 14-20.	5.7	51
118	The ELISA Situation Frame extraction for low resource languages pipeline for LoReHLT™2016. Machine Translation, 2018, 32, 127-142.	1.3	3
119	Multimodal Interaction Modeling of Child Forensic Interviewing. , 2018, , .		6
120	Role Annotated Speech Recognition for Conversational Interactions. , 2018, , .		1
121	A Novel Method for Human Bias Correction of Continuous- Time Annotations. , 2018, , .		11
122	Discovering Latent Psychological Structures from Self-Report Assessments of Hospital Workers. , 2018, , .		4
123	Multimodal Representation of Advertisements Using Segment-level Autoencoders. , 2018, , .		5
124	Participatory methods to support team science development for predictive analytics in health. Journal of Clinical and Translational Science, 2018, 2, 178-182.	0.3	4
125	An Evaluation of EEG-based Metrics for Engagement Assessment of Distance Learners. , 2018, 2018, 307-310.		8
126	Improving Semi-Supervised Classification for Low-Resource Speech Interaction Applications. , 2018, , .		2



#	ARTICLE	IF	CITATIONS
127	TILES audio recorder. , 2018, , .		23
128	Semi-Supervised and Transfer Learning Approaches for Low Resource Sentiment Classification. , 2018, , .		18
129	Speed-accuracy tradeoffs in human speech production. PLoS ONE, 2018, 13, e0202180.	1.1	12
130	Irregularity-Aware Graph Fourier Transforms. IEEE Transactions on Signal Processing, 2018, 66, 5746-5761.	3.2	42
131	An individually tailored family-centered intervention for pediatric obesity in primary care: study protocol of a randomized type II hybrid effectivenessâ€”implementation trial (Raising Healthy Children) Tj ETQq1 1 257843141gBT /Over	10.5	149
132	A Multi-task Approach to Learning Multilingual Representations. , 2018, , .		11
133	Using Prosodic and Lexical Information for Learning Utterance-level Behaviors in Psychotherapy. , 2018, 2018, 3413-3417.		11
134	Computational Modeling of Conversational Humor in Psychotherapy. , 2018, 2018, 2344-2348.		4
135	Multiple Instance Learning for Behavioral Coding. IEEE Transactions on Affective Computing, 2017, 8, 81-94.	5.7	5
136	A fast and flexible MRI system for the study of dynamic vocal tract shaping. Magnetic Resonance in Medicine, 2017, 77, 112-125.	1.9	53
137	Modeling Dynamics of Expressive Body Gestures In Dyadic Interactions. IEEE Transactions on Affective Computing, 2017, 8, 369-381.	5.7	22
138	Feasibility of throughâ€”time spiral generalized autocalibrating partial parallel acquisition for low latency accelerated realâ€”time MRI of speech. Magnetic Resonance in Medicine, 2017, 78, 2275-2282.	1.9	17
139	MUPETâ€”Mouse Ultrasonic Profile ExTraction: AÂ”Signal Processing Tool for Rapid and Unsupervised Analysis of Ultrasonic Vocalizations. Neuron, 2017, 94, 465-485.e5.	3.8	131
140	New Frontiers in Ambulatory Assessment. Social Psychological and Personality Science, 2017, 8, 552-563.	2.4	30
141	Characterizing Articulation in Apraxic Speech Using Real-Time Magnetic Resonance Imaging. Journal of Speech, Language, and Hearing Research, 2017, 60, 877-891.	0.7	20
142	Signal Processing and Machine Learning for Mental Health Research and Clinical Applications [Perspectives]. IEEE Signal Processing Magazine, 2017, 34, 196-195.	4.6	73
143	Testâ€”retest repeatability of human speech biomarkers from static and real-time dynamic magnetic resonance imaging. Journal of the Acoustical Society of America, 2017, 141, 3323-3336.	0.5	16
144	Designing Contestability. , 2017, 2017, 95-99.		56

