

Publications by Björn W. Schuller

26 April 2017

Current h-index: **56** (source: Google Scholar)

Current citation count: **14052** (source: Google Scholar)

Citations are only provided for papers with citation count ≥ 100 .

(IF): Journal Impact Factor according to Journal Citation Reports, Thomson Reuters.

(IF*): Conference Impact Factor according to <http://citescholar.org>.

Acceptance rates and impact factors of satellite workshops may be subsumed with the main conference.

A) BOOKS

Books Authored (6):

- 1) Schuller, B. W. (2018). *I Know What You're Thinking: The Making of Emotional Machines*. Princeton University Press. to appear
- 2) Balahur-Dobrescu, A., Taboada, M., and Schuller, B. W. (2017). *Computational Methods for Affect Detection from Natural Language*. Computational Social Sciences. Springer. to appear
- 3) Schuller, B. (2013b). *Intelligent Audio Analysis*. Signals and Communication Technology. Springer. 350 pages
- 4) Schuller, B. and Batliner, A. (2013). *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing*. Wiley
- 5) Kroschel, K., Rigoll, G., and Schuller, B. (2011). *Statistische Informationstechnik*. Springer, Berlin/Heidelberg, 5th edition
- 6) Schuller, B. (2007). *Mensch, Maschine, Emotion – Erkennung aus sprachlicher und manueller Interaktion*. VDM Verlag Dr Müller, Saarbrücken. 239 pages

Books Edited (2):

- 7) Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., and Krüger, A., editors (2017). *Handbook of Multimodal-Multisensor Interfaces (3 volumes)*. ACM Books, Morgan & Claypool. to appear
- 8) D'Mello, S., Graesser, A., Schuller, B., and Martin, J.-C., editors (2011). *Proceedings of the 4th International HUMAINE Association Conference on Affective Computing and Intelligent Interaction 2011, ACII 2011*, volume 6974/6975, Part I / Part II of *Lecture Notes on Computer Science (LNCS)*, Memphis, TN. HUMAINE Association, Springer

Contributions to Books (42):

- 9) Schuller, B., Elkins, A., and Scherer, K. (2017b). Computational Analysis of Vocal Expression of Affect: Trends and Challenges. In Burgoon, J., Magnenat-Thalmann, N., Pantic, M., and Vinciarelli, A., editors, *Social Signal Processing*, chapter 6, pages 56–68. Cambridge University Press
- 10) Schuller, B. (2017d). Multimodal User State & Trait Recognition: An Overview. In Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., and Krüger, A., editors, *Handbook of Multimodal-Multisensor Interfaces*. ACM Books, Morgan & Claypool. 30 pages, to appear
- 11) Schuller, B., Bengio, S., and Morency, L.-P. (2017a). Perspectives on Predictive Power of Multimodal Deep Learning: Surprises and Future Directions. In Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., and Krüger, A., editors, *Handbook of Multimodal-Multisensor Interfaces*. ACM Books, Morgan & Claypool. 15 pages, to appear
- 12) Schuller, B., Ernst, M., Ross, A., and Ramos, F. (2017c). Perspectives on Strategic Fusion: What Can Neuroscience, Engineering, and Applications Experience Teach Us about Optimizing Multimodal Multi-sensor System Reliability? In Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., and Krüger, A., editors, *Handbook of Multimodal-Multisensor Interfaces*. ACM Books, Morgan & Claypool. 15 pages, to appear
- 13) Bengio, S., Deng, L., Morency, L.-P., and Schuller, B. (2017). Multidisciplinary Challenge Topic: Perspectives on Predictive Power of Multimodal Deep Learning: Surprises and Future Directions. In Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., and Krüger, A., editors, *Handbook of Multimodal-Multisensor Interfaces*. ACM Books, Morgan & Claypool. 30 pages, to appear
- 14) Ernst, M., Ramos, F., Ross, A., and Schuller, B. (2017). Multidisciplinary Challenge Topic: Perspectives on Strategic Fusion: What Can Neuroscience, Engineering, and Applications Experience Teach Us about Optimizing Multimodal Multi-sensor System Reliability? In Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., and Krüger, A., editors, *Handbook of Multimodal-Multisensor Interfaces*. ACM Books, Morgan & Claypool. to appear
- 15) Gunes, H. and Schuller, B. (2017a). Automatic Analysis of Aesthetics: Human Beauty, Attractiveness, and Likability. In Burgoon, J., Magnenat-Thalmann, N., Pantic, M., and Vinciarelli, A., editors, *Social Signal Processing*, chapter 14, pages 183–201. Cambridge University Press
- 16) Gunes, H. and Schuller, B. (2017b). Automatic Analysis of Social Emotions. In Burgoon, J., Magnenat-Thalmann, N., Pantic, M., and Vinciarelli, A., editors, *Social Signal Processing*, chapter 16, pages 213–224. Cambridge University Press
- 17) Hantke, S., Amiriparian, S., Schmitt, M., Qian, K., Guo, J., Matsuoka, S., and Schuller, B. (2017). Big Data Multimedia Mining: Equipped for Victory facing Volume, Velocity, Variety, Veracity, and Value. In Vrochidis, S., Huet, B., Chang, E., and Kompatsiaris, I., editors, *Big Data Analytics for Large-Scale Multimedia Search*. Wiley. to appear
- 18) Schuller, B. (2016b). Acquisition of affect. In Tkalcic, M., De Carolis, B., de Gemmis, M., Odić, A., and Kosir, A., editors, *Emotions and Personality in Personalized Services*, HumanComputer Interaction Series, pages 57–80. Springer, 1st edition
- 19) Burkhardt, F., Pelachaud, C., Schuller, B., and Zovato, E. (2017). Emotion Markup Language. In Dahl, D., editor, *Multimodal Interaction with W3C Standards: Towards Natural User Interfaces to Everything*, pages 65–80. Springer, Berlin/Heidelberg
- 20) Keren, G., Mousa, A. E.-D., Pietquin, O., Zafeiriou, S., and Schuller, B. (2017b). Deep Learning for Multisensorial and Multimodal Interaction. In Oviatt, S., Schuller, B., Cohen, P., Sonntag, D., Potamianos, G., and Krüger, A., editors, *Handbook of Multimodal-Multisensor Interfaces*. ACM Books, Morgan & Claypool. 30 pages, to appear
- 21) Schmitt, M. and Schuller, B. (2015). Machine-based decoding of paralinguistic vocal features. In Frühholz, S. and Belin, P., editors, *The Oxford Handbook of Voice Perception*, chapter 40. Oxford University Press. invited contribution, to appear
- 22) Schuller, B. (2015b). Deep Learning our Everyday Emotions – A Short Overview. In Bassis, S., Esposito, A., and Morabito, F. C., editors, *Advances in Neural Networks: Computational and Theoretical Issues Emotional Expressions and Daily Cognitive Functions*, volume 37 of *Smart Innovation Systems and Technologies*, pages 339–346. Springer, Berlin Heidelberg. invited contribution
- 23) Schuller, B. and Weninger, F. (2015). Human Affect Recognition – Audio-Based Methods. In Webster, J. G., editor, *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–13. John Wiley & Sons, New York. invited contribution
- 24) Schuller, B. (2015g). Multimodal Affect Databases - Collection, Challenges & Chances. In Calvo, R. A., D'Mello, S., Gratch, J., and Kappas, A., editors, *Handbook of Affective Computing*, Oxford Library of Psychology, chapter 23, pages 323–333.

- Oxford University Press. invited contribution
- 25) Schuller, B. (2015d). Emotion Modeling via Speech Content and Prosody – in Computer Games and Elsewhere. In Yannakakis, G. and Karpouzis, K., editors, *Emotion in Games – Theory and Practice*, volume 4 of *Socio-Affective Computing*, chapter 5, pages 85–102. Springer. invited contribution
 - 26) Batliner, A., Steidl, S., Eyben, F., and Schuller, B. (2015). On Laughter and Speech Laugh, Based on Observations of Child-Robot Interaction. In Trouvain, J. and Campbell, N., editors, *Phonetics of Laughing*. universaar – Saarland University Press, Saarbrücken, Germany. 29 pages, to appear
 - 27) Brückner, R. and Schuller, B. (2015). Being at Odds? – Deep and Hierarchical Neural Networks for Classification and Regression of Conflict in Speech. In D’Errico, F., Poggi, I., Vinciarelli, A., and Vincze, L., editors, *Conflict and Multimodal Communication – Social research and machine intelligence*, Computational Social Sciences, pages 403–429. Springer, Berlin/Heidelberg. invited contribution
 - 28) Castellano, G., Gunes, H., Peters, C., and Schuller, B. (2015). Multimodal Affect Detection for Naturalistic Human-Computer and Human-Robot Interactions. In Calvo, R. A., D’Mello, S., Gratch, J., and Kappas, A., editors, *Handbook of Affective Computing*, Oxford Library of Psychology, chapter 17, pages 246–260. Oxford University Press. invited contribution
 - 29) Marchi, E., Zhang, Y., Eyben, F., Ringeval, F., and Schuller, B. (2015e). Autism and Speech, Language, and Emotion - a Survey. In Patil, H. and Kulshreshtha, M., editors, *Evaluating the Role of Speech Technology in Medical Case Management*. De Gruyter, Berlin. invited contribution, to appear
 - 30) Weninger, F., Wöllmer, M., and Schuller, B. (2015c). Emotion Recognition in Naturalistic Speech and Language - A Survey. In Konar, A. and Chakraborty, A., editors, *Emotion Recognition: A Pattern Analysis Approach*, chapter 10, pages 237–267. Wiley, 1st edition
 - 31) Marchi, E., Ringeval, F., and Schuller, B. (2014c). Voice-enabled assistive robots for handling autism spectrum conditions: an examination of the role of prosody. In Neustein, A., editor, *Speech and Automata in Health Care (Speech Technology and Text Mining in Medicine and Healthcare)*, pages 207–236. De Gruyter, Boston/Berlin/Munich. invited contribution
 - 32) Schuller, B. (2013d). Prosody and Phonemes: On the Influence of Speaking Style. In Hancil, S. and Hirst, D., editors, *Prosody and Iconicity*, chapter 13, pages 233–250. Benjamins
 - 33) Schuller, B. and Weninger, F. (2012). Ten Recent Trends in Computational Paralinguistics. In Esposito, A., Vinciarelli, A., Hoffmann, R., and Müller, V. C., editors, *4th COST 2102 International Training School on Cognitive Behavioural Systems*, volume 7403/2012 of *Lecture Notes on Computer Science (LNCS)* 7403, pages 35–49. Springer, Berlin Heidelberg
 - 34) Rotili, R., Principi, E., Wöllmer, M., Squartini, S., and Schuller, B. (2012c). Conversational Speech Recognition In Non-Stationary Reverberated Environments. In Esposito, A., Vinciarelli, A., Hoffmann, R., and Müller, V. C., editors, *4th COST 2102 International Training School on Cognitive Behavioural Systems*, volume 7403/2012 of *Lecture Notes on Computer Science (LNCS)*, pages 50–59. Springer, Berlin Heidelberg
 - 35) Rotili, R., Principi, E., Squartini, S., and Schuller, B. (2012b). Real-Time Speech Recognition in a Multi-Talker Reverberated Acoustic Scenario. In *Advanced Intelligent Computing Theories and Applications. With Aspects of Artificial Intelligence. Proc. Seventh International Conference on Intelligent Computing (ICIC 2011)*, volume 6839 of *Lecture Notes on Computer Science (LNCS)*, pages 379–386. Springer
 - 36) Schuller, B. (2011c). Voice and Speech Analysis in Search of States and Traits. In Salah, A. A. and Gevers, T., editors, *Computer Analysis of Human Behavior*, Advances in Pattern Recognition, chapter 9, pages 227–253. Springer
 - 37) Schuller, B. and Knaup, T. (2011). Learning and Knowledge-based Sentiment Analysis in Movie Review Key Excerpts. In Esposito, A., Esposito, A. M., Martone, R., Müller, V., and Scarpetta, G., editors, *Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces: Theoretical and Practical Issues: Third COST 2102 International Training School, Caserta, Italy, March 15-19, 2010, Revised Selected Papers*, volume 6456/2010 of *Lecture Notes on Computer Science (LNCS)*, pages 448–472. Springer, Heidelberg, 1st edition
 - 38) Schuller, B., Wöllmer, M., Eyben, F., and Rigoll, G. (2011g). Retrieval of Paralinguistic Information in Broadcasts. In Maybury, M. T., editor, *Multimedia Information Extraction: Advances in video, audio, and imagery extraction for search, data mining, surveillance, and authoring*, chapter 17, pages 273–288. Wiley, IEEE Computer Society Press
 - 39) Arsić, D. and Schuller, B. (2011). Real Time Person Tracking and Behavior Interpretation in Multi Camera Scenarios Applying Homography and Coupled HMMs. In Esposito, A., Vinciarelli, A., Vicsi, K., Pelachaud, C., and Nijholt, A., editors, *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issues, COST 2102 International Conference, Budapest, Hungary, September 7-10, 2010, Revised Selected Papers*, volume 6800/2011 of *Lecture Notes on Computer Science (LNCS)*, pages 1–18. Springer, Heidelberg
 - 40) Schuller, B., Dibiasi, F., Eyben, F., and Rigoll, G. (2010a). Music Thumbnailing Incorporating Harmony- and Rhythm Structure. In Detyniecki, M., Leiner, U., and Nürnberger, A., editors, *Adaptive Multimedia Retrieval: 6th International Workshop, AMR 2008, Berlin, Germany, June 26-27, 2008. Revised Selected Papers*, volume 5811/2010 of *Lecture Notes in Computer Science (LNCS)*, pages 78–88. Springer, Berlin/Heidelberg
 - 41) Batliner, A., Schuller, B., Seppi, D., Steidl, S., Devillers, L., Vidrascu, L., Vogt, T., Aharonson, V., and Amir, N. (2010a). The Automatic Recognition of Emotions in Speech. In Cowie, R., Petta, P., and Pelachaud, C., editors, *Emotion-Oriented Systems: The HUMAINE Handbook*, Cognitive Technologies, pages 71–99. Springer, 1st edition
 - 42) Wöllmer, M., Eyben, F., Graves, A., Schuller, B., and Rigoll, G. (2010b). Improving Keyword Spotting with a Tandem BLSTM-DBN Architecture. In Sole-Casals, J. and Zaiats, V., editors, *Advances in Non-Linear Speech Processing: International Conference on Nonlinear Speech Processing, NOLISP 2009, Vic, Spain, June 25-27, 2009, Revised Selected Papers*, volume 5933/2010 of *Lecture Notes on Computer Science (LNCS)*, pages 68–75. Springer
 - 43) Schuller, B., Wöllmer, M., Eyben, F., and Rigoll, G. (2009k). Spectral or Voice Quality? Feature Type Relevance for the Discrimination of Emotion Pairs. In Hancil, S., editor, *The Role of Prosody in Affective Speech*, volume 97 of *Linguistic Insights, Studies in Language and Communication*, pages 285–307. Peter Lang Publishing Group
 - 44) Schuller, B., Eyben, F., and Rigoll, G. (2008e). Static and Dynamic Modelling for the Recognition of Non-Verbal Vocalisations in Conversational Speech. In André, E., Dybkjaer, L., Neumann, H., Pieraccini, R., and Weber, M., editors, *Perception in Multimodal Dialogue Systems: 4th IEEE Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems, PIT 2008, Kloster Irsee, Germany, June 16-18, 2008, Proceedings*, volume 5078/2008 of *Lecture Notes on Computer Science (LNCS)*, pages 99–110. Springer, Berlin/Heidelberg
 - 45) Schuller, B., Wöllmer, M., Moosmayr, T., Ruske, G., and Rigoll, G. (2008m). Switching Linear Dynamic Models for Noise Robust In-Car Speech Recognition. In Rigoll, G., editor, *Pattern Recognition: 30th DAGM Symposium Munich, Germany, June 10-13, 2008 Proceedings*, volume 5096 of *Lecture Notes on Computer Science (LNCS)*, pages 244–253. Springer, Berlin/Heidelberg. (acceptance rate: 39%, IF* 1.13 (2010))
 - 46) Vlasenko, B., Schuller, B., Wendemuth, A., and Rigoll, G. (2008b). On the Influence of Phonetic Content Variation for Acoustic Emotion Recognition. In André, E., Dybkjaer, L.,