#	ARTICLE	IF	CITATIONS
145	A knowledge transfer and boosting approach to the prediction of affect in movies. , 2017, , .		1
146	Multimodal detection of fake social media use through a fusion of classification and pairwise ranking systems. , 2017, , .		12
147	Estimation of vocal tract area function from volumetric Magnetic Resonance Imaging. , 2017, , .		8
148	Weighted geodesic flow kernel for interpersonal mutual influence modeling and emotion recognition in dyadic interactions. , 2017, , .		4
149	The Promise and the Challenge of Technology-Facilitated Methods for Assessing Behavioral and Cognitive Markers of Risk for Suicide among U.S. Army National Guard Personnel. International Journal of Environmental Research and Public Health, 2017, 14, 361.	1.2	5
150	Predicting couple therapy outcomes based on speech acoustic features. PLoS ONE, 2017, 12, e0185123.	1.1	22
151	Linguistic analysis of differences in portrayal of movie characters. , 2017, , .		12
152	Bilabial Substitution Patterns during Consonant Production in a Case of Congenital Aglossia. Journal of Communication Disorders Deaf Studies & Hearing Aids, 2017, 05, .	0.2	5
153	Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. Journal of Child Psychology and Psychiatry and Allied Disciplines, 2016, 57, 927-937.	3.1	140
154	Opening big in box office? Trailer content can help. , 2016, , .		7
155	Comparison of feature-level and kernel-level data fusion methods in multi-sensory fall detection. , 2016, , .		1
156	Speech and language processing for mental health research and care. , 2016, , .		1
157	Computational Analysis and Simulation of Empathic Behaviors: a Survey of Empathy Modeling with Behavioral Signal Processing Framework. Current Psychiatry Reports, 2016, 18, 49.	2.1	27
158	Advances in real-time magnetic resonance imaging of the vocal tract for speech science and technology research. APSIPA Transactions on Signal and Information Processing, 2016, 5, .	2.6	18
159	Markov Chain Monte Carlo Inference of Parametric Dictionaries for Sparse Bayesian Approximations. IEEE Transactions on Signal Processing, 2016, 64, 3077-3092.	3.2	8
160	A multimodal mixture-of-experts model for dynamic emotion prediction in movies. , 2016, , .		18
161	Lightly-supervised utterance-level emotion identification using latent topic modeling of multimodal words. , 2016, , .		2
162	Long-Term SNR Estimation of Speech Signals in Known and Unknown Channel Conditions. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 2495-2506.	4.0	16

#	ARTICLE	IF	CITATIONS
163	Online Affect Tracking with Multimodal Kalman Filters. , 2016, , .		13
164	Comparison of feature-level and kernel-level data fusion methods in multi-sensory fall detection. , 2016, , .		0
165	Strategies for Disseminating Information on Biomedical Research on Autism to Hispanic Parents. Journal of Autism and Developmental Disorders, 2016, 46, 1038-1050.	1.7	13
166	Directly data-derived articulatory gesture-like representations retain discriminatory information about phone categories. Computer Speech and Language, 2016, 36, 330-346.	2.9	7
167	Pathological speech processing: State-of-the-art, current challenges, and future directions. , 2016, , .		22
168	Analysis of engagement behavior in children during dyadic interactions using prosodic cues. Computer Speech and Language, 2016, 37, 47-66.	2.9	15
169	Detecting paralinguistic events in audio stream using context in features and probabilistic decisions. Computer Speech and Language, 2016, 36, 72-92.	2.9	11
170	Speaker verification based on the fusion of speech acoustics and inverted articulatory signals. Computer Speech and Language, 2016, 36, 196-211.	2.9	22
171	The USC CreativeIT database of multimodal dyadic interactions: from speech and full body motion capture to continuous emotional annotations. Language Resources and Evaluation, 2016, 50, 497-521.	1.8	41
172	“It sounds like...” A natural language processing approach to detecting counselor reflections in motivational interviewing.. Journal of Counseling Psychology, 2016, 63, 343-350.	1.4	33
173	Novel affective features for multiscale prediction of emotion in music. , 2016, , .		2
174	A technology prototype system for rating therapist empathy from audio recordings in addiction counseling. PeerJ Computer Science, 2016, 2, e59.	2.7	32
175	Gender Representation in Cinematic Content. , 2015, , .		11
176	"Rate My Therapist": Automated Detection of Empathy in Drug and Alcohol Counseling via Speech and Language Processing. PLoS ONE, 2015, 10, e0143055.	1.1	72
177	Speech in Affective Computing. , 2015, , .		17
178	On Short-Time Estimation of Vocal Tract Length from Formant Frequencies. PLoS ONE, 2015, 10, e0132193.	1.1	44
179	Affect prediction in music using boosted ensemble of filters. , 2015, , .		2
180	Advancing Methods for Reliably Assessing Motivational Interviewing Fidelity using the Motivational Interviewing Skills Code. Journal of Substance Abuse Treatment, 2015, 49, 50-57.	1.5	20

#	ARTICLE	IF	CITATIONS
181	Sparse Representation of Electrodermal Activity With Knowledge-Driven Dictionaries. IEEE Transactions on Biomedical Engineering, 2015, 62, 960-971.	2.5	42
182	Behaviorally-based couple therapies reduce emotional arousal during couple conflict. Behaviour Research and Therapy, 2015, 72, 49-55.	1.6	29
183	A kinematic study of critical and non-critical articulators in emotional speech production. Journal of the Acoustical Society of America, 2015, 137, 1411-1429.	0.5	14
184	Rapid Language Identification. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, 23, 1118-1129.	4.0	17
185	Head Motion Modeling for Human Behavior Analysis in Dyadic Interaction. IEEE Transactions on Multimedia, 2015, 17, 1107-1119.	5.2	36
186	Modeling mutual influence of multimodal behavior in affective dyadic interactions. , 2015, , .		4
187	On quantifying facial expression-related atypicality of children with Autism Spectrum Disorder. , 2015, 2015, 803-807.		24
188	Redundancy analysis of behavioral coding for couples therapy and improved estimation of behavior from noisy annotations. , 2015, , .		7
189	Computationally deconstructing movie narratives: An informatics approach. , 2015, , .		12
190	Automatic intelligibility classification of sentence-level pathological speech. Computer Speech and Language, 2015, 29, 132-144.	2.9	59
191	Applying Machine Learning to Facilitate Autism Diagnostics: Pitfalls and Promises. Journal of Autism and Developmental Disorders, 2015, 45, 1121-1136.	1.7	149
192	Are Articulatory Settings Mechanically Advantageous for Speech Motor Control?. PLoS ONE, 2014, 9, e104168.	1.1	9
193	Integration and Automation of Data Preparation and Data Mining. , 2014, , .		2
194	Learning multiple concepts with incremental diverse density. , 2014, , .		2
195	Energy-constrained minimum variance response filter for robust vowel spectral estimation. , 2014, , .		3
196	Power-spectral analysis of head motion signal for behavioral modeling in human interaction. , 2014, , .		7
197	Semi-supervised term-weighted value rescoring for keyword search. , 2014, , .		6
198	Multimodal Prediction of Affective Dimensions and Depression in Human-Computer Interactions. , 2014, , .		63

#	ARTICLE	IF	CITATIONS
199	Co-registration of speech production datasets from electromagnetic articulography and real-time magnetic resonance imaging. Journal of the Acoustical Society of America, 2014, 135, EL115-EL121.	0.5	12
200	Gesture dynamics modeling for attitude analysis using graph based transform. , 2014, , .		5
201	Graph-based approach for motion capture data representation and analysis. , 2014, , .		6
202	Barista: A framework for concurrent speech processing by usc-sail. , 2014, 2014, 3306-3310.		5
203	Analysis of interaction attitudes using data-driven hand gesture phrases. , 2014, , .		15
204	Developmental acoustic study of American English diphthongs. Journal of the Acoustical Society of America, 2014, 136, 1880-1894.	0.5	13
205	Real-time magnetic resonance imaging and electromagnetic articulography database for speech production research (TC). Journal of the Acoustical Society of America, 2014, 136, 1307-1311.	0.5	120
206	Hull detection based on largest empty sector angle with application to analysis of realtime MR images. , 2014, , .		1
207	Theoretical Analysis of Diversity in an Ensemble of Automatic Speech Recognition Systems. IEEE/ACM Transactions on Audio Speech and Language Processing, 2014, 22, 711-726.	4.0	59
208	Fusion of diverse denoising systems for robust automatic speech recognition. , 2014, , .		3
209	A supervised signal-to-noise ratio estimation of speech signals. , 2014, , .		12
210	Robust Unsupervised Arousal Rating:A Rule-Based Framework withKnowledge-Inspired Vocal Features. IEEE Transactions on Affective Computing, 2014, 5, 201-213.	5.7	52
211	Intoxicated speech detection: A fusion framework with speaker-normalized hierarchical functionals and GMM supervectors. Computer Speech and Language, 2014, 28, 375-391.	2.9	18
212	Simplified supervised i-vector modeling with application to robust and efficient language identification and speaker verification. Computer Speech and Language, 2014, 28, 940-958.	2.9	41
213	Tracking continuous emotional trends of participants during affective dyadic interactions using body language and speech information. Image and Vision Computing, 2013, 31, 137-152.	2.7	87
214	Automatic speaker age and gender recognition using acoustic and prosodic level information fusion. Computer Speech and Language, 2013, 27, 151-167.	2.9	135
215	Dynamic 3-D Visualization of Vocal Tract Shaping During Speech. IEEE Transactions on Medical Imaging, 2013, 32, 838-848.	5.4	31
216	High-quality bilingual subtitle document alignments with application to spontaneous speech translation. Computer Speech and Language, 2013, 27, 572-591.	2.9	1