- Neumann, H., Pieraccini, R., and Weber, M., editors, *Perception in Multimodal Dialogue Systems: 4th IEEE Tutorial and Research Workshop on Perception and Interactive Technologies for Speech-Based Systems, PIT 2008, Kloster Irsee, Germany, June 16-18, 2008, Proceedings*, volume 5078/2008 of *Lecture Notes on Computer Science (LNCS)*, pages 217–220. Springer, Berlin/Heidelberg
- 47) Grimm, M., Kroschel, K., Harris, H., Nass, C., Schuller, B., Rigoll, G., and Moosmayr, T. (2007a). On the Necessity and Feasibility of Detecting a Driver’s Emotional State While Driving. In Paiva, A., Picard, R. W., and Prada, R., editors, *Affective Computing and Intelligent Interaction: Second International Conference, ACII 2007, Lisbon, Portugal, September 12-14, 2007, Proceedings*, volume 4738/2007 of *Lecture Notes on Computer Science (LNCS)*, pages 126–138. Springer, Berlin/Heidelberg
- 48) Vlasenko, B., Schuller, B., Wendemuth, A., and Rigoll, G. (2007b). Frame vs. Turn-Level: Emotion Recognition from Speech Considering Static and Dynamic Processing. In Paiva, A., Picard, R. W., and Prada, R., editors, *Affective Computing and Intelligent Interaction: Second International Conference, ACII 2007, Lisbon, Portugal, September 12-14, 2007, Proceedings*, volume 4738/2007 of *Lecture Notes on Computer Science (LNCS)*, pages 139–147. Springer, Berlin/Heidelberg
- 49) Schröder, M., Devillers, L., Karpouzis, K., Martin, J.-C., Pelachaud, C., Peter, C., Pirker, H., Schuller, B., Tao, J., and Wilson, I. (2007). What should a generic emotion markup language be able to represent? In Paiva, A., Picard, R. W., and Prada, R., editors, *Affective Computing and Intelligent Interaction: Second International Conference, ACII 2007, Lisbon, Portugal, September 12-14, 2007, Proceedings*, volume 4738/2007 of *Lecture Notes on Computer Science (LNCS)*, pages 440–451. Springer, Berlin/Heidelberg
- 50) Schuller, B., Ablameier, M., Müller, R., Reifinger, S., Poitschke, T., and Rigoll, G. (2006a). Speech Communication and Multimodal Interfaces. In Kraiss, K.-F., editor, *Advanced Man Machine Interaction*, Signals and Communication Technology, chapter 4, pages 141–190. Springer, Berlin/Heidelberg
- B) REFEREED JOURNAL PAPERS (94)
- 51) Schuller, B. (2017b). Can Affective Computing Save Lives? Meet Mobile Health. *IEEE Computer Magazine*, 50:40. (IF: 1.438 (2013))
- 52) Deng, J., Xu, X., Zhang, Z., Frühholz, S., and Schuller, B. (2017b). Universum Autoencoder-based Domain Adaptation for Speech Emotion Recognition. *IEEE Signal Processing Letters*, 24. 5 pages, to appear (acceptance rate: 17 %, IF: 1.674, 5-year IF: 1.595 (2012))
- 53) Guo, J., Qian, K., Zhang, G., Xu, H., and Schuller, B. (2017c). Accelerating biomedical signal processing using GPU: A case study of snore sounds feature extraction. *Interdisciplinary Sciences – Computational Life Sciences*, 9. 11 pages, to appear (IF: 0.853 (2015))
- 54) Janott, C., Schuller, B., and Heiser, C. (2017). Akustische Informationen von Schnarchgeräuschen. *HNO, Leitthemenheft “Schlafmedizin”*, 22(2):1–10. (IF: 0.852 (2015))
- 55) Marchi, E., Vesperini, F., Squartini, S., and Schuller, B. (2017). Deep Recurrent Neural Network-based Autoencoders for Acoustic Novelty Detection. *Computational Intelligence and Neuroscience*, 2017. 14 pages (IF: 0.430 (2015))
- 56) Marschik, P. B., Pokorný, F. B., Peharz, R., Zhang, D., O’Muircheartaigh, J., Roeyers, H., Bölte, S., Spittle, A. J., Urlesberger, B., Schuller, B., Poustka, L., Ozonoff, S., Pernkopf, F., Pock, T., Tammimies, K., Enzinger, C., Kriebler, M., Tomantschger, I., Bartl-Pokorný, K. D., Sigafos, J., Roche, L., Esposito, G., Gugatschka, M., Nielsen-Saines, K., Einspieler, C., Kaufmann, W. E., and The BEE-PRI study group (2017). A Novel Way to Measure and Predict Development: A Heuristic Approach to Facilitate the Early Detection of Neurodevelopmental Disorders. *Current Neurology and Neuroscience Reports*, 17(43). 15 pages (IF: 2.961, 5-year IF: 3.064 (2015))
- 57) Qian, K., Zhang, Z., Baird, A., and Schuller, B. (2017b). Active Learning for Bird Sounds Classification. *Acta Acustica united with Acustica*, 103:361–364. (IF: 0.897 (2015))
- 58) Rudovic, O., Lee, J., Mascarell-Maricic, L., Schuller, B. W., and Picard, R. (2017). Measuring Engagement in Autism Therapy with Social Robots: a Cross-cultural Study. *Frontiers in Robotics and AI, section Humanoid Robotics, Special Issue on Affective and Social Signals for HRI*. to appear
- 59) Trigeorgis, G., Bousmalis, K., Zafeiriou, S., and Schuller, B. (2017a). A deep matrix factorization method for learning attribute representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(3):417–429. (IF: 6.077, 5-year IF: 8.693 (2015))
- 60) Trigeorgis, G., Nicolaou, M. A., Zafeiriou, S., and Schuller, B. (2017b). Deep Canonical Time Warping for simultaneous alignment and representation learning of sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39. 12 pages, to appear (IF: 6.077, 5-year IF: 8.693 (2015))
- 61) Xu, X., Deng, J., Cummins, N., Zhang, Z., Wu, C., Zhao, L., and Schuller, B. (2017). A Two-Dimensional Framework of Multiple Kernel Subspace Learning for Recognising Emotion in Speech. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 25. 15 pages, to appear (acceptance rate: 25 %, IF: 2.625, 5-year IF: 2.583 (2013))
- 62) Zhang, Z., Cummins, N., and Schuller, B. (2017b). Efficient Data Exploration for Automatic Speech Analysis: Challenges and State of the Art. *IEEE Signal Processing Magazine*, 34. 15 pages, to appear (IF: 6.671 (2015))
- 63) Schuller, B. (2016c). Can Virtual Human Interviewers “Hear” Real Humans’ Depression? *IEEE Computer Magazine*, 49(7):8. (IF: 1.438 (2013))
- 64) Schuller, B. (2016f). Speech Emotion Recognition: 20 Years and Beyond. *Communications of the ACM*. 8 pages, to appear (IF: 3.301, 5-year IF: 4.425 (2015))
- 65) Deng, J., Xu, X., Zhang, Z., Frühholz, S., and Schuller, B. (2016c). Exploitation of Phase-based Features for Whispered Speech Emotion Recognition. *IEEE Access*, 4:4299–4309. (IF: 1.270, 5-year IF: 1.276 (2015))
- 66) Eyben, F., Scherer, K., Schuller, B., Sundberg, J., André, E., Busso, C., Devillers, L., Epps, J., Laukka, P., Narayanan, S., and Truong, K. (2016). The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. *IEEE Transactions on Affective Computing*, 7(2):190–202. (acceptance rate: 22 %, IF: 3.466, 5-year IF: 3.871 (2013))
- 67) Gross, F., Jordan, J., Wenginger, F., Klanner, F., and Schuller, B. (2016). Route and Stopping Intent Prediction at Intersections from Car Fleet Data. *IEEE Transactions on Intelligent Vehicles*, 1(2):177–186
- 68) Han, W., Coutinho, E., Ruan, H., Li, H., Schuller, B., Yu, X., and Zhu, X. (2016b). Semi-Supervised Active Learning for Sound Classification in Hybrid Learning Environments. *PLoS ONE*, 11(9). 23 pages (IF: 3.234 (2014))
- 69) Han, J., Zhang, Z., Cummins, N., Ringeval, F., and Schuller, B. (2016a). Strength Modelling for Real-World Automatic Continuous Affect Recognition from Audiovisual Signals. *Image and Vision Computing, Special Issue on Multimodal Sentiment Analysis and Mining in the Wild*, 34. 12 pages, to appear (IF: 1.766, 5-Year IF: 2.584 (2015))
- 70) Hantke, S., Wenginger, F., Kurler, R., Ringeval, F., Batliner, A., El-Desoky Mousa, A., and Schuller, B. (2016c). I Hear You Eat and Speak: Automatic Recognition of Eating Condition and Food Types, Use-Cases, and Impact on ASR Performance. *PLoS ONE*, 11(5):1–24. (IF: 3.234 (2014))
- 71) Lingenfelser, F., Wagner, J., Deng, J., Brueckner, R., Schuller, B., and André, E. (2016). Asynchronous and Event-based Fusion Systems for Affect Recognition on Naturalistic Data in

- Comparison to Conventional Approaches. *IEEE Transactions on Affective Computing*, 7. 13 pages, to appear (IF: 3.466, 5-year IF: 3.871 (2013))
- 72) Marchi, E., Frühholz, S., and Schuller, B. (2016b). The Effect of Narrow-band Transmission on Recognition of Paralinguistic Information from Human Vocalizations. *IEEE Access*, 4:6059–6072. (IF: 1.270, 5-year IF: 1.276 (2015))
- 73) Mencattini, A., Martinelli, E., Ringeval, F., Schuller, B., and Di Natale, C. (2016). Continuous Estimation of Emotions in Speech by Dynamic Cooperative Speaker Models. *IEEE Transactions on Affective Computing*, 7. 14 pages, to appear (IF: 3.466, 5-year IF: 3.871 (2013))
- 74) Qian, K., Janott, C., Pandit, V., Zhang, Z., Heiser, C., Hohenhorst, W., Herzog, M., Hemmert, W., and Schuller, B. (2016a). Classification of the Excitation Location of Snore Sounds in the Upper Airway by Acoustic Multi-Feature Analysis. *IEEE Transactions on Biomedical Engineering*. 11 pages, to appear (IF: 2.468, 5-year IF: 2.719 (2015))
- 75) Rynkiewicz, A., Schuller, B., Marchi, E., Piana, S., Camurri, A., Lassalle, A., and Baron-Cohen, S. (2016a). An investigation of the ‘female camouflage effect’ in autism using a computerized ADOS-2, and a test of sex/gender differences. *Molecular Autism*, 7(10). 10 pages (IF: 5.413 (2014))
- 76) Sagha, H., Cummins, N., and Schuller, B. (2017). Autoencoders for Sentiment Analysis: A Review. *WIREs Data Mining and Knowledge Discovery*, 7. to appear, invited review article (IF: 1.759 (2015))
- 77) Sagha, H., Li, F., del R. Millán, E. V. J., Chavarriaga, R., and Schuller, B. (2016b). Stream fusion for multi-stream automatic speech recognition. *International Journal of Speech Technology*, 19(4):669–675
- 78) Schuller, B., Steidl, S., Batliner, A., Nöth, E., Vinciarelli, A., Burkhardt, F., van Son, R., Wenginger, F., Eyben, F., Bocklet, T., Mohammadi, G., and Weiss, B. (2015e). A Survey on Perceived Speaker Traits: Personality, Likability, Pathology, and the First Challenge. *Computer Speech and Language, Special Issue on Next Generation Computational Paralinguistics*, 29(1):100–131. (acceptance rate: 23 %, IF: 1.812, 5-year IF: 1.776 (2013))
- 79) Schuller, B. (2015c). Do Computers Have Personality? *IEEE Computer Magazine*, 48(3):6–7. (IF: 1.438 (2013))
- 80) Schuller, B., Mousa, A. E.-D., and Vasileios, V. (2015c). Sentiment Analysis and Opinion Mining: On Optimal Parameters and Performances. *WIREs Data Mining and Knowledge Discovery*, 5:255–263. invited focus article (IF: 1.759 (2015))
- 81) Coutinho, E. and Schuller, B. (2015). Automatic estimation of biosignals from the human voice. *Science, Special Supplement on Advances in Computational Psychophysiology*, 350(6256):114:48–50. invited contribution
- 82) Eyben, F., Salomo, G. L., Sundberg, J., Scherer, K., and Schuller, B. (2015b). Emotion in the singing voice - a deeper look at acoustic features in the light of automatic classification. *EURASIP Journal on Audio, Speech, and Music Processing, Special Issue on Scalable Audio-Content Analysis*, 2015. 9 pages (acceptance rate: 33 %, IF: 0.709 (2011))
- 83) Mencattini, A., Ringeval, F., Schuller, B., Martinelli, E., and Natale, C. D. (2015). Continuous monitoring of emotions by a multimodal cooperative sensor system. *Procedia Engineering, Special Issue Euroensors 2015*, 120:556–559
- 84) Wenginger, F., Bergmann, J., and Schuller, B. (2015a). Introducing CURRENNT: the Munich Open-Source CUDA Recurrent Neural Network Toolkit. *Journal of Machine Learning Research*, 16:547–551. 5 pages, (acceptance rate: 18 %, IF: 2.853, 5-year IF: 4.649 (2013))
- 85) Zhang, Z., Coutinho, E., Deng, J., and Schuller, B. (2015d). Cooperative Learning and its Application to Emotion Recognition from Speech. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 23(1):115–126. (acceptance rate: 25 %, IF: 2.625, 5-year IF: 2.583 (2013))
- 86) Schuller, B., Steidl, S., Batliner, A., Schiel, F., Krajewski, J., Wenginger, F., and Eyben, F. (2014e). Medium-Term Speaker States – A Review on Intoxication, Sleepiness and the First Challenge. *Computer Speech and Language, Special Issue on Broadening the View on Speaker Analysis*, 28(2):346–374. (acceptance rate: 23 %, IF: 1.812, 5-year IF: 1.776 (2013))
- 87) Ringeval, F., Eyben, F., Kroupi, E., Yuce, A., Thiran, J.-P., Ebrahimi, T., Lalanne, D., and Schuller, B. (2015a). Prediction of Asynchronous Dimensional Emotion Ratings from Audiovisual and Physiological Data. *Pattern Recognition Letters*, 66:22–30. (acceptance rate: 25 %, IF: 1.062, 5-year IF: 1.466 (2013))
- 88) Rosner, A., Schuller, B., and Kostek, B. (2014). Classification of music genres based on music separation into harmonic and drum components. *Archives of Acoustics*, 39(4):629–638. (IF: 0.829, 5-year IF: 0.500 (2012))
- 89) Zhang, Z., Coutinho, E., Deng, J., and Schuller, B. (2014a). Distributing Recognition in Computational Paralinguistics. *IEEE Transactions on Affective Computing*, 5(4):406–417. (acceptance rate: 22 %, IF: 3.466, 5-year IF: 3.871 (2013))
- 90) Deng, J., Zhang, Z., Eyben, F., and Schuller, B. (2014b). Autoencoder-based Unsupervised Domain Adaptation for Speech Emotion Recognition. *IEEE Signal Processing Letters*, 21(9):1068–1072. (acceptance rate: 17 %, IF: 1.674, 5-year IF: 1.595 (2012))
- 91) Geiger, J. T., Wenginger, F., Gemmeke, J. F., Wöllmer, M., Schuller, B., and Rigoll, G. (2014d). Memory-Enhanced Neural Networks and NMF for Robust ASR. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 22(6):1037–1046. (acceptance rate: 25 %, IF: 2.625, 5-year IF: 2.583 (2013))
- 92) Wenginger, F., Geiger, J., Wöllmer, M., Schuller, B., and Rigoll, G. (2014a). Feature Enhancement by Deep LSTM Networks for ASR in Reverberant Multisource Environments. *Computer Speech and Language*, 28(4):888–902. (acceptance rate: 23 %, IF: 1.812, 5-year IF: 1.776 (2013))
- 93) Wöllmer, M. and Schuller, B. (2014). Probabilistic Speech Feature Extraction with Context-Sensitive Bottleneck Neural Networks. *Neurocomputing, Special Issue on Machines Learning for Non-Linear Processing Selected papers from the 2011 International Conference on Non-Linear Speech Processing (NoLISP 2011)*, 132:113–120. (acceptance rate: 31 %, IF: 2.005 (2013))
- 94) Zhang, Z., Pinto, J., Plahl, C., Schuller, B., and Willett, D. (2014c). Channel Mapping using Bidirectional Long Short-Term Memory for Dereverberation in Hands-Free Voice Controlled Devices. *IEEE Transactions on Consumer Electronics*, 60(3):525–533. (acceptance rate: 15 %, IF: 1.157 (2013))
- 95) Schuller, B., Dunwell, I., Wenginger, F., and Paletta, L. (2013a). Serious Gaming for Behavior Change – The State of Play. *IEEE Pervasive Computing Magazine, Special Issue on Understanding and Changing Behavior*, 12(3):48–55. (IF: 2.103 (2013))
- 96) Schuller, B., Steidl, S., Batliner, A., Burkhardt, F., Devillers, L., Müller, C., and Narayanan, S. (2013h). Paralinguistics in Speech and Language – State-of-the-Art and the Challenge. *Computer Speech and Language, Special Issue on Paralinguistics in Naturalistic Speech and Language*, 27(1):4–39. (CSL most downloaded article 2012–2014 (3 208 downloads) (acceptance rate: 36 %, IF: 1.812 (2013), 5-year IF: 1.776, > 100 citations)
- 97) Cambria, E., Schuller, B., Xia, Y., and Havasi, C. (2013c). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems Magazine*, 28(2):15–21. (IF: 2.570, 5-year IF: 2.632 (2010), > 300 citations)
- 98) Eyben, F., Wenginger, F., Lehment, N., Schuller, B., and Rigoll, G. (2013b). Affective Video Retrieval: Violence Detection in Hollywood Movies by Large-Scale Segmental Feature Extraction. *PLOS ONE*, 8(12):1–9. (acceptance rate: 50 %, IF: 3.534 (2013))
- 99) Gunes, H. and Schuller, B. (2013a). Categorical and Dimensional Affect Analysis in Continuous Input: Current Trends and Future Directions. *Image and Vision Computing, Special*