#	ARTICLE	IF	CITATIONS
217	A Globally-Variant Locally-Constant Model for Fusion of Labels from Multiple Diverse Experts without Using Reference Labels. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35, 769-783.	9.7	26
218	Annotation and processing of continuous emotional attributes: Challenges and opportunities. , 2013, , .		99
219	An audio-visual approach to learning salient behaviors in couples' problem solving discussions. , 2013, , .		5
220	Paralinguistics in speech and languageâ€”State-of-the-art and the challenge. Computer Speech and Language, 2013, 27, 4-39.	2.9	207
221	Enriching machine-mediated speech-to-speech translation using contextual information. Computer Speech and Language, 2013, 27, 492-508.	2.9	2
222	Behavioral Signal Processing: Deriving Human Behavioral Informatics From Speech and Language. Proceedings of the IEEE, 2013, 101, 1203-1233.	16.4	225
223	Toward automating a human behavioral coding system for married couplesâ€™ interactions using speech acoustic features. Speech Communication, 2013, 55, 1-21.	1.6	84
224	Distributional Semantic Models for Affective Text Analysis. IEEE Transactions on Audio Speech and Language Processing, 2013, 21, 2379-2392.	3.8	58
225	Iterative Feature Normalization Scheme for Automatic Emotion Detection from Speech. IEEE Transactions on Affective Computing, 2013, 4, 386-397.	5.7	46
226	Morphological Variation in the Adult Hard Palate and Posterior Pharyngeal Wall. Journal of Speech, Language, and Hearing Research, 2013, 56, 521-530.	0.7	34
227	On smoothing articulatory trajectories obtained from Gaussian mixture model based acoustic-to-articulatory inversion. Journal of the Acoustical Society of America, 2013, 134, EL258-EL264.	0.5	5
228	Interspeaker Variability in Hard Palate Morphology and Vowel Production. Journal of Speech, Language, and Hearing Research, 2013, 56, 1924-1933.	0.7	35
229	Evaluation of Swallow Function After Tongue Cancer Treatment Using Real-Time Magnetic Resonance Imaging. JAMA Otolaryngology - Head and Neck Surgery, 2013, 139, 1312.	1.2	25
230	Speaker verification using simplified and supervised i-vector modeling. , 2013, , .		14
231	A study on the effect of prosodic emphasis transfer on overall speech translation quality. , 2013, , .		6
232	Quantifying atypicality in affective facial expressions of children with autism spectrum disorders. , 2013, 2013, 1-6.		22
233	Using physiology and language cues for modeling verbal response latencies of children with ASD. , 2013, , .		12
234	Data driven modeling of head motion towards analysis of behaviors in couple interactions. , 2013, , .		8

#	ARTICLE	IF	CITATIONS
235	Using emotional noise to uncloud audio-visual emotion perceptual evaluation. , 2013, , .		7
236	A robust frontend for ASR: Combining denoising, noise masking and feature normalization. , 2013, , .		3
237	Combining window predictions efficiently - A new imputation approach for noise robust automatic speech recognition. , 2013, , .		1
238	Toward body language generation in dyadic interaction settings from interlocutor multimodal cues. , 2013, , .		4
239	Head motion synchrony and its correlation to affectivity in dyadic interactions. , 2013, , .		7
240	Spatial and temporal alignment of multimodal human speech production data: Real time imaging, flesh point tracking and audio. , 2013, , .		4
241	Object classification in sidescan sonar images with sparse representation techniques. , 2012, , .		21
242	Speaker states recognition using latent factor analysis based Eigenchannel factor vector modeling. , 2012, , .		13
243	Analyzing the memory of BLSTM Neural Networks for enhanced emotion classification in dyadic spoken interactions. , 2012, , .		12
244	Features for comparing tune similarity of songs across different languages. , 2012, , .		2
245	Automatic recognition of emotion evoked by general sound events. , 2012, , .		27
246	Classification of emotional content of sighs in dyadic human interactions. , 2012, , .		2
247	Improvements in predicting children's overall reading ability by modeling variability in evaluators' subjective judgments. , 2012, , .		3
248	An acoustic analysis of shared enjoyment in ECA interactions of children with autism. , 2012, , .		8
249	Creating ensemble of diverse maximum entropy models. , 2012, , .		3
250	Context-Sensitive Learning for Enhanced Audiovisual Emotion Classification. IEEE Transactions on Affective Computing, 2012, 3, 184-198.	5.7	131
251	Supervised acoustic topic model with a consequent classifier for unstructured audio classification. , 2012, , .		4
252	Improved imaging of lingual articulation using real-time multislice MRI. Journal of Magnetic Resonance Imaging, 2012, 35, 943-948.	1.9	27