- Issue on Affect Analysis in Continuous Input*, 31(2):120–136. (acceptance rate: 20 %, IF: 1.581 (2013), > 100 citations)
- 100) Hofmann, M., Geiger, J., Bachmann, S., Schuller, B., and Rigoll, G. (2013). The TUM Gait from Audio, Image and Depth (GAID) Database: Multimodal Recognition of Subjects and Traits. *Journal of Visual Communication and Image Representation, Special Issue on Visual Understanding and Applications with RGB-D Cameras*, 25(1):195–206. (IF: 1.361 (2013))
- 101) Weninger, F., Eyben, F., Schuller, B. W., Mortillaro, M., and Scherer, K. R. (2013b). On the Acoustics of Emotion in Audio: What Speech, Music and Sound have in Common. *Frontiers in Psychology, section Emotion Science, Special Issue on Expression of emotion in music and vocal communication*, 4(Article ID 292):1–12. (IF: 2.843 (2013))
- 102) Weninger, F., Staudt, P., and Schuller, B. (2013e). Words that Fascinate the Listener: Predicting Affective Ratings of On-Line Lectures. *International Journal of Distance Education Technologies, Special Issue on Emotional Intelligence for Online Learning*, 11(2):110–123
- 103) Wöllmer, M., Schuller, B., and Rigoll, G. (2013b). Keyword Spotting Exploiting Long Short-Term Memory. *Speech Communication*, 55(2):252–265. (acceptance rate: 38 %, IF: 1.267, 5-year IF: 1.526 (2011))
- 104) Wöllmer, M., Weninger, F., Knaup, T., Schuller, B., Sun, C., Sagae, K., and Morency, L.-P. (2013e). YouTube Movie Reviews: Sentiment Analysis in an Audiovisual Context. *IEEE Intelligent Systems Magazine, Special Issue on Statistical Approaches to Concept-Level Sentiment Analysis*, 28(3):46–53. (acceptance rate: 17 %, IF: 2.570, 5-year IF: 2.632 (2010))
- 105) Wöllmer, M., Kaiser, M., Eyben, F., Schuller, B., and Rigoll, G. (2013a). LSTM-Modeling of Continuous Emotions in an Audiovisual Affect Recognition Framework. *Image and Vision Computing, Special Issue on Affect Analysis in Continuous Input*, 31(2):153–163. (acceptance rate: 20 %, IF: 1.581 (2013))
- 106) Schuller, B. (2012c). The Computational Paralinguistics Challenge. *IEEE Signal Processing Magazine*, 29(4):97–101. (IF: 6.000, 5-year IF: 7.043 (2011))
- 107) Schuller, B. and Gollan, B. (2012). Music Theoretic and Perception-based Features for Audio Key Determination. *Journal of New Music Research*, 41(2):175–193. (IF: 0.755, 5-year IF: 0.768 (2010))
- 108) Schuller, B. (2012b). Recognizing Affect from Linguistic Information in 3D Continuous Space. *IEEE Transactions on Affective Computing*, 2(4):192–205. (acceptance rate: 22 %, IF: 3.466, 5-year IF: 3.871 (2013))
- 109) Schuller, B., Zhang, Z., Weninger, F., and Burkhardt, F. (2012h). Synthesized Speech for Model Training in Cross-Corpus Recognition of Human Emotion. *International Journal of Speech Technology, Special Issue on New and Improved Advances in Speaker Recognition Technologies*, 15(3):313–323
- 110) Rotili, R., Principi, E., Squartini, S., and Schuller, B. (2012a). A Real-Time Speech Enhancement Framework in Noisy and Reverberated Acoustic Scenarios. *Cognitive Computation*, 5(4):504–516. (IF: 1.000 (2011))
- 111) Eyben, F., Batliner, A., and Schuller, B. (2012a). Towards a standard set of acoustic features for the processing of emotion in speech. *Proceedings of Meetings on Acoustics*, 16. 11 pages
- 112) Weninger, F., Krajewski, J., Batliner, A., and Schuller, B. (2012c). The Voice of Leadership: Models and Performances of Automatic Analysis in On-Line Speeches. *IEEE Transactions on Affective Computing*, 3(4):496–508. (acceptance rate: 22 %, IF: 3.466, 5-year IF: 3.871 (2013))
- 113) Wöllmer, M., Weninger, F., Geiger, J., Schuller, B., and Rigoll, G. (2013d). Noise Robust ASR in Reverberated Multisource Environments Applying Convolutional NMF and Long Short-Term Memory. *Computer Speech and Language, Special Issue on Speech Separation and Recognition in Multisource Environments*, 27(3):780–797. (acceptance rate: 36 %, IF: 1.812, 5-year IF: 1.776 (2013))
- 114) Weninger, F. and Schuller, B. (2012b). Optimization and Parallelization of Monaural Source Separation Algorithms in the openBLISSART Toolkit. *Journal of Signal Processing Systems*, 69(3):267–277. (IF: 0.672, 5-year IF: 0.684 (2011))
- 115) Principi, E., Rotili, R., Wöllmer, M., Eyben, F., Squartini, S., and Schuller, B. (2012a). Real-Time Activity Detection in a Multi-Talker Reverberated Environment. *Cognitive Computation, Special Issue on Cognitive and Emotional Information Processing for Human-Machine Interaction*, 4(4):386–397. (IF: 1.000 (2011))
- 116) Krajewski, J., Schnieder, S., Sommer, D., Batliner, A., and Schuller, B. (2012). Applying Multiple Classifiers and Non-Linear Dynamics Features for Detecting Sleepiness from Speech. *Neurocomputing, Special Issue From neuron to behavior: evidence from behavioral measurements*, 84:65–75. (acceptance rate: 31 %, IF: 2.005 (2013))
- 117) Metallinou, A., Wöllmer, M., Katsamanis, A., Eyben, F., Schuller, B., and Narayanan, S. (2012). Context-Sensitive Learning for Enhanced Audiovisual Emotion Classification. *IEEE Transactions on Affective Computing*, 3(2):184–198. (acceptance rate: 22 %, IF: 3.466, 5-year IF: 3.871 (2013))
- 118) Eyben, F., Wöllmer, M., and Schuller, B. (2012e). A Multi-Task Approach to Continuous Five-Dimensional Affect Sensing in Natural Speech. *ACM Transactions on Interactive Intelligent Systems, Special Issue on Affective Interaction in Natural Environments*, 2(1). 29 pages
- 119) Grosche, P., Schuller, B., Müller, M., and Rigoll, G. (2012). Automatic Transcription of Recorded Music. *Acta Acustica united with Acustica*, 98(2):199–215(17). (IF: 0.714 (2012))
- 120) Schröder, M., Bevacqua, E., Cowie, R., Eyben, F., Gunes, H., Heylen, D., ter Maat, M., McKeown, G., Pammi, S., Pantic, M., Pelachaud, C., Schuller, B., de Sevin, E., Valstar, M., and Wöllmer, M. (2012). Building Autonomous Sensitive Artificial Listeners. *IEEE Transactions on Affective Computing*, 3(2):165–183. (acceptance rate: 22 %, IF: 3.466, 5-year IF: 3.871 (2013), > 100 citations)
- 121) Schuller, B. (2011a). Affective Speaker State Analysis in the Presence of Reverberation. *International Journal of Speech Technology*, 14(2):77–87
- 122) Schuller, B., Batliner, A., Steidl, S., and Seppi, D. (2011b). Recognising Realistic Emotions and Affect in Speech: State of the Art and Lessons Learnt from the First Challenge. *Speech Communication, Special Issue on Sensing Emotion and Affect - Facing Realism in Speech Processing*, 53(9/10):1062–1087. (acceptance rate: 38 %, IF: 1.267, 5-year IF: 1.526 (2011), > 300 citations)
- 123) Wöllmer, M., Marchi, E., Squartini, S., and Schuller, B. (2011c). Multi-Stream LSTM-HMM Decoding and Histogram Equalization for Noise Robust Keyword Spotting. *Cognitive Neurodynamics*, 5(3):253–264. (acceptance rate: 50 %, IF: 1.77 (2013))
- 124) Wöllmer, M., Weninger, F., Eyben, F., and Schuller, B. (2011i). Computational Assessment of Interest in Speech – Facing the Real-Life Challenge. *Künstliche Intelligenz (German Journal on Artificial Intelligence), Special Issue on Emotion and Computing*, 25(3):227–236
- 125) Wöllmer, M., Blaschke, C., Schindl, T., Schuller, B., Färber, B., Mayer, S., and Trefflich, B. (2011a). On-line Driver Distraction Detection using Long Short-Term Memory. *IEEE Transactions on Intelligent Transportation Systems*, 12(2):574–582. (IF: 3.452, 5-year IF: 4.090 (2011))
- 126) Wöllmer, M., Schuller, B., Batliner, A., Steidl, S., and Seppi, D. (2011e). Tandem Decoding of Children’s Speech for Keyword Detection in a Child-Robot Interaction Scenario. *ACM Transactions on Speech and Language Processing, Special Issue on Speech and Language Processing of Children’s Speech for Child-machine Interaction Applications*, 7(4). 22 pages, Article 12
- 127) Weninger, F., Schuller, B., Batliner, A., Steidl, S., and Seppi,

- D. (2011d). Recognition of Non-Prototypical Emotions in Reverberated and Noisy Speech by Non-Negative Matrix Factorization. *EURASIP Journal on Advances in Signal Processing, Special Issue on Emotion and Mental State Recognition from Speech*, 2011(Article ID 838790). 16 pages (acceptance rate: 38 %, IF: 1.012, 5-Year IF: 1.136 (2010))
- 128) Batliner, A., Steidl, S., Schuller, B., Seppi, D., Vogt, T., Wagner, J., Devillers, L., Vidrascu, L., Aharonson, V., Kessous, L., and Amir, N. (2011). Whodunnit – Searching for the Most Important Feature Types Signalling Emotion-Related User States in Speech. *Computer Speech and Language, Special Issue on Affective Speech in real-life interactions*, 25(1):4–28. (acceptance rate: 31 %, IF: 1.812, 5-year IF: 1.776 (2013), > 100 citations)
- 129) Schuller, B., Vlasenko, B., Eyben, F., Wöllmer, M., Stuhlsatz, A., Wendemuth, A., and Rigoll, G. (2010h). Cross-Corpus Acoustic Emotion Recognition: Variances and Strategies. *IEEE Transactions on Affective Computing*, 1(2):119–131. (acceptance rate: 32 %, IF: 3.466, 5-year IF: 3.871 (2013), > 100 citations)
- 130) Schuller, B., Hage, C., Schuller, D., and Rigoll, G. (2010d). “Mister D.J., Cheer Me Up!”: Musical and Textual Features for Automatic Mood Classification. *Journal of New Music Research*, 39(1):13–34. (IF: 0.755, 5-year IF: 0.768 (2010))
- 131) Schuller, B. (2010). On the acoustics of emotion in speech: Desperately seeking a standard. *Journal of the Acoustical Society of America*, 127(3):1995–1995. (IF: 1.644, 5-year IF: 1.899 (2010))
- 132) Schuller, B., Dorfner, J., and Rigoll, G. (2010b). Determination of Non-Prototypical Valence and Arousal in Popular Music: Features and Performances. *EURASIP Journal on Audio, Speech, and Music Processing, Special Issue on Scalable Audio-Content Analysis*, 2010(Article ID 735854). 19 pages, (acceptance rate: 33 %, IF: 0.709 (2011))
- 133) Steidl, S., Batliner, A., Seppi, D., and Schuller, B. (2010). On the Impact of Children’s Emotional Speech on Acoustic and Language Models. *EURASIP Journal on Audio, Speech, and Music Processing, Special Issue on Atypical Speech*, 2010(Article ID 783954). 14 pages (acceptance rate: 33 %, IF: 0.709 (2011))
- 134) Batliner, A., Seppi, D., Steidl, S., and Schuller, B. (2010b). Segmenting into adequate units for automatic recognition of emotion-related episodes: a speech-based approach. *Advances in Human Computer Interaction, Special Issue on Emotion-Aware Natural Interaction*, 2010(Article ID 782802). 15 pages
- 135) Eyben, F., Wöllmer, M., Graves, A., Schuller, B., Douglas-Cowie, E., and Cowie, R. (2010c). On-line Emotion Recognition in a 3-D Activation-Valence-Time Continuum using Acoustic and Linguistic Cues. *Journal on Multimodal User Interfaces, Special Issue on Real-Time Affect Analysis and Interpretation: Closing the Affective Loop in Virtual Agents and Robots*, 3(1–2):7–12. (IF: 0.833 (2012))
- 136) Can, S., Schuller, B., Kranzfelder, M., and Feussner, H. (2010). Emotional factors in speech based human-machine interaction in the operating room. *International Journal of Computer Assisted Radiology and Surgery*, 5(Supplement 1):188–189. (acceptance rate: 75 %, IF: 1.659 (2013))
- 137) Wöllmer, M., Eyben, F., Graves, A., Schuller, B., and Rigoll, G. (2010a). Bidirectional LSTM Networks for Context-Sensitive Keyword Detection in a Cognitive Virtual Agent Framework. *Cognitive Computation, Special Issue on Non-Linear and Non-Conventional Speech Processing*, 2(3):180–190. (IF: 1.1 (2013))
- 138) Eyben, F., Wöllmer, M., Poitschke, T., Schuller, B., Blaschke, C., Färber, B., and Nguyen-Thien, N. (2010d). Emotion on the Road – Necessity, Acceptance, and Feasibility of Affective Computing in the Car. *Advances in Human Computer Interaction, Special Issue on Emotion-Aware Natural Interaction*, 2010(Article ID 263593). 17 pages
- 139) Wöllmer, M., Schuller, B., Eyben, F., and Rigoll, G. (2010g). Combining Long Short-Term Memory and Dynamic Bayesian Networks for Incremental Emotion-Sensitive Artificial Listening. *IEEE Journal of Selected Topics in Signal Processing, Special Issue on Speech Processing for Natural Interaction with Intelligent Environments*, 4(5):867–881. (IF: 2.571 (2010), > 100 citations)
- 140) Schuller, B., Müller, R., Eyben, F., Gast, J., Hörnler, B., Wöllmer, M., Rigoll, G., Höthker, A., and Konosu, H. (2009f). Being Bored? Recognising Natural Interest by Extensive Audio-visual Integration for Real-Life Application. *Image and Vision Computing, Special Issue on Visual and Multimodal Analysis of Human Spontaneous Behavior*, 27(12):1760–1774. (IF: 1.474, 5-Year IF: 1.767 (2009), > 100 citations)
- 141) Schuller, B., Wöllmer, M., Moosmayr, T., and Rigoll, G. (2009l). Recognition of Noisy Speech: A Comparative Survey of Robust Model Architecture and Feature Enhancement. *EURASIP Journal on Audio, Speech, and Music Processing*, 2009(Article ID 942617). 17 pages (acceptance rate: 33 %, IF: 0.709 (2011))
- 142) Wöllmer, M., Al-Hames, M., Eyben, F., Schuller, B., and Rigoll, G. (2009a). A Multidimensional Dynamic Time Warping Algorithm for Efficient Multimodal Fusion of Asynchronous Data Streams. *Neurocomputing*, 73(1–3):366–380. (IF: 1.440, 5-year IF: 1.459 (2009))
- 143) Schuller, B., Eyben, F., and Rigoll, G. (2008f). Tango or Waltz? – Putting Ballroom Dance Style into Tempo Detection. *EURASIP Journal on Audio, Speech, and Music Processing, Special Issue on Intelligent Audio, Speech, and Music Processing Applications*, 2008(Article ID 846135). 12 pages (acceptance rate: 33 %, IF: 0.709 (2011))
- 144) Nieschulz, R., Schuller, B., Geiger, M., and Neuss, R. (2002). Aspects of Efficient Usability Engineering. *Information Technology, Special Issue on Usability Engineering*, 44(1):23–30

C) REFEREED CONFERENCE PROCEEDINGS

Contributions to Conference Proceedings (353):

- 145) Schuller, B. (2017a). Big Data, Deep Learning – At the Edge of X-Ray Speaker Analysis. In *Proceedings 19th International Conference on Speech and Computer, SPECOM 2017, Hatfield, UK, 12.-16.09.2017*, Lecture Notes in Computer Science (LNCS). Springer, Berlin/Heidelberg. 10 pages, to appear
- 146) Schuller, B. (2017e). Reading the Author and Speaker: Towards a Holistic Approach on Automatic Assessment of What is in One’s Words. In *Proceedings 18th International Conference on Intelligent Text Processing and Computational Linguistics, CI-Ling 2017, Budapest, Hungary, 17.-23.04.2017*, Lecture Notes in Computer Science (LNCS). Springer, Berlin/Heidelberg. 12 pages, to appear
- 147) Schuller, B., Steidl, S., Batliner, A., Bergelson, E., Krajewski, J., Janott, C., Amatuni, A., Casillas, M., Seidl, A., Soderstrom, M., Warlaumont, A., Hidalgo, G., Schnieder, S., Heiser, C., Hohenhorst, W., Herzog, M., Schmitt, M., Qian, K., Zhang, Y., Trigeorgis, G., Tzirakis, P., and Zafeiriou, S. (2017d). The INTERSPEECH 2017 Computational Paralinguistics Challenge: Addressee, Cold & Snoring. In *Proceedings INTERSPEECH 2017, 18th Annual Conference of the International Speech Communication Association*, Stockholm, Sweden. ISCA, ISCA. 5 pages, to appear
- 148) Böhm, J., Eyben, F., Schmitt, M., Kosch, H., and Schuller, B. (2017). Seeking the SuperStar: Automatic Assessment of Perceived Singing Quality. In *Proceedings 30th International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK. IEEE, IEEE. 10 pages, to appear
- 149) Cummins, N., Vlasenko, B., Sagha, H., and Schuller, B. (2017c). Enhancing speech-based depression detection through gender dependent vowel level formant features. In *Proc. 16th Conference on Artificial Intelligence in Medicine (AIME)*,