#	ARTICLE	IF	CITATIONS
253	On signal representations within the Bayes decision framework. Pattern Recognition, 2012, 45, 1853-1865.	5.1	4
254	Novel Variations of Group Sparse Regularization Techniques With Applications to Noise Robust Automatic Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2012, 20, 1337-1346.	3.8	18
255	EmotiWord: Affective Lexicon Creation with Application to Interaction and Multimedia Data. Lecture Notes in Computer Science, 2012, , 30-41.	1.0	3
256	Analyzing the Language of Therapist Empathy in Motivational Interview based Psychotherapy. , 2012, .		5
257	Directional descriptors using zernike moment phases for object orientation estimation in underwater sonar images. , 2011, , .		8
258	A hierarchical static-dynamic framework for emotion classification. , 2011, , .		17
259	Emotion recognition using a hierarchical binary decision tree approach. Speech Communication, 2011, 53, 1162-1171.	1.6	274
260	A Generative Student Model for Scoring Word Reading Skills. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 348-360.	3.8	18
261	Robust Voice Activity Detection Using Long-Term Signal Variability. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 600-613.	3.8	153
262	Automatic Prediction of Children's Reading Ability for High-Level Literacy Assessment. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 1015-1028.	3.8	36
263	Enhanced Sparse Imputation Techniques for a Robust Speech Recognition Front-End. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 2418-2429.	3.8	19
264	Auditory-like filterbank: An optimal speech processor for efficient human speech communication. Sadhana - Academy Proceedings in Engineering Sciences, 2011, 36, 699-712.	0.8	0
265	Flexible retrospective selection of temporal resolution in real-time speech MRI using a golden-ratio spiral view order. Magnetic Resonance in Medicine, 2011, 65, 1365-1371.	1.9	47
266	Novel 16-channel receive coil array for accelerated upper airway MRI at 3 Tesla. Magnetic Resonance in Medicine, 2011, 65, 1711-1717.	1.9	17
267	Detecting emotional state of a child in a conversational computer game. Computer Speech and Language, 2011, 25, 29-44.	2.9	72
268	Accurate transcription of broadcast news speech using multiple noisy transcribers and unsupervised reliability metrics. , 2011, , .		16
269	Robust talking face video verification using joint factor analysis and sparse representation on GMM mean shifted supervectors. , 2011, , .		9
270	Emotion classification from speech using evaluator reliability-weighted combination of ranked lists. , 2011, , .		6



#	ARTICLE	IF	CITATIONS
271	Iterative feature normalization for emotional speech detection. , 2011, , .		40
272	Overlapped speech detection using long-term spectro-temporal similarity in stereo recording. , 2011, , .		2
273	Estimation of ordinal approach-avoidance labels in dyadic interactions: Ordinal logistic regression approach. , 2011, , .		11
274	Automatically assessing the ABCs. ACM Transactions on Speech and Language Processing, 2011, 7, 1-17.	0.9	3
275	Bilingual audio-subtitle extraction using automatic segmentation of movie audio. , 2011, , .		4
276	Rachel: Design of an emotionally targeted interactive agent for children with autism. , 2011, , .		48
277	Processing speech signal using auditory-like filterbank provides least uncertainty about articulatory gestures. Journal of the Acoustical Society of America, 2011, 129, 4014-4022.	0.5	13
278	Automatic speech recognition using articulatory features from subject-independent acoustic-to-articulatory inversion. Journal of the Acoustical Society of America, 2011, 130, EL251-EL257.	0.5	33
279	“That’s Aggravating, Very Aggravating”: Is It Possible to Classify Behaviors in Couple Interactions Using Automatically Derived Lexical Features?. Lecture Notes in Computer Science, 2011, , 87-96.	1.0	27
280	Optimal Arousal Identification and Classification for Affective Computing Using Physiological Signals: Virtual Reality Stroop Task. IEEE Transactions on Affective Computing, 2010, 1, 109-118.	5.7	120
281	Towards modeling user behavior in interactions mediated through an automated bidirectional speech translation system. Computer Speech and Language, 2010, 24, 232-256.	2.9	5
282	Decision level combination of multiple modalities for recognition and analysis of emotional expression. , 2010, , .		64
283	Predicting interruptions in dyadic spoken interactions. , 2010, , .		16
284	Visual emotion recognition using compact facial representations and viseme information. , 2010, , .		40
285	A near-optimal (minimax) tree-structured partition for mutual information estimation. , 2010, , .		0
286	An exploratory study of manifolds of emotional speech. , 2010, , .		12
287	Language model adaptation using WWW documents obtained by utterance-based queries. , 2010, , .		7
288	Using naïve text queries for robust audio information retrieval. , 2010, , .		1

#	ARTICLE	IF	CITATIONS
289	Robust ECG Biometrics by Fusing Temporal and Cepstral Information. , 2010, , .		68
290	A generalized smoothness criterion for acoustic-to-articulatory inversion. Journal of the Acoustical Society of America, 2010, 128, 2162-2172.	0.5	56
291	Robust representations for out-of-domain emotions using Emotion Profiles. , 2010, , .		3
292	Speech emotion estimation in 3D space. , 2010, , .		53
293	On data-driven histogram-based estimation for mutual information. , 2010, , .		0
294	Bark Frequency Transform Using an Arbitrary Order Allpass Filter. IEEE Signal Processing Letters, 2010, 17, 543-546.	2.1	3
295	An analysis of articulatory-acoustic data based on articulatory strokes. , 2009, , .		5
296	A divide-and-conquer approach to Latent Perceptual Indexing of audio for large Web 2.0 applications. , 2009, , .		1
297	Automatic pronunciation verification of english letter-names for early literacy assessment of preliterate children. , 2009, , .		5
298	A robust harmony structure modeling scheme for classical music opus identification. , 2009, , .		1
299	A semi-supervised learning approach to online audio background detection. , 2009, , .		7
300	Region Segmentation in the Frequency Domain Applied to Upper Airway Real-Time Magnetic Resonance Images. IEEE Transactions on Medical Imaging, 2009, 28, 323-338.	5.4	87
301	Automatic Detection of Disfluency Boundaries in Spontaneous Speech of Children Using Audio&#x2013;Visual Information. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 2-12.	3.8	14
302	Analysis of Emotionally Salient Aspects of Fundamental Frequency for Emotion Detection. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 582-596.	3.8	211
303	Environmental Sound Recognition With Timeâ€“Frequency Audio Features. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 1142-1158.	3.8	475
304	Timing effects of syllable structure and stress on nasals: A real-time MRI examination. Journal of Phonetics, 2009, 37, 97-110.	0.6	64
305	Accelerated threeâ€“dimensional upper airway MRI using compressed sensing. Magnetic Resonance in Medicine, 2009, 61, 1434-1440.	1.9	63
306	Assessment of emerging reading skills in young native speakers and language learners. Speech Communication, 2009, 51, 968-984.	1.6	13

#	ARTICLE	IF	CITATIONS
307	Combining lexical, syntactic and prosodic cues for improved online dialog act tagging. Computer Speech and Language, 2009, 23, 407-422.	2.9	30
308	Lattice-based lexical cues for word fragment detection in conversational speech. , 2009, , .		1
309	Interpreting ambiguous emotional expressions. , 2009, , .		89
310	Histogram-based estimation for the divergence revisited. , 2009, , .		3
311	Pitch Contour Stylization Using an Optimal Piecewise Polynomial Approximation. IEEE Signal Processing Letters, 2009, 16, 810-813.	2.1	11
312	Accelerated 3D MRI of vocal tract shaping using compressed sensing and parallel imaging. , 2009, , .		5
313	Saliency-driven unstructured acoustic scene classification using latent perceptual indexing. , 2009, , .		11
314	IEMOCAP: interactive emotional dyadic motion capture database. Language Resources and Evaluation, 2008, 42, 335-359.	1.8	1,756
315	Strategies to Improve the Robustness of Agglomerative Hierarchical Clustering Under Data Source Variation for Speaker Diarization. IEEE Transactions on Audio Speech and Language Processing, 2008, 16, 1590-1601.	3.8	43
316	Seeing speech: Capturing vocal tract shaping using real-time magnetic resonance imaging [Exploratory DSP]. IEEE Signal Processing Magazine, 2008, 25, 123-132.	4.6	82
317	Environmental sound recognition using MP-based features. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	44
318	The Vera am Mittag German audio-visual emotional speech database. , 2008, , .		246
319	Audio-Visual Emotion Recognition Using Gaussian Mixture Models for Face and Voice. , 2008, , .		63
320	Selection of Emotionally Salient Audio-Visual Features for Modeling Human Evaluations of Synthetic Character Emotion Displays. , 2008, , .		2
321	The SAIL speaker diarization system for analysis of spontaneous meetings. , 2008, , .		5
322	Incorporating discourse context in spoken language translation through dialog acts. , 2008, , .		0
323	Audio retrieval by latent perceptual indexing. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	25
324	Novel inter-cluster distance measure combining GLR and ICR for improved agglomerative hierarchical speaker clustering. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	8