- Vienna, Austria. Society for Artificial Intelligence in MEDicine (AIME). 5 pages, to appear
- 150) Cummins, N., Schmitt, M., Amiriparian, S., Krajewski, J., and Schuller, B. (2017b). You sound ill, take the day off: Classification of speech affected by Upper Respiratory Tract Infection. In *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine & Biology Society, EMBC 2017*, Jeju Island, South Korea. IEEE, IEEE. 4 pages, to appear
 - 151) Deng, J., Cummins, N., Schmitt, M., Qian, K., Ringeval, F., and Schuller, B. (2017a). Speech-based Diagnosis of Autism Spectrum Condition by Generative Adversarial Network Representations. In *Proceedings of the 7th International Digital Health Conference, DH 2017*, London, U.K. ACM, ACM. 5 pages, to appear
 - 152) Eyben, F., Unfried, M., Hagerer, G., and Schuller, B. (2017). Automatic Multi-lingual Arousal Detection from Voice Applied to Real Product Testing Applications. In *Proceedings 42nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2017*, pages 5155–5159, New Orleans, LA. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
 - 153) Guo, J., Qian, K., Schuller, B. W., and Matsuoka, S. (2017b). GPU-based Training of Autoencoders for Bird Sound Data Processing. In *Proceedings IEEE International Conference on Consumer Electronics Taiwan, ICCE-TW 2017*, Taipei, Taiwan. IEEE, IEEE. 2 pages, to appear
 - 154) Hagerer, G., Pandit, V., Eyben, F., and Schuller, B. (2017). Enhancing LSTM RNN-based Speech Overlap Detection by Artificially Mixed Data. In *Proceedings AES 56th International Conference on Semantic Audio*, pages 1–8, Erlangen, Germany. AES, Audio Engineering Society. to appear
 - 155) Han, J., Zhang, Z., Ringeval, F., and Schuller, B. (2017a). Prediction-based Learning from Continuous Emotion Recognition in Speech. In *Proceedings 42nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2017*, pages 5005–5009, New Orleans, LA. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
 - 156) Han, J., Zhang, Z., Ringeval, F., and Schuller, B. (2017b). Reconstruction-error-based Learning for Continuous Emotion Recognition in Speech. In *Proceedings 42nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2017*, pages 2367–2371, New Orleans, LA. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
 - 157) Keren, G., Kirschstein, T., Marchi, E., Ringeval, F., and Schuller, B. (2017a). End-to-end learning for dimensional emotion recognition from physiological signals. In *Proceedings 18th IEEE International Conference on Multimedia and Expo, ICME 2017*, Hong Kong, P.R. China. IEEE, IEEE. 6 pages, to appear (acceptance rate: 30 %, IF* 0.88 (2010))
 - 158) Mousa, A. E.-D. and Schuller, B. (2017). Contextual Bidirectional Long Short-Term Memory Recurrent Neural Network Language Models: A Generative Approach to Sentiment Analysis. In *Proceedings EACL 2017, 15th Conference of the European Chapter of the Association for Computational Linguistics*, Valencia, Spain. ACL, ACL. 9 pages, to appear
 - 159) Parada-Cabaleiro, E., Baird, A. E., Cummins, N., and Schuller, B. (2017). Stimulation of Psychological Listener Experiences by Semi-Automatically Composed Electroacoustic Environments. In *Proceedings 18th IEEE International Conference on Multimedia and Expo, ICME 2017*, Hong Kong, P.R. China. IEEE, IEEE. 7 pages, to appear (acceptance rate: 30 %, IF* 0.88 (2010))
 - 160) Qian, K., Janott, C., Deng, J., Heiser, C., Hohenhorst, W., Cummins, N., and Schuller, B. (2017a). Snore Sound Recognition: On Wavelets and Classifiers from Deep Nets to Kernels. In *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine & Biology Society, EMBC 2017*, Jeju Island, South Korea. IEEE, IEEE. 4 pages, to appear
 - 161) Sabathé, R., Coutinho, E., and Schuller, B. (2017). Deep Recurrent Music Writer: Memory-enhanced Variational Autoencoder-based Musical Score Composition and an Objective Measure. In *Proceedings 30th International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK. IEEE, IEEE. 8 pages, to appear
 - 162) Walecki, R., Rudovic, O., Pavlovic, V., Schuller, B., and Pantic, M. (2017). Deep Structured Ordinal Regression for Facial Action Unit Intensity Estimation. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Honolulu, HI. IEEE, IEEE. (acceptance rate: 20 %, IF* 5.97 (2010))
 - 163) Zhang, Y., Liu, Y., Weninger, F., and Schuller, B. (2017a). Multi-Task Deep Neural Network with Shared Hidden Layers: Breaking Down the Wall between Emotion Representations. In *Proceedings 42nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2017*, pages 4990–4994, New Orleans, LA. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
 - 164) Zhang, Z., Weninger, F., Wöllmer, M., Han, J., and Schuller, B. (2017c). Towards Intoxicated Speech Recognition. In *Proceedings 30th International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK. IEEE, IEEE. 8 pages, to appear
 - 165) Schuller, B. (2016a). 7 Essential Principles to Make Multimodal Sentiment Analysis Work in the Wild. In Bandyopadhyay, S., Das, D., Cambria, E., and Patra, B. G., editors, *Proceedings of the 4th Workshop on Sentiment Analysis where AI meets Psychology (SAAIP 2016)*, satellite of the 25th International Joint Conference on Artificial Intelligence, IJCAI 2016, volume 1619, page 1, New York City, NY. IJCAI/AAAI, CEUR. invited contribution
 - 166) Schuller, B., Ganascia, J.-G., and Devillers, L. (2016a). Multimodal Sentiment Analysis in the Wild: Ethical considerations on Data Collection, Annotation, and Exploitation. In Devillers, L., Schuller, B., Mower Provost, E., Robinson, P., Mariani, J., and Delaborde, A., editors, *Proceedings of the 1st International Workshop on ETHics In Corpus Collection, Annotation and Application (ETHI-CA² 2016)*, satellite of the 10th Language Resources and Evaluation Conference (LREC 2016), pages 29–34, Portoroz, Slovenia. ELRA, ELRA
 - 167) Schuller, B. and McTear, M. (2016). Sociocognitive Language Processing - Emphasising the Soft Factors. In *Proceedings of the Seventh International Workshop on Spoken Dialogue Systems (IWSDS)*, Saariselkä, Finland. 6 pages
 - 168) Schuller, B., Steidl, S., Batliner, A., Hirschberg, J., Burgoon, J. K., Baird, A., Elkins, A., Zhang, Y., Coutinho, E., and Evanini, K. (2016b). The INTERSPEECH 2016 Computational Paralinguistics Challenge: Deception, Sincerity & Native Language. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 2001–2005, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
 - 169) Abdić, I., Fridman, L., McDuff, D., Marchi, E., Reimer, B., and Schuller, B. (2016b). Driver Frustration Detection From Audio and Video. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence, IJCAI 2016*, pages 1354–1360, New York City, NY. IJCAI/AAAI. (acceptance rate: 25 %)
 - 170) Abdić, I., Fridman, L., Brown, D. E., Angell, W., Reimer, B., Marchi, E., and Schuller, B. (2016a). Detecting Road Surface Wetness from Audio: A Deep Learning Approach. In *Proceedings 23rd International Conference on Pattern Recognition (ICPR 2016)*, Cancun, Mexico. IAPR, IAPR. 6 pages
 - 171) Abdić, I., Fridman, L., McDuff, D., Marchi, E., Reimer, B., and Schuller, B. (2016c). Driver Frustration Detection From Audio and Video (Extended Abstract). In Friedrich, G., Helmert, M., and Wotawa, F., editors, *KI 2016: Advances in Artificial Intelligence 39th Annual German Conference on AI*, volume LNCS Volume 9904/2016, pages 237–243, Klagenfurt, Austria.

- GfI / ÖGAI, Springer. (acceptance rate: 45 %)
- 172) Amiriparian, S., Pohjalainen, J., Marchi, E., Pugachevskiy, S., and Schuller, B. (2016). Is deception emotional? An emotion-driven predictive approach. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 2011–2015, San Francisco, CA. ISCA, ISCA. nominated for best student paper award (12 nominations for 3 awards, overall conference 50 % acceptance rate)
- 173) Cambria, E., Poria, S., Bajpai, R., and Schuller, B. (2016a). SenticNet4: A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives. In *Proceedings of the 26th International Conference on Computational Linguistics, COLING*, pages 2666–2677, Osaka, Japan. ICCL, ANLP
- 174) Coutinho, E., Hönig, F., Zhang, Y., Hantke, S., Batliner, A., Nöth, E., and Schuller, B. (2016). Assessing the Prosody of Non-Native Speakers of English: Measures and Feature Sets. In *Proceedings 10th Language Resources and Evaluation Conference (LREC 2016)*, pages 1328–1332, Portoroz, Slovenia. ELRA, ELRA
- 175) Deng, J., Cummins, N., Han, J., Xu, X., Ren, Z., Pandit, V., Zhang, Z., and Schuller, B. (2016a). The University of Passau Open Emotion Recognition System for the Multimodal Emotion Challenge. In *Proceedings of the 7th Chinese Conference on Pattern Recognition, CCPR*, pages 652–666, Chengdu, P.R. China. Springer
- 176) Dong, B., Zhang, Z., and Schuller, B. (2016). Empirical Mode Decomposition: A Data-Enrichment Perspective on Speech Emotion Recognition. In Sánchez-Rada, J. and Schuller, B., editors, *Proceedings of the 6th International Workshop on Emotion and Sentiment Analysis (ESA 2016), satellite of the 10th Language Resources and Evaluation Conference (LREC 2016)*, pages 71–75, Portoroz, Slovenia. ELRA, ELRA. (74 % acceptance rate)
- 177) Guo, J., Qian, K., Xu, H., Janott, C., Schuller, B. W., and Matsuoka, S. (2016). GPU-Based Fast Signal Processing for Large Amounts of Snore Sound Data. In *Proceedings IEEE 5th Global Conference on Consumer Electronics, GCCE 2016*, pages 523–524, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 66 %)
- 178) Hantke, S., Batliner, A., and Schuller, B. (2016a). Ethics for Crowdsourced Corpus Collection, Data Annotation and its Application in the Web-based Game iHEARu-PLAY. In Devillers, L., Schuller, B., Mower Provost, E., Robinson, P., Mariani, J., and Delaborde, A., editors, *Proceedings of the 1st International Workshop on ETHics In Corpus Collection, Annotation and Application (ETHI-CA² 2016), satellite of the 10th Language Resources and Evaluation Conference (LREC 2016)*, pages 54–59, Portoroz, Slovenia. ELRA, ELRA. (oral acceptance rate: 45 %)
- 179) Hantke, S., Marchi, E., and Schuller, B. (2016b). Introducing the Weighted Trustability Evaluator for Crowdsourcing Exemplified by Speaker Likability Classification. In *Proceedings 10th Language Resources and Evaluation Conference (LREC 2016)*, pages 2156–2161, Portoroz, Slovenia. ELRA, ELRA
- 180) Keren, G., Deng, J., Pohjalainen, J., and Schuller, B. (2016a). Convolutional Neural Networks and Data Augmentation for Classifying Speakers' Native Language. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 2393–2397, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 181) Keren, G. and Schuller, B. (2016a). Convolutional RNN: an Enhanced Model for Extracting Features from Sequential Data. In *Proceedings 2016 International Joint Conference on Neural Networks (IJCNN) as part of the IEEE World Congress on Computational Intelligence (IEEE WCCI)*, pages 3412–3419, Vancouver, Canada. IEEE, IEEE
- 182) Keren, G., Sabato, S., and Schuller, B. (2017c). Tunable Sensitivity to Large Errors in Neural Network Training. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence, AAAI 17*, San Francisco, CA. AAAI. 10 pages, to appear (acceptance rate: 25 %)
- 183) Li, Y., Tao, J., Schuller, B., Shan, S., Jiang, D., and Jia, J. (2016). MEC 2016: The Multimodal Emotion Recognition Challenge of CCPR 2016. In *Proceedings of the 7th Chinese Conference on Pattern Recognition, CCPR*, pages 667–678, Chengdu, P.R. China. Springer
- 184) Marchi, E., Tonelli, D., Xu, X., Ringeval, F., Deng, J., Squartini, S., and Schuller, B. (2016c). Pairwise Decomposition with Deep Neural Networks and Multiscale Kernel Subspace Learning for Acoustic Scene Classification. In *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2016 IEEE AASP Challenge Workshop (DCASE 2016), satellite to EUSIPCO 2016*, pages 1–5, Budapest, Hungary. EUSIPCO, IEEE
- 185) Marchi, E., Eyben, F., Hagerer, G., and Schuller, B. W. (2016a). Real-time Tracking of Speakers' Emotions, States, and Traits on Mobile Platforms. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 1182–1183, San Francisco, CA. ISCA, ISCA. Show & Tell demonstration (50 % acceptance rate)
- 186) Marchi, E., Tonelli, D., Xu, X., Ringeval, F., Deng, J., Squartini, S., and Schuller, B. (2016d). The UP System for the 2016 DCASE Challenge using Deep Recurrent Neural Network and Multiscale Kernel Subspace Learning. In *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2016 IEEE AASP Challenge Workshop (DCASE 2016), satellite to EUSIPCO 2016*, Budapest, Hungary. EUSIPCO, IEEE. 1 page
- 187) Mousa, A. E.-D. and Schuller, B. (2016). Deep Bidirectional Long Short-Term Memory Recurrent Neural Networks for Grapheme-to-Phoneme Conversion utilizing Complex Many-to-Many Alignments. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 2836–2840, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 188) Pokorny, F. B., Marschik, P. B., Einspieler, C., and Schuller, B. W. (2016a). Does She Speak RTT? Towards an Earlier Identification of Rett Syndrome Through Intelligent Pre-linguistic Vocalisation Analysis. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 1953–1957, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 189) Pokorny, F. B., Peharz, R., Roth, W., Zohrer, M., Pernkopf, F., Marschik, P. B., and Schuller, B. W. (2016b). Manual Versus Automated: The Challenging Routine of Infant Vocalisation Segmentation in Home Videos to Study Neuro(mal)development. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 2997–3001, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 190) Qian, K., Janott, C., Zhang, Z., Heiser, C., and Schuller, B. (2016b). Wavelet Features for Classification of VOTE Snore Sounds. In *Proceedings 41st IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2016*, pages 221–225, Shanghai, P.R. China. IEEE, IEEE. (acceptance rate: 45 %, IF* 1.16 (2010))
- 191) Pantic, M., Evers, V., Deisenroth, M., Merino, L., and Schuller, B. (2016). Social and Affective Robotics Tutorial. In *Proceedings of the 24th ACM International Conference on Multimedia, MM 2016*, pages 1477–1478, Amsterdam, The Netherlands. ACM, ACM. (acceptance rate short paper: 30 %, IF* 2.45 (2010))
- 192) Pohjalainen, J., Ringeval, F., Zhang, Z., and Schuller, B. (2016). Spectral and Cepstral Audio Noise Reduction Techniques in Speech Emotion Recognition. In *Proceedings of the 24th ACM International Conference on Multimedia, MM 2016*, pages 670–674, Amsterdam, The Netherlands. ACM, ACM. (acceptance

- rate short paper: 30 %, IF* 2.45 (2010))
- 193) Ringeval, F., Marchi, E., Grossard, C., Xavier, J., Chetouani, M., Cohen, D., and Schuller, B. (2016a). Automatic Analysis of Typical and Atypical Encoding of Spontaneous Emotion in the Voice of Children. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 1210–1214, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 194) Rynkiewicz, A., Schuller, B., Marchi, E., Piana, S., Camurri, A., Lassalle, A., and Baron-Cohen, S. (2016b). An Investigation of the Female Camouflage Effect’ in Autism Using a New Computerized Test Showing Sex/Gender Differences during ADOS-2. In *Proceedings 15th Annual International Meeting For Autism Research (IMFAR 2016)*, Baltimore, MD. International Society for Autism Research (INSAR), INSAR. 1 page
- 195) Sagha, H., Deng, J., Gavryukova, M., Han, J., and Schuller, B. (2016a). Cross Lingual Speech Emotion Recognition using Canonical Correlation Analysis on Principal Component Subspace. In *Proceedings 41st IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2016*, pages 5800–5804, Shanghai, P.R. China. IEEE, IEEE. (acceptance rate: 45 %, IF* 1.16 (2010))
- 196) Sagha, H., Matejka, P., Gavryukova, M., Povolny, F., Marchi, E., and Schuller, B. (2016c). Enhancing multilingual recognition of emotion in speech by language identification. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 2949–2953, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 197) Sánchez-Rada, J., Schuller, B., Patti, V., Buitelaar, P., Vulcu, G., Burkhardt, F., Clavel, C., Petychakis, M., and Iglesias, C. A. (2016). Towards a Common Linked Data Model for Sentiment and Emotion Analysis. In Sánchez-Rada, J. and Schuller, B., editors, *Proceedings of the 6th International Workshop on Emotion and Sentiment Analysis (ESA 2016), satellite of the 10th Language Resources and Evaluation Conference (LREC 2016)*, pages 48–54, Portoroz, Slovenia. ELRA, ELRA. (42 % long paper acceptance rate)
- 198) Schmitt, M., Janott, C., Pandit, V., Qian, K., Heiser, C., Hemmert, W., and Schuller, B. (2016a). A Bag-of-Audio-Words Approach for Snore Sounds’ Excitation Localisation. In *Proceedings 14th ITG Conference on Speech Communication*, volume 267 of *ITG-Fachbericht*, pages 230–234, Paderborn, Germany. ITG/VDE, IEEE/VDE. nominated for best student paper award (4 nominations for 2 awards)
- 199) Schmitt, M., Ringeval, F., and Schuller, B. (2016c). At the Border of Acoustics and Linguistics: Bag-of-Audio-Words for the Recognition of Emotions in Speech. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 495–499, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 200) Schmitt, M., Marchi, E., Ringeval, F., and Schuller, B. (2016b). Towards Cross-lingual Automatic Diagnosis of Autism Spectrum Condition in Children’s Voices. In *Proceedings 14th ITG Conference on Speech Communication*, volume 267 of *ITG-Fachbericht*, pages 264–268, Paderborn, Germany. ITG/VDE, IEEE/VDE
- 201) Telespan, M. and Schuller, B. (2016). Audio Watermarking Based on Empirical Mode Decomposition and Beat Detection. In *Proceedings 41st IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2016*, pages 2124–2128, Shanghai, P.R. China. IEEE, IEEE. (acceptance rate: 45 %, IF* 1.16 (2010))
- 202) Trigeorgis, G., Ringeval, F., Brückner, R., Marchi, E., Nicolaou, M., Schuller, B., and Zafeiriou, S. (2016b). Adieu Features? End-to-End Speech Emotion Recognition using a Deep Convolutional Recurrent Network. In *Proceedings 41st IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2016*, pages 5200–5204, Shanghai, P.R. China. IEEE, IEEE. winner of the IEEE Spoken Language Processing Student Travel Grant 2016 (acceptance rate: 45 %, IF* 1.16 (2010))
- 203) Trigeorgis, G., Nicolaou, M. A., Zafeiriou, S., and Schuller, B. (2016a). Deep Canonical Time Warping. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, pages 5110–5118, Las Vegas, NV. IEEE, IEEE. (acceptance rate: ~25 %, IF* 5.97 (2010))
- 204) Valstar, M., Gratch, J., Schuller, B., Ringeval, F., Lalanne, D., Torres Torres, M., Scherer, S., Stratou, G., Cowie, R., and Pantic, M. (2016c). AVEC 2016: Depression, Mood, and Emotion Recognition Workshop and Challenge. In Valstar, M., Gratch, J., Schuller, B., Ringeval, F., Cowie, R., and Pantic, M., editors, *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge, AVEC’16, co-located with the 24th ACM International Conference on Multimedia, MM 2016*, pages 3–10, Amsterdam, The Netherlands. ACM, ACM
- 205) Valstar, M., Pelachaud, C., Heylen, D., Cafaro, A., Dermouche, S., Ghitulescu, A., André, E., Bauer, T., Wagner, J., Durieu, L., Aylett, M., Blaise, P., Coutinho, E., Schuller, B., Zhang, Y., Theune, M., and van Waterschoot, J. (2016d). Ask Alice; an Artificial Retrieval of Information Agent. In Morency, L.-P., Busso, C., and Pelachaud, C., editors, *Proceedings of the 18th ACM International Conference on Multimodal Interaction, ICMI*, pages 419–420, Tokyo, Japan. ACM, ACM. (IF* 1.77 (2010))
- 206) Valstar, M., Gratch, J., Schuller, B., Ringeval, F., Cowie, R., and Pantic, M. (2016b). Summary for AVEC 2016: Depression, Mood, and Emotion Recognition Workshop and Challenge. In *Proceedings of the 24th ACM International Conference on Multimedia, MM 2016*, pages 1483–1484, Amsterdam, The Netherlands. ACM, ACM. (acceptance rate short paper: 30 %, IF* 2.45 (2010))
- 207) Vlasenko, B., Schuller, B., and Wendemuth, A. (2016). Tendencies Regarding the Effect of Emotional Intensity in Inter Corpus Phoneme-Level Speech Emotion Modelling. In *Proceedings 2016 IEEE International Workshop on Machine Learning for Signal Processing, MLSP*, pages 1–6, Salerno, Italy. IEEE, IEEE
- 208) Wegener, R., Kohlschein, C., Jeschke, S., and Schuller, B. (2016). Automatic Detection of Textual Triggers of Reader Emotion in Short Stories. In Sánchez-Rada, J. and Schuller, B., editors, *Proceedings of the 6th International Workshop on Emotion and Sentiment Analysis (ESA 2016), satellite of the 10th Language Resources and Evaluation Conference (LREC 2016)*, pages 80–84, Portoroz, Slovenia. ELRA, ELRA. (74 % acceptance rate)
- 209) Weninger, F., Ringeval, F., Marchi, E., and Schuller, B. (2016a). Discriminatively trained recurrent neural networks for continuous dimensional emotion recognition from audio. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence, IJCAI 2016*, pages 2196–2202, New York City, NY. IJCAI/AAAI. (acceptance rate: 25 %)
- 210) Weninger, F., Ringeval, F., Marchi, E., and Schuller, B. (2016b). Discriminatively Trained Recurrent Neural Networks for Continuous Dimensional Emotion Recognition from Audio (Extended Abstract). In Friedrich, G., Helmert, M., and Wotawa, F., editors, *KI 2016: Advances in Artificial Intelligence 39th Annual German Conference on AI*, volume LNCS Volume 9904/2016, pages 310–315, Klagenfurt, Austria. GfI / ÖGAI, Springer. (acceptance rate: 45 %)
- 211) Xu, X., Deng, J., Gavryukova, M., Zhang, Z., Zhao, L., and Schuller, B. (2016). Multiscale Kernel Locally Penalised Discriminant Analysis Exemplified by Emotion Recognition in Speech. In Morency, L.-P., Busso, C., and Pelachaud, C., editors, *Proceedings of the 18th ACM International Conference on Multimodal Interaction, ICMI*, pages 233–237, Tokyo, Japan. ACM, ACM. (IF* 1.77 (2010))
- 212) Zhang, Z., Ringeval, F., Dong, B., Coutinho, E., Marchi, E., and Schuller, B. (2016d). Enhanced Semi-Supervised Learning for Multimodal Emotion Recognition. In *Proceedings 41st IEEE*