#	ARTICLE	IF	CITATIONS
325	Recognition for synthesis: Automatic parameter selection for resynthesis of emotional speech from neutral speech. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	5
326	Joint-processing of audio-visual signals in human perception of conflicting synthetic character emotions. , 2008, , .		6
327	Minimum Probability of Error Signal Representation. IEEE International Workshop on Machine Learning for Signal Processing, 2007, , .	0.0	5
328	Discriminating Two Types of Noise Sources using Cortical Representation and Dimension Reduction Technique. , 2007, , .		3
329	Optimal Wavelet Packets Decomposition Based on a Rate-Distortion Optimality Criterion. , 2007, , .		0
330	Analysis of Audio Clustering using Word Descriptions. , 2007, , .		7
331	Universal Consistency of Data-Driven Partitions for Divergence Estimation. , 2007, , .		12
332	Real-time Emotion Detection System using Speech: Multi-modal Fusion of Different Timescale Features. , 2007, , .		61
333	Data Driven Approach for Language Model Adaptation using Stepwise Relative Entropy Minimization. , 2007, , .		7
334	Real-Time Monitoring of Participants' Interaction in a Meeting using Audio-Visual Sensors. , 2007, , .		18
335	Experiments in Automatic Genre Classification of Full-length Music Tracks using Audio Activity Rate. , 2007, , .		3
336	Improved Speech Recognition using Acoustic and Lexical Correlates of Pitch Accent in a N-Best Rescoring Framework. , 2007, , .		19
337	Investigating Implicit Cues for User State Estimation in Human-Robot Interaction Using Physiological Measurements. , 2007, , .		30
338	A Statistical Approach for Modeling Prosody Features using POS Tags for Emotional Speech Synthesis. , 2007, , .		16
339	Robust speaker clustering strategies to data source variation for improved speaker diarization. , 2007, , .		9
340	Statistical Modeling and Retrieval of Polyphonic Music. , 2007, , .		8
341	Analyzing the Multimodal Behaviors of Users of a Speech-to-Speech Translation Device by using Concept Matching Scores. , 2007, , .		0
342	Multimodal Meeting Monitoring: Improvements on Speaker Tracking and Segmentation through a Modified Mixture Particle Filter. , 2007, , .		4

#	ARTICLE	IF	CITATIONS
343	Support Vector Regression for Automatic Recognition of Spontaneous Emotions in Speech. , 2007, , .		68
344	Joint Analysis of the Emotional Fingerprint in the Face and Speech: A single subject study. , 2007, , .		14
345	A System for Technology Based Assessment of Language and Literacy in Young Children: the Role of Multiple Information Sources. , 2007, , .		26
346	A review of the acoustic and linguistic properties of children's speech. , 2007, , .		15
347	Primitives-based evaluation and estimation of emotions in speech. Speech Communication, 2007, 49, 787-800.	1.6	308
348	Rigid Head Motion in Expressive Speech Animation: Analysis and Synthesis. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 1075-1086.	3.8	147
349	Interrelation Between Speech and Facial Gestures in Emotional Utterances: A Single Subject Study. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 2331-2347.	3.8	101
350	Pathological Voice Assessment. , 2006, 2006, 1669-73.		45
351	ACOUSTIC-SYNTACTIC MAXIMUM ENTROPY MODEL FOR AUTOMATIC PROSODY LABELING. , 2006, , .		4
352	An attribute-based approach to audio description applied to segmenting vocal sections in popular music songs. , 2006, , .		2
353	Efficient scalable encoding for distributed speech recognition. Speech Communication, 2006, 48, 888-902.	1.6	16
354	Synchronized and noise-robust audio recordings during realtime magnetic resonance imaging scans. Journal of the Acoustical Society of America, 2006, 120, 1791-1794.	0.5	104
355	Upper Bound Kullback-Leibler Divergence for Hidden Markov Models with Application as Discrimination Measure for Speech Recognition. , 2006, , .		7
356	An approach to real-time magnetic resonance imaging for speech production. Journal of the Acoustical Society of America, 2004, 115, 1771-1776.	0.5	256
357	Adaptive speaker identification with audiovisual cues for movie content analysis. Pattern Recognition Letters, 2004, 25, 777-791.	2.6	15
358	Movie Content Analysis, Indexing and Skimming Via Multimodal Information. , 2003, , 123-154.		5
359	Identification of speakers in movie dialogs using audiovisual cues. , 2002, , .		4
360	Acoustics of children's speech: Developmental changes of temporal and spectral parameters. Journal of the Acoustical Society of America, 1999, 105, 1455-1468.	0.5	609