- International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2016*, pages 5185–5189, Shanghai, P.R. China. IEEE, IEEE. (acceptance rate: 45 %, IF* 1.16 (2010))
- 213) Zhang, Z., Ringeval, F., Han, J., Deng, J., Marchi, E., and Schuller, B. (2016e). Facing Realism in Spontaneous Emotion Recognition from Speech: Feature Enhancement by Autoencoder with LSTM Neural Networks. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 3593–3597, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 214) Zhang, Y., Weninger, F., Batliner, A., Hönl, F., and Schuller, B. (2016a). Language Proficiency Assessment of English L2 Speakers Based on Joint Analysis of Prosody and Native Language. In Morency, L.-P., Busso, C., and Pelachaud, C., editors, *Proceedings of the 18th ACM International Conference on Multimodal Interaction, ICMI*, pages 274–278, Tokyo, Japan. ACM, ACM. (IF* 1.77 (2010))
- 215) Zhang, Y., Weninger, F., Ren, Z., and Schuller, B. (2016b). Sincerity and Deception in Speech: Two Sides of the Same Coin? A Transfer- and Multi-Task Learning Perspective. In *Proceedings INTERSPEECH 2016, 17th Annual Conference of the International Speech Communication Association*, pages 2041–2045, San Francisco, CA. ISCA, ISCA. (50 % acceptance rate)
- 216) Zhang, Y., Zhou, Y., Shen, J., and Schuller, B. (2016c). Semi-autonomous Data Enrichment Based on Cross-task Labelling of Missing Targets for Holistic Speech Analysis. In *Proceedings 41st IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2016*, pages 6090–6094, Shanghai, P.R. China. IEEE, IEEE. (acceptance rate: 45 %, IF* 1.16 (2010))
- 217) Zhang, Y. and Schuller, B. (2016). Towards Human-Like Holistic Machine Perception of Speaker States and Traits. In *Proceedings of the Human-Like Computing Machine Intelligence Workshop, MI20-HLC*, Windsor, U.K. Springer. 3 pages
- 218) Deng, J., Xu, X., Zhang, Z., Frühholz, S., Grandjean, D., and Schuller, B. (2016b). Fisher Kernels on Phase-based Features for Speech Emotion Recognition. In *Proceedings of the Seventh International Workshop on Spoken Dialogue Systems (IWSDS)*, Saariselkä, Finland. Springer. 6 pages
- 219) Schuller, B., Vlasenko, B., Eyben, F., Wöllmer, M., Stuhlsatz, A., Wendemuth, A., and Rigoll, G. (2015g). Cross-Corpus Acoustic Emotion Recognition: Variances and Strategies (Extended Abstract). In *Proc. 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 470–476, Xi'an, P.R. China. AAAC, IEEE. invited for the Special Session on Most Influential Articles in IEEE Transactions on Affective Computing
- 220) Schuller, B., Marchi, E., Baron-Cohen, S., Lassalle, A., O'Reilly, H., Pigat, D., Robinson, P., Davies, I., Baltrusaitis, T., Mahmoud, M., Golan, O., Friedenson, S., Tal, S., Newman, S., Meir, N., Shillo, R., Camurri, A., Piana, S., Staglianò, A., Bölte, S., Lundqvist, D., Berggren, S., Baranger, A., Sullings, N., Sezgin, M., Alyuz, N., Rynkiewicz, A., Ptaszek, K., and Ligmann, K. (2015b). Recent developments and results of ASC-Inclusion: An Integrated Internet-Based Environment for Social Inclusion of Children with Autism Spectrum Conditions. In Paletta, L., Schuller, B., Robinson, P., and Sabouret, N., editors, *Proceedings of the of the 3rd International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2015) as part of the 20th ACM International Conference on Intelligent User Interfaces, IUI 2015*, Atlanta, GA. ACM, ACM. 9 pages, best paper award (long talk acceptance rate: 36 %)
- 221) Schuller, B. (2015h). Speech Analysis in the Big Data Era. In *Text, Speech, and Dialogue – Proceedings of the 18th International Conference on Text, Speech and Dialogue, TSD 2015*, volume 9302 of *Lecture Notes in Computer Science (LNCS)*, pages 3–11. Springer. satellite event of INTERSPEECH 2015, invited contribution (acceptance rate: 50 %)
- 222) Schuller, B., Steidl, S., Batliner, A., Hantke, S., Hönl, F., Orozco-Arroyave, J. R., Nöth, E., Zhang, Y., and Weninger, F. (2015d). The INTERSPEECH 2015 Computational Paralinguistics Challenge: Degree of Nativeness, Parkinson's & Eating Condition. In *Proceedings INTERSPEECH 2015, 16th Annual Conference of the International Speech Communication Association*, pages 478–482, Dresden, Germany. ISCA, ISCA. (51 % acceptance rate)
- 223) Azais, L., Payan, A., Sun, T., Vidal, G., Zhang, T., Coutinho, E., Eyben, F., and Schuller, B. (2015). Does my Speech Rock? Automatic Assessment of Public Speaking Skills. In *Proceedings INTERSPEECH 2015, 16th Annual Conference of the International Speech Communication Association*, pages 2519–2523, Dresden, Germany. ISCA, ISCA. (51 % acceptance rate)
- 224) Coutinho, E., Trigeorgis, G., Zafeiriou, S., and Schuller, B. (2015). Automatically Estimating Emotion in Music with Deep Long-Short Term Memory Recurrent Neural Networks. In Larson, M., Ionescu, B., Sjöberg, M., Anguera, X., Poignant, J., Riegler, M., Eskevich, M., Hauff, C., Sutcliffe, R., Jones, G. J., Yang, Y.-H., Soleymani, M., and Papadopoulos, S., editors, *Proceedings of the MediaEval 2015 Multimedia Benchmark Workshop, satellite of Interspeech 2015*, volume 1436, Wurzen, Germany. CEUR. 3 pages
- 225) Deng, J., Zhang, Z., Eyben, F., and Schuller, B. (2015). Autoencoder-based Unsupervised Domain Adaptation for Speech Emotion Recognition. In *Proceedings 40th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2015*, pages 1068–1072, Brisbane, Australia. IEEE, IEEE. (IF* 1.16 (2010))
- 226) Eyben, F., Huber, B., Marchi, E., Schuller, D., and Schuller, B. (2015a). Robust Real-time Affect Analysis and Speaker Characterisation on Mobile Devices. In *Proc. 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 778–780, Xi'an, P.R. China. AAAC, IEEE. (acceptance rate: 55 %)
- 227) Feraru, S., Schuller, D., and Schuller, B. (2015). Cross-Language Acoustic Emotion Recognition: An Overview and Some Tendencies. In *Proc. 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 125–131, Xi'an, P.R. China. AAAC, IEEE. (acceptance rate oral: 28 %)
- 228) Gentsch, K., Coutinho, E., Eyben, F., Schuller, B., and Scherer, K. R. (2015). Classifying Emotion-Antecedent Appraisal in Brain Activity using Machine Learning Methods. In *Proceedings of the International Society for Research on Emotions Conference (ISRE 2015)*, Geneva, Switzerland. ISRE, ISRE. 1 page
- 229) Hantke, S., Eyben, F., Appel, T., and Schuller, B. (2015). iHEARU-PLAY: Introducing a game for crowdsourced data collection for affective computing. In *Proc. 1st International Workshop on Automatic Sentiment Analysis in the Wild (WASA 2015) held in conjunction with the 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 891–897, Xi'an, P.R. China. AAAC, IEEE. (acceptance rate: 60 %)
- 230) Marchi, E., Vesperini, F., Eyben, F., Squartini, S., and Schuller, B. (2015c). A Novel Approach for Automatic Acoustic Novelty Detection Using a Denoising Autoencoder with Bidirectional LSTM Neural Networks. In *Proceedings 40th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2015*, pages 1996–2000, Brisbane, Australia. IEEE, IEEE. (IF* 1.16 (2010))
- 231) Marchi, E., Vesperini, F., Weninger, F., Eyben, F., Squartini, S., and Schuller, B. (2015d). Non-Linear Prediction with LSTM Recurrent Neural Networks for Acoustic Novelty Detection. In *Proceedings 2015 International Joint Conference on Neural Networks (IJCNN)*, Killarney, Ireland. IEEE, IEEE

- 232) Marchi, E., Schuller, B., Baron-Cohen, S., Golan, O., Bölte, S., Arora, P., and Hüb-Umbach, R. (2015a). Typicality and Emotion in the Voice of Children with Autism Spectrum Condition: Evidence Across Three Languages. In *Proceedings INTER-SPEECH 2015, 16th Annual Conference of the International Speech Communication Association*, pages 115–119, Dresden, Germany. ISCA, ISCA. (51 % acceptance rate)
- 233) Marchi, E., Schuller, B., Baron-Cohen, S., Lassalle, A., O'Reilly, H., Pigat, D., Golan, O., Friedenson, S., Tal, S., Bölte, S., Berggren, S., Lundqvist, D., and Elfström, M. S. (2015b). Voice Emotion Games: Language and Emotion in the Voice of Children with Autism Spectrum Condition. In Paletta, L., Schuller, B., Robinson, P., and Sabouret, N., editors, *Proceedings of the 3rd International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2015) as part of the 20th ACM International Conference on Intelligent User Interfaces, IUI 2015*, Atlanta, GA. ACM, ACM. 9 pages (long talk acceptance rate: 36 %)
- 234) Metallinou, A., Wöllmer, M., Katsamanis, A., Eyben, F., Schuller, B., and Narayanan, S. (2015). Context-Sensitive Learning for Enhanced Audiovisual Emotion Classification (Extended Abstract). In *Proc. 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 463–469, Xi'an, P.R. China. AAAC, IEEE. invited for the Special Session on Most Influential Articles in IEEE Transactions on Affective Computing
- 235) Newman, S., Golan, O., Baron-Cohen, S., Bölte, S., Rynkiewicz, A., Baranger, A., Schuller, B., Robinson, P., Camurri, A., Sezgin, M., Meir-Goren, N., Tal, S., Friedenson-Hayo, S., Lassalle, A., Berggren, S., Sullings, N., Pigat, D., Ptaszek, K., Marchi, E., Piana, S., and Baltrusaitis, T. (2015). ASC-Inclusion - a Virtual World Teaching Children with ASC about Emotions. In *Proceedings 14th Annual International Meeting For Autism Research (IMFAR 2015)*, Salt Lake City, UT. International Society for Autism Research (INSAR), INSAR. 1 page
- 236) Pohl, P. and Schuller, B. (2015). Digital Analysis of Vocal Operants. In *Proceedings 2015 Meeting of the Experimental Analysis of Behaviour Group (EABG)*, London, UK. EABG, EABG. 1 page, to appear
- 237) Pokorny, F., Graf, F., Pernkopf, F., and Schuller, B. (2015). Detection of Negative Emotions in Speech Signals Using Bags-of-Audio-Words. In *Proc. 1st International Workshop on Automatic Sentiment Analysis in the Wild (WASA 2015) held in conjunction with the 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 879–884, Xi'an, P.R. China. AAAC, IEEE. (acceptance rate: 60 %)
- 238) Qian, K., Zhang, Z., Ringeval, F., and Schuller, B. (2015). Bird Sounds Classification by Large Scale Acoustic Features and Extreme Learning Machine. In *Proceedings 3rd IEEE Global Conference on Signal and Information Processing, GlobalSIP, Machine Learning Applications in Speech Processing Symposium*, pages 1317–1321, Orlando, FL. IEEE, IEEE. (acceptance rate: 45 %)
- 239) Ringeval, F., Marchi, E., Méhu, M., Scherer, K., and Schuller, B. (2015b). Face Reading from Speech - Predicting Facial Action Units from Audio Cues. In *Proceedings INTERSPEECH 2015, 16th Annual Conference of the International Speech Communication Association*, pages 1977–1981, Dresden, Germany. ISCA, ISCA. (51 % acceptance rate)
- 240) Ringeval, F., Schuller, B., Valstar, M., Cowie, R., and Pantic, M. (2015d). AVEC 2015: The 5th International Audio/Visual Emotion Challenge and Workshop. In *Proceedings of the 23rd ACM International Conference on Multimedia, MM 2015*, pages 1335–1336, Brisbane, Australia. ACM, ACM. (25 % acceptance rate)
- 241) Ringeval, F., Schuller, B., Valstar, M., Jaiswal, S., Marchi, E., Lalanne, D., Cowie, R., and Pantic, M. (2015f). AV+EC 2015 - The First Affect Recognition Challenge Bridging Across Audio, Video, and Physiological Data. In Ringeval, F., Schuller, B., Valstar, M., Cowie, R., and Pantic, M., editors, *Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge, AVEC'15, co-located with the 23rd ACM International Conference on Multimedia, MM 2015*, pages 3–8, Brisbane, Australia. ACM, ACM
- 242) Sabouret, N., Schuller, B., Paletta, L., Marchi, E., Jones, H., and Youssef, A. B. (2015). Intelligent User Interfaces in Digital Games for Empowerment and Inclusion. In *Proceedings of the 12th International Conference on Advancement in Computer Entertainment Technology, ACE 2015*, Iskandar, Malaysia. ACM, ACM. 8 pages, Gold Paper Award
- 243) Sagha, H., Coutinho, E., and Schuller, B. (2015). The importance of individual differences in the prediction of emotions induced by music. In Ringeval, F., Schuller, B., Valstar, M., Cowie, R., and Pantic, M., editors, *Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge, AVEC'15, co-located with the 23rd ACM International Conference on Multimedia, MM 2015*, pages 57–63, Brisbane, Australia. ACM, ACM. (acceptance rate: 60 %)
- 244) Schröder, M., Bevacqua, E., Cowie, R., Eyben, F., Gunes, H., Heylen, D., ter Maat, M., McKeown, G., Pammi, S., Pantic, M., Pelachaud, C., Schuller, B., de Sevin, E., Valstar, M., and Wöllmer, M. (2015). Building Autonomous Sensitive Artificial Listeners (Extended Abstract). In *Proc. 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, pages 456–462, Xi'an, P.R. China. AAAC, IEEE. invited for the Special Session on Most Influential Articles in IEEE Transactions on Affective Computing
- 245) Trigeorgis, G., Coutinho, E., Ringeval, F., Marchi, E., Zafeiriou, S., and Schuller, B. (2015b). The ICL-TUM-PASSAU approach for the MediaEval 2015 "Affective Impact of Movies" Task. In Larson, M., Ionescu, B., Sjberg, M., Anguera, X., Poignant, J., Riegler, M., Eskevich, M., Hauff, C., Sutcliffe, R., Jones, G. J., Yang, Y.-H., Soleymani, M., and Papadopoulos, S., editors, *Proceedings of the MediaEval 2015 Multimedia Benchmark Workshop, satellite of Interspeech 2015*, volume 1436, Wurzen, Germany. CEUR. 3 pages, best result
- 246) Trigeorgis, G., Nicolaou, M. A., Zafeiriou, S., and Schuller, B. (2015c). Towards Deep Alignment of Multimodal Data. In *Proceedings 2015 Multimodal Machine Learning Workshop held in conjunction with NIPS 2015 (MMML@NIPS)*, Montral, QC. NIPS, NIPS. 4 pages
- 247) Weninger, F., Erdogan, H., Watanabe, S., Vincent, E., Le Roux, J., Hershey, J. R., and Schuller, B. (2015b). Speech Enhancement with LSTM Recurrent Neural Networks and its Application to Noise-Robust ASR. In Vincent, E., Yeredor, A., Koldovsk, Z., and Tichavsk, P., editors, *Latent Variable Analysis and Signal Separation – Proceedings 12th International Conference on Latent Variable Analysis and Signal Separation, LVA/ICA 2015*, volume 9237 of *Lecture Notes in Computer Science*, pages 91–99, Liberec, Czech Republic. Springer
- 248) Xu, X., Deng, J., Zheng, W., Zhao, L., and Schuller, B. (2015). Dimensionality Reduction for Speech Emotion Features by Multiscale Kernels. In *Proceedings INTERSPEECH 2015, 16th Annual Conference of the International Speech Communication Association*, pages 1532–1536, Dresden, Germany. ISCA, ISCA. (51 % acceptance rate)
- 249) Zhang, Y., Coutinho, E., Zhang, Z., Quan, C., and Schuller, B. (2015b). Agreement-based Dynamic Active Learning with Least and Medium Certainty Query Strategy. In Krishnamurthy, A., Ramdas, A., Balcan, N., and Singh, A., editors, *Proceedings Advances in Active Learning : Bridging Theory and Practice Workshop held in conjunction with the 32nd International Conference on Machine Learning, ICML 2015*, Lille, France. International Machine Learning Society, IMLS. 5 pages
- 250) Zhang, Y., Coutinho, E., Zhang, Z., Adam, M., and Schuller, B. (2015a). On Rater Reliability and Agreement Based Dynamic Active Learning. In *Proc. 6th biannual Conference on Affective*