#	ARTICLE	IF	CITATIONS
361	Geometry, kinematics, and acoustics of Tamil liquid consonants. Journal of the Acoustical Society of America, 1999, 106, 1993-2007.	0.5	54
362	Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part II. The rhotics. Journal of the Acoustical Society of America, 1997, 101, 1078-1089.	0.5	100
363	Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part I. The laterals. Journal of the Acoustical Society of America, 1997, 101, 1064-1077.	0.5	162
364	Acoustic analysis and automatic recognition of spontaneous children's speech. , 0, , .		5
365	A Bayesian network classifier for word-level reading assessment. , 0, , .		13
366	An investigation of vocal arousal dynamics in child-psychologist interactions using synchrony measures and a conversation-based model. , 0, , .		6
367	Perceptual Lateralization of Coda Rhotic Production in Puerto Rican Spanish. , 0, , .		2
368	Behavioral Coding of Therapist Language in Addiction Counseling Using Recurrent Neural Networks. , 0, , .		24
369	An Expectation Maximization Approach to Joint Modeling of Multidimensional Ratings Derived from Multiple Annotators. , 0, , .		9
370	A Deep Learning Approach to Modeling Empathy in Addiction Counseling. , 0, , .		24
371	State-of-the-Art MRI Protocol for Comprehensive Assessment of Vocal Tract Structure and Function. , 0, , .		15
372	Convex Hull Convolutional Non-Negative Matrix Factorization for Uncovering Temporal Patterns in Multivariate Time-Series Data. , 0, , .		9
373	L2 Acquisition and Production of the English Rhotic Pharyngeal Gesture. , 0, , .		4
374	Improved Depiction of Tissue Boundaries in Vocal Tract Real-Time MRI Using Automatic Off-Resonance Correction. , 0, , .		3
375	An Affect Prediction Approach Through Depression Severity Parameter Incorporation in Neural Networks. , 0, , .		8
376	Semantic Edge Detection for Tracking Vocal Tract Air-Tissue Boundaries in Real-Time Magnetic Resonance Images. , 0, , .		20
377	Extracting Situation Frames from Non-English Speech: Evaluation Framework and Pilot Results. , 0, , .		4
378	Multi-Scale Context Adaptation for Improving Child Automatic Speech Recognition in Child-Adult Spoken Interactions. , 0, , .		5

#	ARTICLE	IF	CITATIONS
379	Database of Volumetric and Real-Time Vocal Tract MRI for Speech Science. , 0, , .		14
380	Improving Gender Identification in Movie Audio Using Cross-Domain Data. , 0, , .		10
381	Language Features for Automated Evaluation of Cognitive Behavior Psychotherapy Sessions. , 0, , .		14
382	Combined Speaker Clustering and Role Recognition in Conversational Speech. , 0, , .		6
383	Modeling Interpersonal Linguistic Coordination in Conversations Using Word Moverâ€™s Distance. , 0, , .		8
384	Data Augmentation Using GANs for Speech Emotion Recognition. , 0, , .		73
385	Identifying Therapist and Client Personae for Therapeutic Alliance Estimation. , 0, , .		5
386	Multi-Task Discriminative Training of Hybrid DNN-TVM Model for Speaker Verification with Noisy and Far-Field Speech. , 0, , .		7
387	Pronunciation variations of Spanish-accented English spoken by young children. , 0, , .		15
388	TBALL data collection: the making of a young children's speech corpus. , 0, , .		23
389	Exploiting prosodic features for dialog act tagging in a discriminative modeling framework. , 0, , .		10
390	Factored translation models for enriching spoken language translation with prosody. , 0, , .		6
391	Pronunciation verification of English letter-sounds in preliterate children. , 0, , .		6
392	Emotion recognition using a hierarchical binary decision tree approach. , 0, , .		37
393	Modeling therapist empathy and vocal entrainment in drug addiction counseling. , 0, , .		24
394	Modeling therapist empathy through prosody in drug addiction counseling. , 0, , .		26
395	Predicting therapist empathy in motivational interviews using language features inspired by psycholinguistic norms. , 0, , .		25
396	Analyzing speech rate entrainment and its relation to therapist empathy in drug addiction counseling. , 0, , .		23

#	ARTICLE	IF	CITATIONS
397	A dialog act tagging approach to behavioral coding: a case study of addiction counseling conversations. , 0, , .		18
398	Test-Retest Repeatability of Articulatory Strategies Using Real-Time Magnetic Resonance Imaging. , 0, , .		0
399	Exploiting Intra-Annotator Rating Consistency Through Copelandâ€™s Method for Estimation of Ground Truth Labels in Couplesâ€™ Therapy. , 0, , .		0
400	A Distribution Free Formulation of the Total Variability Model. , 0, , .		2
401	A Knowledge Driven Structural Segmentation Approach for Play-Talk Classification During Autism Assessment. , 0, , .		2
402	Exploring the Relationship between Conic Affinity of NMF Dictionaries and Speech Enhancement Metrics. , 0, , .		1
403	Towards an Unsupervised Entrainment Distance in Conversational Speech Using Deep Neural Networks. , 0, , .		5
404	Stochastic Shake-Shake Regularization for Affective Learning from Speech. , 0, , .		2
405	Denoising and Raw-waveform Networks for Weakly-Supervised Gender Identification on Noisy Speech. , 0, , .		1
406	Speaker Diarization with Lexical Information. , 0, , .		13
407	The Second DIHARD Challenge: System Description for USC-SAIL Team. , 0, , .		6
408	Multiview Shared Subspace Learning Across Speakers and Speech Commands. , 0, , .		4