- Computing and Intelligent Interaction (ACII 2015)*, pages 70–76, Xi'an, P.R. China. AAAC, IEEE. (acceptance rate oral: 28 %)
- 251) Zhang, Y., Coutinho, E., Zhang, Z., Quan, C., and Schuller, B. (2015c). Dynamic Active Learning Based on Agreement and Applied to Emotion Recognition in Spoken Interactions. In *Proceedings 17th International Conference on Multimodal Interaction, ICMI 2015*, pages 275–278, Seattle, WA. ACM, ACM. (IF* 1.77 (2010))
- 252) Schuller, B., Zhang, Y., Eyben, F., and Weninger, F. (2014f). Intelligent systems' Holistic Evolving Analysis of Real-life Universal speaker characteristics. In Schuller, B., Buitelaar, P., Devillers, L., Pelachaud, C., Declerck, T., Batliner, A., Rosso, P., and Gaines, S., editors, *Proceedings of the 5th International Workshop on Emotion Social Signals, Sentiment & Linked Open Data (ES³LOD 2014), satellite of the 9th Language Resources and Evaluation Conference (LREC 2014)*, pages 14–20, Reykjavik, Iceland. ELRA, ELRA
- 253) Schuller, B., Steidl, S., Batliner, A., Epps, J., Eyben, F., Ringeval, F., Marchi, E., and Zhang, Y. (2014d). The INTERSPEECH 2014 Computational Paralinguistics Challenge: Cognitive & Physical Load. In *Proceedings INTERSPEECH 2014, 15th Annual Conference of the International Speech Communication Association*, Singapore, Singapore. ISCA, ISCA. 5 pages (52 % acceptance rate)
- 254) Schuller, B., Friedmann, F., and Eyben, F. (2014b). The Munich BioVoice Corpus: Effects of Physical Exercising, Heart Rate, and Skin Conductance on Human Speech Production. In *Proceedings 9th Language Resources and Evaluation Conference (LREC 2014)*, pages 1506–1510, Reykjavik, Iceland. ELRA, ELRA
- 255) Schuller, B., Marchi, E., Baron-Cohen, S., O'Reilly, H., Pigat, D., Robinson, P., Davies, I., Golan, O., Fridenson, S., Tal, S., Newman, S., Meir, N., Shillo, R., Camurri, A., Piana, S., Staglianò, A., Bölte, S., Lundqvist, D., Berggren, S., Baranger, A., and Sullings, N. (2014c). The state of play of ASC-Inclusion: An Integrated Internet-Based Environment for Social Inclusion of Children with Autism Spectrum Conditions. In Paletta, L., Schuller, B., Robinson, P., and Sabouret, N., editors, *Proceedings 2nd International Workshop on Digital Games for Empowerment and Inclusion (IDGEI 2014)*, Haifa, Israel. ACM, ACM. 8 pages, held in conjunction with the 19th International Conference on Intelligent User Interfaces (IUI 2014)
- 256) Batliner, A. and Schuller, B. (2014). More Than Fifty Years of Speech Processing – The Rise of Computational Paralinguistics and Ethical Demands. In *Proceedings ETHICOMP 2014*, Paris, France. Commission de réflexion sur l'Ethique de la Recherche en sciences et technologies du Numérique d'Allistene, CERNA
- 257) Brückner, R. and Schuller, B. (2014). Social Signal Classification Using Deep BLSTM Recurrent Neural Networks. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 4856–4860, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 258) Celiktutan, O., Eyben, F., Sariyanidi, E., Gunes, H., and Schuller, B. (2014a). MAPTRAITS 2014: The First Audio/Visual Mapping Personality Traits Challenge. In *Proceedings of the Personality Mapping Challenge & Workshop (MAPTRAITS 2014), Satellite of the 16th ACM International Conference on Multimodal Interaction (ICMI 2014)*, pages 3–9, Istanbul, Turkey. ACM, ACM
- 259) Celiktutan, O., Eyben, F., Sariyanidi, E., Gunes, H., and Schuller, B. (2014b). MAPTRAITS 2014: The First Audio/Visual Mapping Personality Traits Challenge – An Introduction. In *Proceedings of the Personality Mapping Challenge & Workshop (MAPTRAITS 2014), Satellite of the 16th ACM International Conference on Multimodal Interaction (ICMI 2014)*, pages 529–530, Istanbul, Turkey. ACM, ACM
- 260) Coutinho, E., Weninger, F., Scherer, K., and Schuller, B. (2014b). The Munich LSTM-RNN Approach to the MediaEval 2014 “Emotion in Music” Task. In Larson, M., Ionescu, B., Anguera, X., Eskevich, M., Korshunov, P., Schedl, M., Soleymani, M., Petkos, G., Sutcliffe, R., Choi, J., and Jones, G. J., editors, *Proceedings of the MediaEval 2014 Multimedia Benchmark Workshop*, Barcelona, Spain. CEUR. 2 pages, best result
- 261) Coutinho, E., Deng, J., and Schuller, B. (2014a). Transfer Learning Emotion Manifestation Across Music and Speech. In *Proceedings 2014 International Joint Conference on Neural Networks (IJCNN) as part of the IEEE World Congress on Computational Intelligence (IEEE WCCI)*, pages 3592–3598, Beijing, China. IEEE, IEEE. (acceptance rate: 30 %)
- 262) Deng, J., Xia, R., Zhang, Z., Liu, Y., and Schuller, B. (2014a). Introducing Shared-Hidden-Layer Autoencoders for Transfer Learning and their Application in Acoustic Emotion Recognition. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 4851–4855, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 263) Deng, J., Zhang, Z., and Schuller, B. (2014c). Linked Source and Target Domain Subspace Feature Transfer Learning – Exemplified by Speech Emotion Recognition. In *Proceedings 22nd International Conference on Pattern Recognition (ICPR 2014)*, pages 761–766, Stockholm, Sweden. IAPR, IAPR. acceptance rate: 56 %
- 264) Geiger, J. T., Kneissl, M., Schuller, B., and Rigoll, G. (2014c). Acoustic Gait-based Person Identification using Hidden Markov Models. In *Proceedings of the Personality Mapping Challenge & Workshop (MAPTRAITS 2014), Satellite of the 16th ACM International Conference on Multimodal Interaction (ICMI 2014)*, pages 25–30, Istanbul, Turkey. ACM, ACM
- 265) Geiger, J. T., Gemmeke, J. F., Schuller, B., and Rigoll, G. (2014b). Investigating NMF Speech Enhancement for Neural Network based Acoustic Models. In *Proceedings INTERSPEECH 2014, 15th Annual Conference of the International Speech Communication Association*, Singapore, Singapore. ISCA, ISCA. 5 pages, (52 % acceptance rate)
- 266) Geiger, J. T., Weninger, F., Gemmeke, J. F., Wöllmer, M., Schuller, B., and Rigoll, G. (2014e). Memory-Enhanced Neural Networks and NMF for Robust ASR. In *Proceedings 2nd IEEE Global Conference on Signal and Information Processing, GlobalSIP, Machine Learning Applications in Speech Processing Symposium*, Atlanta, GA. IEEE, IEEE. 10 pages (acceptance rate: 45 %)
- 267) Geiger, J. T., Zhang, Z., Weninger, F., Schuller, B., and Rigoll, G. (2014g). Robust Speech Recognition using Long Short-Term Memory Recurrent Neural Networks for Hybrid Acoustic Modelling. In *Proceedings INTERSPEECH 2014, 15th Annual Conference of the International Speech Communication Association*, Singapore, Singapore. ISCA, ISCA. 5 pages, (52 % acceptance rate)
- 268) Geiger, J., Marchi, E., Weninger, F., Schuller, B., and Rigoll, G. (2014a). The TUM system for the REVERB Challenge: Recognition of Reverberated Speech using Multi-Channel Correlation Shaping Dereverberation and BLSTM Recurrent Neural Networks. In *Proceedings REVERB Workshop, held in conjunction with ICASSP 2014 and HSCMA 2014*, pages 1–8, Florence, Italy. IEEE, IEEE
- 269) Geiger, J. T., Zhang, B., Schuller, B., and Rigoll, G. (2014f). On the Influence of Alcohol Intoxication on Speaker Recognition. In *Proceedings AES 53rd International Conference on Semantic Audio*, pages 1–7, London, UK. AES, Audio Engineering Society
- 270) Hartmann, K., Böck, R., and Schuller, B. (2014a). ERM4HCI 2014 – The 2nd Workshop on Emotion Representation and Modelling in Human-Computer-Interaction-Systems. In Hartmann, K., Böck, R., and Schuller, B., editors, *Proceedings of the 2nd Workshop on Emotion representation and modelling in*

- Human-Computer-Interaction-Systems, ERM4HCI 2014*, pages 525–526, Istanbul, Turkey. ACM, ACM. held in conjunction with the 16th ACM International Conference on Multimodal Interaction, ICMI 2014
- 271) Kaya, H., Eyben, F., Salah, A. A., and Schuller, B. (2014). CCA Based Feature Selection with Application to Continuous Depression Recognition from Acoustic Speech Features. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 3757–3761, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 272) Kirst, C., Weninger, F., Joder, C., Grosche, P., Geiger, J., and Schuller, B. (2014). On-line NMF-based Stereo Up-Mixing of Speech Improves Perceived Reduction of Non-Stationary Noise. In Brandenburg, K. and Sandler, M., editors, *Proceedings AES 53rd International Conference on Semantic Audio*, pages 1–7, London, UK. AES, Audio Engineering Society. Best Student Paper Award
- 273) Marchi, E., Ferroni, G., Eyben, F., Gabrielli, L., Squartini, S., and Schuller, B. (2014a). Multi-resolution Linear Prediction Based Features for Audio Onset Detection with Bidirectional LSTM Neural Networks. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 2183–2187, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 274) Marchi, E., Ferroni, G., Eyben, F., Squartini, S., and Schuller, B. (2014b). Audio Onset Detection: A Wavelet Packet Based Approach with Recurrent Neural Networks. In *Proceedings 2014 International Joint Conference on Neural Networks (IJCNN) as part of the IEEE World Congress on Computational Intelligence (IEEE WCCI)*, pages 3585–3591, Beijing, China. IEEE, IEEE. (acceptance rate: 30 %)
- 275) Newman, S., Golan, O., Baron-Cohen, S., Bölte, S., Baranger, A., Schuller, B., Robinson, P., Camurri, A., Meir-Goren, N., Skurnik, M., Fridenson, S., Tal, S., Eshchar, E., O’Reilly, H., Pigat, D., Berggren, S., Lundqvist, D., Sullings, N., Davies, I., and Piana, S. (2014). ASC-Inclusion – Interactive Software to Help Children with ASC Understand and Express Emotions. In *Proceedings 13th Annual International Meeting For Autism Research (IMFAR 2014)*, Atlanta, GA. International Society for Autism Research (INSAR), INSAR. 1 page
- 276) Ringeval, F., Amiriparian, S., Eyben, F., Scherer, K., and Schuller, B. (2014). Emotion Recognition in the Wild: Incorporating Voice and Lip Activity in Multimodal Decision-Level Fusion. In *Proceedings of the ICMI 2014 EmotiW – Emotion Recognition In The Wild Challenge and Workshop (EmotiW 2014)*, *Satellite of the 16th ACM International Conference on Multimodal Interaction (ICMI 2014)*, pages 473–480, Istanbul, Turkey. ACM, ACM
- 277) Soleymani, M., Aljanaki, A., Yang, Y.-H., Caro, M. N., Eyben, F., Markov, K., Schuller, B., Veltkamp, R., Weninger, F., and Wiering, F. (2014). Emotional Analysis of Music: A Comparison of Methods. In *Proceedings of the 22nd ACM International Conference on Multimedia, MM 2014*, pages 1161–1164, Orlando, FL. ACM, ACM. 4 pages
- 278) Trigeorgis, G., Bousmalis, K., Zafeiriou, S., and Schuller, B. (2014). A Deep Semi-NMF Model for Learning Hidden Representations. In Xing, E. P. and Jebara, T., editors, *Proceedings 31st International Conference on Machine Learning, ICML 2014*, volume 32, Beijing, China. International Machine Learning Society, IMLS. 9 pages (acceptance rate: 25 %)
- 279) Weninger, F., Hershey, J. R., Le Roux, J., and Schuller, B. (2014b). Discriminatively Trained Recurrent Neural Networks for Single-Channel Speech Separation. In *Proceedings 2nd IEEE Global Conference on Signal and Information Processing, GlobalSIP, Machine Learning Applications in Speech Processing Symposium*, pages 577–581, Atlanta, GA. IEEE, IEEE. (acceptance rate: 45 %)
- 280) Weninger, F., Watanabe, S., Le Roux, J., Hershey, J. R., Tachioka, Y., Geiger, J., Schuller, B., and Rigoll, G. (2014d). The MERL/MELCO/TUM system for the REVERB Challenge using Deep Recurrent Neural Network Feature Enhancement. In *Proceedings REVERB Workshop, held in conjunction with ICASSP 2014 and HSCMA 2014*, pages 1–8, Florence, Italy. IEEE, IEEE. second best result
- 281) Zhang, Z., Eyben, F., Deng, J., and Schuller, B. (2014b). An Agreement and Sparseness-based Learning Instance Selection and its Application to Subjective Speech Phenomena. In Schuller, B., Buitelaar, P., Devillers, L., Pelachaud, C., Declerck, T., Batliner, A., Rosso, P., and Gaines, S., editors, *Proceedings of the 5th International Workshop on Emotion Social Signals, Sentiment & Linked Open Data (ES³LOD 2014)*, *satellite of the 9th Language Resources and Evaluation Conference (LREC 2014)*, pages 21–26, Reykjavik, Iceland. ELRA, ELRA
- 282) Valstar, M., Schuller, B., Smith, K., Almaev, T., Eyben, F., Krajewski, J., Cowie, R., and Pantic, M. (2014b). AVEC 2014 – The Three Dimensional Affect and Depression Challenge. In *Proceedings of the 4th ACM international workshop on Audio/Visual Emotion Challenge*, Orlando, FL. ACM, ACM. 9 pages
- 283) Weninger, F. J., Watanabe, S., Tachioka, Y., and Schuller, B. (2014g). Deep Recurrent De-Noiseing Auto-Encoder and blind de-reverberation for reverberated speech recognition. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 4656–4660, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 284) Weninger, F. J., Eyben, F., and Schuller, B. (2014e). On-Line Continuous-Time Music Mood Regression with Deep Recurrent Neural Networks. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 5449–5453, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 285) Weninger, F. J., Eyben, F., and Schuller, B. (2014f). Single-Channel Speech Separation With Memory-Enhanced Recurrent Neural Networks. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 3737–3741, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 286) Xia, R., Deng, J., Schuller, B., and Liu, Y. (2014). Modeling Gender Information for Emotion Recognition Using Denoising Autoencoders. In *Proceedings 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pages 990–994, Florence, Italy. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 287) Schuller, B., Marchi, E., Baron-Cohen, S., O’Reilly, H., Robinson, P., Davies, I., Golan, O., Fridenson, S., Tal, S., Newman, S., Meir, N., Shillo, R., Camurri, A., Piana, S., Bölte, S., Lundqvist, D., Berggren, S., Baranger, A., and Sullings, N. (2013c). ASC-Inclusion: Interactive Emotion Games for Social Inclusion of Children with Autism Spectrum Conditions. In Schuller, B., Paletta, L., and Sabouret, N., editors, *Proceedings 1st International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2013) held in conjunction with the 8th Foundations of Digital Games 2013 (FDG)*, Chania, Greece. ACM, SASDG. 8 pages (acceptance rate: 69 %)
- 288) Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Scherer, K., Ringeval, F., Chetouani, M., Weninger, F., Eyben, F., Marchi, E., Mortillaro, M., Salamin, H., Polychroniou, A., Valente, F., and Kim, S. (2013i). The INTERSPEECH 2013 Computational Paralinguistics Challenge: Social Signals, Conflict, Emotion, Autism. In *Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association*, pages 148–152, Lyon, France. ISCA, ISCA. (acceptance rate: 52 %, IF* 1.05 (2010), > 200 citations)
- 289) Schuller, B., Friedmann, F., and Eyben, F. (2013b). Automatic Recognition of Physiological Parameters in the Human Voice: Heart Rate and Skin Conductance. In *Proceedings 38th IEEE*

- International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 7219–7223, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 290) Schuller, B., Pokorny, F., Ladstätter, S., Fellner, M., Graf, F., and Paletta, L. (2013e). Acoustic Geo-Sensing: Recognising Cyclists' Route, Route Direction, and Route Progress from Cell-Phone Audio. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 453–457, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 291) Brückner, R. and Schuller, B. (2013). Hierarchical Neural Networks and Enhanced Class Posteriors for Social Signal Classification. In *Proceedings 13th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2013*, pages 362–367, Olomouc, Czech Republic. IEEE, IEEE. 6 pages (acceptance rate: 47 %)
- 292) Deng, J., Zhang, Z., Marchi, E., and Schuller, B. (2013). Sparse Autoencoder-based Feature Transfer Learning for Speech Emotion Recognition. In *Proc. 5th biannual Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII 2013)*, pages 511–516, Geneva, Switzerland. HUMAINE Association, IEEE. (acceptance rate oral: 31 %)
- 293) Dunwell, I., Lameris, P., Stewart, C., Petridis, P., Arnab, S., Hendrix, M., de Freitas, S., Gaved, M., Schuller, B., and Paletta, L. (2013). Developing a Digital Game to Support Cultural Learning amongst Immigrants. In Schuller, B., Paletta, L., and Sabouret, N., editors, *Proceedings 1st International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2013) held in conjunction with the 8th Foundations of Digital Games 2013 (FDG)*, Chania, Greece. ACM, SASDG. 8 pages (acceptance rate: 69 %)
- 294) Eyben, F., Weninger, F., Paletta, L., and Schuller, B. (2013d). The acoustics of eye contact - Detecting visual attention from conversational audio cues. In *Proceedings 6th Workshop on Eye Gaze in Intelligent Human Machine Interaction: Gaze in Multimodal Interaction (GAZEIN 2013), held in conjunction with the 15th International Conference on Multimodal Interaction, ICMI 2013*, pages 7–12, Sydney, Australia. ACM, ACM. (acceptance rate: 38 %, IF* 1.77 (2010))
- 295) Eyben, F., Weninger, F., and Schuller, B. (2013e). Affect recognition in real-life acoustic conditions - A new perspective on feature selection. In *Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association*, pages 2044–2048, Lyon, France. ISCA, ISCA. (acceptance rate: 52 %, IF* 1.05 (2010))
- 296) Eyben, F., Weninger, F., Groß, F., and Schuller, B. (2013a). Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor. In *Proceedings of the 21st ACM International Conference on Multimedia, MM 2013*, pages 835–838, Barcelona, Spain. ACM, ACM. (Honorable Mention (2nd place) in the ACM MM 2013 Open-source Software Competition, acceptance rate: 28 %, > 200 citations)
- 297) Eyben, F., Weninger, F., Squartini, S., and Schuller, B. (2013f). Real-life Voice Activity Detection with LSTM Recurrent Neural Networks and an Application to Hollywood Movies. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 483–487, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 298) Eyben, F., Weninger, F., Marchi, E., and Schuller, B. (2013c). Likability of human voices: A feature analysis and a neural network regression approach to automatic likability estimation. In *Proceedings 14th International Workshop on Image and Audio Analysis for Multimedia Interactive Services, WIAMIS 2013*, Paris, France. IEEE, IEEE. Special Session on Social Stance Analysis, 4 pages (acceptance rate: 52 %)
- 299) Geiger, J. T., Eyben, F., Schuller, B., and Rigoll, G. (2013b). Detecting Overlapping Speech with Long Short-Term Memory Recurrent Neural Networks. In *Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association*, pages 1668–1672, Lyon, France. ISCA, ISCA. (acceptance rate: 52 %, IF* 1.05 (2010))
- 300) Geiger, J. T., Schuller, B., and Rigoll, G. (2013d). Large-Scale Audio Feature Extraction and SVM for Acoustic Scene Classification. In *Proceedings of the 2013 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, WASPAA 2013*, pages 1–4, New Paltz, NY. IEEE, IEEE
- 301) Geiger, J. T., Eyben, F., Evans, N., Schuller, B., and Rigoll, G. (2013a). Using Linguistic Information to Detect Overlapping Speech. In *Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association*, pages 690–694, Lyon, France. ISCA, ISCA. (acceptance rate: 52 %, IF* 1.05 (2010))
- 302) Geiger, J. T., Hofmann, M., Schuller, B., and Rigoll, G. (2013c). Gait-based Person Identification by Spectral, Cepstral and Energy-related Audio Features. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 458–462, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 303) Geiger, J. T., Weninger, F., Hurmalainen, A., Gemmeke, J. F., Wöllmer, M., Schuller, B., Rigoll, G., and Virtanen, T. (2013e). The TUM+TUT+KUL Approach to the CHiME Challenge 2013: Multi-Stream ASR Exploiting BLSTM Networks and Sparse NMF. In *Proceedings The 2nd CHiME Workshop on Machine Listening in Multisource Environments held in conjunction with ICASSP 2013*, pages 25–30, Vancouver, Canada. IEEE, IEEE. winning paper of track 1 and best paper award
- 304) Han, W., Li, H., Ruan, H., Ma, L., Sun, J., and Schuller, B. (2013). Active Learning for Dimensional Speech Emotion Recognition. In *Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association*, pages 2856–2859, Lyon, France. ISCA, ISCA. (acceptance rate: 52 %, IF* 1.05 (2010))
- 305) Joder, C., Weninger, F., Virette, D., and Schuller, B. (2013a). A Comparative Study on Sparsity Penalties for NMF-based Speech Separation: Beyond LP-Norms. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 858–862, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 306) Joder, C., Weninger, F., Virette, D., and Schuller, B. (2013b). Integrating Noise Estimation and Factorization-based Speech Separation: a Novel Hybrid Approach. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 131–135, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 307) Joder, C. and Schuller, B. (2013). Off-line Refinement of Audio-to-Score Alignment by Observation Template Adaptation. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 206–210, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 308) Newman, S., Golan, O., Baron-Cohen, S., Bölte, S., Baranger, A., Schuller, B., Robinson, P., Camurri, A., Meir, N., Rotman, C., Tal, S., Fridenson, S., O'Reilly, H., Lundqvist, D., Berggren, S., Sullings, N., Marchi, E., Batliner, A., Davies, I., and Piana, S. (2013). ASC-Inclusion - Interactive Software to Help Children with ASC Understand and Express Emotions. In *Proceedings 12th Annual International Meeting For Autism Research (IMFAR 2013)*, San Sebastián, Spain. International Society for Autism Research (INSAR), INSAR. 1 page
- 309) Rosner, A., Weninger, F., Schuller, B., Michalak, M., and Kostek, B. (2013). Influence of Low-Level Features Extracted from Rhythmic and Harmonic Sections on Music Genre Classification. In Gruca, A., Czachurski, T., and Kozielski, S., editors, *Man-Machine Interactions 3*, volume 242 of *Advances in Intelligent Systems and Computing (AISC)*, pages 467–473. Springer
- 310) Valstar, M., Schuller, B., Smith, K., Eyben, F., Jiang, B.,

- Bilakhia, S., Schnieder, S., Cowie, R., and Pantic, M. (2013b). AVEC 2013 - The Continuous Audio/Visual Emotion and Depression Recognition Challenge. In *Proceedings of the 3rd ACM international workshop on Audio/Visual Emotion Challenge*, pages 3–10, Barcelona, Spain. ACM, ACM. (> 100 citations)
- 311) Valstar, M., Schuller, B., Krajewski, J., Cowie, R., and Pantic, M. (2013a). Workshop summary for the 3rd international audio/visual emotion challenge and workshop (AVEC'13). In *Proceedings of the 21st ACM international conference on Multimedia, ACM MM 2013*, pages 1085–1086, Barcelona, Spain. ACM, ACM. (acceptance rate: 28 %)
- 312) Weninger, F., Kirst, C., Schuller, B., and Bungartz, H.-J. (2013d). A Discriminative Approach to Polyphonic Piano Note Transcription using Non-negative Matrix Factorization. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 6–10, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 313) Weninger, F., Wagner, C., Wöllmer, M., Schuller, B., and Morency, L.-P. (2013f). Speaker Trait Characterization in Web Videos: Uniting Speech, Language, and Facial Features. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 3647–3651, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 314) Weninger, F., Geiger, J., Wöllmer, M., Schuller, B., and Rigoll, G. (2013c). The Munich Feature Enhancement Approach to the 2013 CHiME Challenge Using BLSTM Recurrent Neural Networks. In *Proceedings The 2nd CHiME Workshop on Machine Listening in Multisource Environments held in conjunction with ICASSP 2013*, pages 86–90, Vancouver, Canada. IEEE, IEEE
- 315) Weninger, F., Eyben, F., and Schuller, B. (2013a). The TUM Approach to the MediaEval Music Emotion Task Using Generic Affective Audio Features. In Larson, M., Anguera, X., Reuter, T., Jones, G. J., Ionescu, B., Schedl, M., Piatrik, T., Hauff, C., and Soleymani, M., editors, *Proceedings of the MediaEval 2013 Multimedia Benchmark Workshop*, Barcelona, Spain. CEUR. 2 pages, best result
- 316) Wöllmer, M., Zhang, Z., Weninger, F., Schuller, B., and Rigoll, G. (2013f). Feature Enhancement by Bidirectional LSTM Networks for Conversational Speech Recognition in Highly Non-Stationary Noise. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 6822–6826, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 317) Wöllmer, M., Schuller, B., and Rigoll, G. (2013c). Probabilistic ASR Feature Extraction Applying Context-Sensitive Connectionist Temporal Classification Networks. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 7125–7129, Vancouver, Canada
- 318) Zhang, Z., Deng, J., Marchi, E., and Schuller, B. (2013a). Active Learning by Label Uncertainty for Acoustic Emotion Recognition. In *Proceedings INTERSPEECH 2013, 14th Annual Conference of the International Speech Communication Association*, pages 2841–2845, Lyon, France. ISCA, ISCA. (acceptance rate: 52 %, IF* 1.05 (2010))
- 319) Zhang, Z., Deng, J., and Schuller, B. (2013b). Co-Training Succeeds in Computational Paralinguistics. In *Proceedings 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2013*, pages 8505–8509, Vancouver, Canada. IEEE, IEEE. (acceptance rate: 53 %, IF* 1.16 (2010))
- 320) Schuller, B., Steidl, S., Batliner, A., Nöth, E., Vinciarelli, A., Burkhardt, F., van Son, R., Weninger, F., Eyben, F., Bocklet, T., Mohammadi, G., and Weiss, B. (2012d). The INTERSPEECH 2012 Speaker Trait Challenge. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010), > 100 citations)
- 321) Schuller, B., Valstar, M., Eyben, F., Cowie, R., and Pantic, M. (2012g). AVEC 2012 – The Continuous Audio/Visual Emotion Challenge. In Morency, L.-P., Bohus, D., Aghajan, H. K., Cassell, J., Nijholt, A., and Epps, J., editors, *Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI*, pages 449–456, Santa Monica, CA. ACM, ACM. (acceptance rate: 36 %, IF* 1.77 (2010), > 100 citations)
- 322) Schuller, B., Hantke, S., Weninger, F., Han, W., Zhang, Z., and Narayanan, S. (2012c). Automatic Recognition of Emotion Evoked by General Sound Events. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 341–344, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 323) Ungruh, S., Krajewski, J., and Schuller, B. (2012). Maus- und tastaturunterstützte Detektion von Schläfrigkeit Zuständen. In *Proceedings 48. Kongress der Deutschen Gesellschaft für Psychologie*, Bielefeld, Germany. Deutsche Gesellschaft für Psychologie (DGPs), Deutsche Gesellschaft für Psychologie. 1 page
- 324) Eyben, F., Weninger, F., Lehment, N., Rigoll, G., and Schuller, B. (2012d). Violent Scenes Detection with Large, Brute-forced Acoustic and Visual Feature Sets. In Larson, M. A., Schmiedeke, S., Kelm, P., Rae, A., Mezaris, V., Piatrik, T., Soleymani, M., Metz, F., and Jones, G. J., editors, *Working Notes Proceedings of the MediaEval 2012 Workshop*, volume 927, Pisa, Italy. CEUR. 2 pages
- 325) Joder, C., Weninger, F., Wöllmer, M., and Schuller, B. (2012d). The TUM Cumulative DTW Approach for the Mediaeval 2012 Spoken Web Search Task. In Larson, M. A., Schmiedeke, S., Kelm, P., Rae, A., Mezaris, V., Piatrik, T., Soleymani, M., Metz, F., and Jones, G. J., editors, *Working Notes Proceedings of the MediaEval 2012 Workshop*, volume 927, Pisa, Italy. CEUR. 2 pages
- 326) Eyben, F., Schuller, B., and Rigoll, G. (2012c). Improving Generalisation and Robustness of Acoustic Affect Recognition. In Morency, L.-P., Bohus, D., Aghajan, H. K., Cassell, J., Nijholt, A., and Epps, J., editors, *Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI*, pages 517–522, Santa Monica, CA. ACM, ACM. (acceptance rate: 36 %, IF* 1.77 (2010))
- 327) Han, W., Li, H., Ma, L., Zhang, X., Sun, J., Eyben, F., and Schuller, B. (2012b). Preserving Actual Dynamic Trend of Emotion in Dimensional Speech Emotion Recognition. In Morency, L.-P., Bohus, D., Aghajan, H. K., Cassell, J., Nijholt, A., and Epps, J., editors, *Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI*, pages 523–528, Santa Monica, CA. ACM, ACM. (acceptance rate: 36 %, IF* 1.77 (2010))
- 328) Marchi, E., Schuller, B., Batliner, A., Fridenzon, S., Tal, S., and Golan, O. (2012b). Emotion in the Speech of Children with Autism Spectrum Conditions: Prosody and Everything Else. In *Proceedings 3rd Workshop on Child, Computer and Interaction (WOCCI 2012), Satellite Event of INTERSPEECH 2012*, Portland, OR. ISCA, ISCA. 8 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 329) Marchi, E., Batliner, A., Schuller, B., Fridenzon, S., Tal, S., and Golan, O. (2012a). Speech, Emotion, Age, Language, Task, and Typicality: Trying to Disentangle Performance and Feature Relevance. In *Proceedings First International Workshop on Wide Spectrum Social Signal Processing (WSP 2012), held in conjunction with the ASE/IEEE International Conference on Social Computing (SocialCom 2012)*, Amsterdam, The Netherlands. ASE/IEEE, IEEE. 8 pages (acceptance rate: 42 %, IF* 2.6 (2010))
- 330) Deng, J. and Schuller, B. (2012). Confidence Measures in Speech Emotion Recognition Based on Semi-supervised Learning. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF*

- 1.05 (2010))
- 331) Weninger, F., Marchi, E., and Schuller, B. (2012d). Improving Recognition of Speaker States and Traits by Cumulative Evidence: Intoxication, Sleepiness, Age and Gender. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 332) Weninger, F. and Schuller, B. (2012a). Discrimination of Linguistic and Non-Linguistic Vocalizations in Spontaneous Speech: Intra- and Inter-Corpus Perspectives. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 333) Brückner, R. and Schuller, B. (2012). Likability Classification – A not so Deep Neural Network Approach. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 334) Weninger, F., Wöllmer, M., and Schuller, B. (2012g). Combining Bottleneck-BLSTM and Semi-Supervised Sparse NMF for Recognition of Conversational Speech in Highly Instationary Noise. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 335) Zhang, Z. and Schuller, B. (2012a). Active Learning by Sparse Instance Tracking and Classifier Confidence in Acoustic Emotion Recognition. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 336) Geiger, J. T., Vipperla, R., Bozonnet, S., Evans, N., Schuller, B., and Rigoll, G. (2012a). Convolutional Non-Negative Sparse Coding and New Features for Speech Overlap Handling in Speaker Diarization. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 337) Wöllmer, M., Eyben, F., and Schuller, B. (2012a). Temporal and Situational Context Modeling for Improved Dominance Recognition in Meetings. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 338) Ringeval, F., Chetouani, M., and Schuller, B. (2012). Novel Metrics of Speech Rhythm for the Assessment of Emotion. In *Proceedings INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association*, Portland, OR. ISCA, ISCA. 4 pages (acceptance rate: 52 %, IF* 1.05 (2010))
- 339) Joder, C. and Schuller, B. (2012b). Score-Informed Leading Voice Separation from Monaural Audio. In *Proceedings 13th International Society for Music Information Retrieval Conference, ISMIR 2012*, pages 277–282, Porto, Portugal. ISMIR, ISMIR. (acceptance rate: 44 %)
- 340) Geiger, J. T., Vipperla, R., Evans, N., Schuller, B., and Rigoll, G. (2012b). Speech Overlap Handling for Speaker Diarization Using Convolutional Non-negative Sparse Coding and Energy-Related Features. In *Proceedings 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania. EURASIP, EURASIP. 4 pages
- 341) Han, W., Li, H., Ma, L., Zhang, X., and Schuller, B. (2012a). A Ranking-based Emotion Annotation Scheme and Real-life Speech Database. In *Proceedings 4th International Workshop on EMOTION SENTIMENT & SOCIAL SIGNALS 2012 (ES 2012) – Corpora for Research on Emotion, Sentiment & Social Signals, held in conjunction with LREC 2012*, pages 67–71, Istanbul, Turkey. ELRA, ELRA. (acceptance rate: 68 %, IF* 1.37 (2010))
- 342) Principi, E., Rotili, R., Wöllmer, M., Squartini, S., and Schuller, B. (2012b). Dominance Detection in a Reverberated Acoustic Scenario. In *Proceedings 9th International Conference on Advances in Neural Networks, ISNN 2012, Shenyang, China, 11-14.07.2012*, volume 7367 of *Lecture Notes in Computer Science (LNCS)*, pages 394–402. Springer, Berlin/Heidelberg. Special Session on Advances in Cognitive and Emotional Information Processing
- 343) Weninger, F., Wöllmer, M., Geiger, J., Schuller, B., Gemmeke, J., Hurmalainen, A., Virtanen, T., and Rigoll, G. (2012f). Non-Negative Matrix Factorization for Highly Noise-Robust ASR: to Enhance or to Recognize? In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 4681–4684, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 344) Wöllmer, M., Metallinou, A., Katsamanis, N., Schuller, B., and Narayanan, S. (2012c). Analyzing the Memory of BLSTM Neural Networks for Enhanced Emotion Classification in Dyadic Spoken Interactions. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 4157–4160, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 345) Zhang, Z. and Schuller, B. (2012b). Semi-supervised Learning Helps in Sound Event Classification. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 333–336, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 346) Weninger, F., Feliu, J., and Schuller, B. (2012b). Supervised and Semi-Supervised Suppression of Background Music in Monaural Speech Recordings. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 61–64, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 347) Weninger, F., Amir, N., Amir, O., Ronen, I., Eyben, F., and Schuller, B. (2012a). Robust Feature Extraction for Automatic Recognition of Vibrato Singing in Recorded Polyphonic Music. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 85–88, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 348) Prylipko, D., Schuller, B., and Wendemuth, A. (2012). Fine-Tuning HMMs for Nonverbal Vocalizations in Spontaneous Speech: a Multicorpus Perspective. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 4625–4628, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 349) Eyben, F., Petridis, S., Schuller, B., and Pantic, M. (2012b). Audiovisual Vocal Outburst Classification in Noisy Acoustic Conditions. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 5097–5100, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 350) Vipperla, R., Geiger, J., Bozonnet, S., Wang, D., Evans, N., Schuller, B., and Rigoll, G. (2012). Speech Overlap Detection and Attribution Using Convolutional Non-Negative Sparse Coding. In *Proceedings 37th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 4181–4184, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 351) Joder, C., Weninger, F., Eyben, F., Virette, D., and Schuller, B. (2012a). Real-time Speech Separation by Semi-Supervised Nonnegative Matrix Factorization. In Theis, F. J., Cichocki, A., Yeredor, A., and Zibulevsky, M., editors, *Proceedings 10th International Conference on Latent Variable Analysis and Signal Separation, LVA/ICA 2012*, volume 7191 of *Lecture Notes in Computer Science*, pages 322–329, Tel Aviv, Israel. Springer. Special Session Real-world constraints and opportunities in audio source separation

- 352) Schuller, B., Wenginger, F., and Dorfner, J. (2011f). Multi-Modal Non-Prototypical Music Mood Analysis in Continuous Space: Reliability and Performances. In *Proceedings 12th International Society for Music Information Retrieval Conference, ISMIR 2011*, pages 759–764, Miami, FL. ISMIR, ISMIR. (acceptance rate: 59 %)
- 353) Schuller, B., Valstar, M., Eyben, F., McKeown, G., Cowie, R., and Pantic, M. (2011e). AVEC 2011 - The First International Audio/Visual Emotion Challenge. In Schuller, B., Valstar, M., Cowie, R., and Pantic, M., editors, *Proceedings First International Audio/Visual Emotion Challenge and Workshop, AVEC 2011, held in conjunction with the International HU-MAINE Association Conference on Affective Computing and Intelligent Interaction 2011, ACII 2011*, volume II, pages 415–424. Springer, Memphis, TN. (> 200 citations)
- 354) Schuller, B., Zhang, Z., Wenginger, F., and Rigoll, G. (2011i). Selecting Training Data for Cross-Corpus Speech Emotion Recognition: Prototypicality vs. Generalization. In *Proceedings 2011 Speech Processing Conference*, Tel Aviv, Israel. AVIOS, AVIOS. invited contribution, 4 pages
- 355) Schuller, B., Batliner, A., Steidl, S., Schiel, F., and Krajewski, J. (2011a). The INTERSPEECH 2011 Speaker State Challenge. In *Proceedings INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, pages 3201–3204, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010), > 100 citations)
- 356) Schuller, B., Zhang, Z., Wenginger, F., and Rigoll, G. (2011j). Using Multiple Databases for Training in Emotion Recognition: To Unite or to Vote? In *Proceedings INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, pages 1553–1556, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 357) Rotili, R., Principi, E., Squartini, S., and Schuller, B. (2011). A Real-Time Speech Enhancement Framework for Multi-party Meetings. In Travieso-González, C. M. and Alonso-Hernández, J., editors, *Advances in Nonlinear Speech Processing, 5th International Conference on Nonlinear Speech Processing, NoLISP 2011, Las Palmas de Gran Canaria, Spain, November 7-9, 2011, Proceedings*, volume 7015/2011 of *Lecture Notes in Computer Science (LNCS)*, pages 80–87. Springer
- 358) Wöllmer, M. and Schuller, B. (2011). Enhancing Spontaneous Speech Recognition with BLSTM Features. In Travieso-González, C. M. and Alonso-Hernández, J., editors, *Advances in Nonlinear Speech Processing, 5th International Conference on Nonlinear Speech Processing, NoLISP 2011, Las Palmas de Gran Canaria, Spain, November 7-9, 2011, Proceedings*, volume 7015/2011 of *Lecture Notes in Computer Science (LNCS)*, pages 17–24. Springer
- 359) Zhang, Z., Wenginger, F., Wöllmer, M., and Schuller, B. (2011). Unsupervised Learning in Cross-Corpus Acoustic Emotion Recognition. In *Proceedings 12th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2011*, pages 523–528, Big Island, HI. IEEE, IEEE. (acceptance rate: 43 %)
- 360) Wöllmer, M., Schuller, B., and Rigoll, G. (2011f). A Novel Bottleneck-BLSTM Front-End for Feature-Level Context Modeling in Conversational Speech Recognition. In *Proceedings 12th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2011*, pages 36–41, Big Island, HI. IEEE, IEEE. (acceptance rate: 43 %)
- 361) Wenginger, F., Wöllmer, M., and Schuller, B. (2011f). Automatic Assessment of Singer Traits in Popular Music: Gender, Age, Height and Race. In *Proceedings 12th International Society for Music Information Retrieval Conference, ISMIR 2011*, pages 37–42, Miami, FL. ISMIR, ISMIR. (acceptance rate: 59 %)
- 362) Ungruh, S., Krajewski, J., Eyben, F., and Schuller, B. (2011). Maus- und tastaturunterstützte Detektion von Schläfrigkeitstufen. In *Proceedings 7. Tagung der Fachgruppe Arbeits-, Organisations- und Wirtschaftspsychologie, AOW 2011*, Rostock, Germany. Deutsche Gesellschaft für Psychologie (DGPs), Deutsche Gesellschaft für Psychologie. 1 page
- 363) Wenginger, F., Geiger, J., Wöllmer, M., Schuller, B., and Rigoll, G. (2011b). The Munich 2011 CHiME Challenge Contribution: NMF-BLSTM Speech Enhancement and Recognition for Reverberated Multisource Environments. In *Proceedings Machine Listening in Multisource Environments, CHiME 2011, satellite workshop of Interspeech 2011*, pages 24–29, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 364) Wöllmer, M., Wenginger, F., Eyben, F., and Schuller, B. (2011h). Acoustic-Linguistic Recognition of Interest in Speech with Bottleneck-BLSTM Nets. In *Proceedings INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, pages 3201–3204, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 365) Wöllmer, M., Wenginger, F., Steidl, S., Batliner, A., and Schuller, B. (2011j). Speech-based Non-prototypical Affect Recognition for Child-Robot Interaction in Reverberated Environments. In *Proceedings INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, pages 3113–3116, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 366) Wöllmer, M., Schuller, B., and Rigoll, G. (2011g). Feature Frame Stacking in RNN-based Tandem ASR Systems – Learned vs. Predefined Context. In *Proceedings INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, pages 1233–1236, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 367) Burkhardt, F., Schuller, B., Weiss, B., and Wenginger, F. (2011). “Would You Buy A Car From Me?” - On the Likability of Telephone Voices. In *Proceedings INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, pages 1557–1560, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 368) Geiger, J. T., Lakhal, M. A., Schuller, B., and Rigoll, G. (2011). Learning new acoustic events in an HMM-based system using MAP adaptation. In *Proceedings INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, pages 293–296, Florence, Italy. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 369) Gunes, H., Schuller, B., Pantic, M., and Cowie, R. (2011). Emotion Representation, Analysis and Synthesis in Continuous Space: A Survey. In *Proceedings International Workshop on Emotion Synthesis, rePresentation, and Analysis in Continuous space, EmoSPACE 2011, held in conjunction with the 9th IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, FG 2011*, pages 827–834, Santa Barbara, CA. IEEE, IEEE. (> 100 citations)
- 370) Wöllmer, M., Marchi, E., Squartini, S., and Schuller, B. (2011d). Robust Multi-Stream Keyword and Non-Linguistic Vocalization Detection for Computationally Intelligent Virtual Agents. In Liu, D., Zhang, H., Polycarpou, M., Alippi, C., and He, H., editors, *Proceedings 8th International Conference on Advances in Neural Networks, ISNN 2011, Guilin, China, 29.05.-01.06.2011*, volume 6676, Part II of *Lecture Notes in Computer Science (LNCS)*, pages 496–505. Springer, Berlin/Heidelberg
- 371) Eyben, F., Wöllmer, M., Valstar, M., Gunes, H., Schuller, B., and Pantic, M. (2011b). String-based Audiovisual Fusion of Behavioural Events for the Assessment of Dimensional Affect. In *Proceedings International Workshop on Emotion Synthesis, rePresentation, and Analysis in Continuous space, EmoSPACE 2011, held in conjunction with the 9th IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, FG 2011*, pages 322–329, Santa Barbara, CA. IEEE, IEEE
- 372) Eyben, F., Petridis, S., Schuller, B., Tzimiropoulos, G., Zafeiriou, S., and Pantic, M. (2011a). Audiovisual Classification of Vocal Outbursts in Human Conversation Using Long-Short-Term Memory Networks. In *Proceedings 36th IEEE*

- International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 5844–5847, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 373) Weninger, F., Schuller, B., Wöllmer, M., and Rigoll, G. (2011e). Localization of Non-Linguistic Events in Spontaneous Speech by Non-Negative Matrix Factorization and Long Short-Term Memory. In *Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 5840–5843, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 374) Weninger, F. and Schuller, B. (2011a). Audio Recognition in the Wild: Static and Dynamic Classification on a Real-World Database of Animal Vocalizations. In *Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 337–340, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 375) Wöllmer, M., Eyben, F., Schuller, B., and Rigoll, G. (2011b). A Multi-Stream ASR Framework for BLSTM Modeling of Conversational Speech. In *Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 4860–4863, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 376) Weninger, F., Durrieu, J.-L., Eyben, F., Richard, G., and Schuller, B. (2011a). Combining Monaural Source Separation With Long Short-Term Memory for Increased Robustness in Vocalist Gender Recognition. In *Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 2196–2199, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 377) Weninger, F., Lehmann, A., and Schuller, B. (2011c). openBliS-SART: Design and Evaluation of a Research Toolkit for Blind Source Separation in Audio Recognition Tasks. In *Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 1625–1628, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 378) Landsiedel, C., Edlund, J., Eyben, F., Neiberg, D., and Schuller, B. (2011). Syllabification of Conversational Speech Using Bidirectional Long-Short-Term Memory Neural Networks. In *Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 5265–5268, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 379) Stuhlsatz, A., Meyer, C., Eyben, F., Zielke, T., Meier, G., and Schuller, B. (2011). Deep Neural Networks for Acoustic Emotion Recognition: Raising the Benchmarks. In *Proceedings 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011*, pages 5688–5691, Prague, Czech Republic. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 380) Schuller, B., Steidl, S., Batliner, A., Burkhardt, F., Devillers, L., Müller, C., and Narayanan, S. (2010g). The INTERSPEECH 2010 Paralinguistic Challenge. In *Proceedings INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, pages 2794–2797, Makuhari, Japan. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010), > 200 citations)
- 381) Schuller, B. and Devillers, L. (2010b). Incremental Acoustic Valence Recognition: an Inter-Corpus Perspective on Features, Matching, and Performance in a Gating Paradigm. In *Proceedings INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, pages 2794–2797, Makuhari, Japan. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 382) Schuller, B., Kozielski, C., Weninger, F., Eyben, F., and Rigoll, G. (2010e). Vocalist Gender Recognition in Recorded Popular Music. In *Proceedings 11th International Society for Music Information Retrieval Conference, ISMIR 2010*, pages 613–618, Utrecht, The Netherlands. ISMIR, ISMIR. (acceptance rate: 61 %)
- 383) Schuller, B., Zaccarelli, R., Rollet, N., and Devillers, L. (2010j). CINEMO - A French Spoken Language Resource for Complex Emotions: Facts and Baselines. In Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., and Tapias, D., editors, *Proceedings 7th International Conference on Language Resources and Evaluation, LREC 2010*, pages 1643–1647, Valletta, Malta. ELRA, European Language Resources Association. (acceptance rate: 69 %, IF* 1.37 (2010))
- 384) Schuller, B., Eyben, F., Can, S., and Feussner, H. (2010c). Speech in Minimal Invasive Surgery – Towards an Affective Language Resource of Real-life Medical Operations. In Devillers, L., Schuller, B., Cowie, R., Douglas-Cowie, E., and Batliner, A., editors, *Proceedings 3rd International Workshop on EMOTION: Corpora for Research on Emotion and Affect, satellite of LREC 2010*, pages 5–9, Valletta, Malta. ELRA, European Language Resources Association. (acceptance rate: 69 %, IF* 1.37 (2010))
- 385) Schuller, B., Weninger, F., Wöllmer, M., Sun, Y., and Rigoll, G. (2010i). Non-Negative Matrix Factorization as Noise-Robust Feature Extractor for Speech Recognition. In *Proceedings 35th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, pages 4562–4565, Dallas, TX. IEEE, IEEE. (acceptance rate: 48 %, IF* 1.16 (2010))
- 386) Schuller, B. and Burkhardt, F. (2010). Learning with Synthesized Speech for Automatic Emotion Recognition. In *Proceedings 35th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, pages 5150–5155, Dallas, TX. IEEE, IEEE. (acceptance rate: 48 %, IF* 1.16 (2010))
- 387) Schuller, B. and Weninger, F. (2010). Discrimination of Speech and Non-Linguistic Vocalizations by Non-Negative Matrix Factorization. In *Proceedings 35th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, pages 5054–5057, Dallas, TX. IEEE, IEEE. (acceptance rate: 48 %, IF* 1.16 (2010))
- 388) Schuller, B., Metze, F., Steidl, S., Batliner, A., Eyben, F., and Polzehl, T. (2010f). Late Fusion of Individual Engines for Improved Recognition of Negative Emotions in Speech – Learning vs. Democratic Vote. In *Proceedings 35th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, pages 5230–5233, Dallas, TX. IEEE, IEEE. (acceptance rate: 48 %, IF* 1.16 (2010))
- 389) Metze, F., Batliner, A., Eyben, F., Polzehl, T., Schuller, B., and Steidl, S. (2010). Emotion Recognition using Imperfect Speech Recognition. In *Proceedings INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, pages 478–481, Makuhari, Japan. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 390) Wöllmer, M., Eyben, F., Schuller, B., and Rigoll, G. (2010c). Recognition of Spontaneous Conversational Speech using Long Short-Term Memory Phoneme Predictions. In *Proceedings INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, pages 1946–1949, Makuhari, Japan. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 391) Wöllmer, M., Sun, Y., Eyben, F., and Schuller, B. (2010h). Long Short-Term Memory Networks for Noise Robust Speech Recognition. In *Proceedings INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, pages 2966–2969, Makuhari, Japan. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 392) Wöllmer, M., Metallinou, A., Eyben, F., Schuller, B., and Narayanan, S. (2010f). Context-Sensitive Multimodal Emotion Recognition from Speech and Facial Expression using Bidirectional LSTM Modeling. In *Proceedings INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, pages 2362–2365, Makuhari, Japan. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 393) Wöllmer, M., Eyben, F., Schuller, B., and Rigoll, G. (2010d).

- Spoken Term Detection with Connectionist Temporal Classification: a Novel Hybrid CTC-DBN Decoder. In *Proceedings 35th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, pages 5274–5277, Dallas, TX. IEEE, IEEE. (acceptance rate: 48 %, IF* 1.16 (2010))
- 394) Eyben, F., Wöllmer, M., and Schuller, B. (2010e). openSMILE – The Munich Versatile and Fast Open-Source Audio Feature Extractor. In *Proceedings of the 18th ACM International Conference on Multimedia, MM 2010*, pages 1459–1462, Florence, Italy. ACM, ACM. (Honorable Mention (2nd place) in the ACM MM 2010 Open-source Software Competition, acceptance rate short paper: about 30 %, IF* 2.45 (2010), > 600 citations)
- 395) Arsić, D., Wöllmer, M., Rigoll, G., Roalter, L., Kranz, M., Kaiser, M., Eyben, F., and Schuller, B. (2010). Automated 3D Gesture Recognition Applying Long Short-Term Memory and Contextual Knowledge in a CAVE. In *Proceedings 1st Workshop on Multimodal Pervasive Video Analysis, MPVA 2010, held in conjunction with ACM Multimedia 2010*, pages 33–36, Florence, Italy. ACM, ACM. (acceptance rate short paper: about 30 %, IF* 2.45 (2010))
- 396) Eyben, F., Böck, S., Schuller, B., and Graves, A. (2010b). Universal Onset Detection with Bidirectional Long-Short Term Memory Neural Networks. In *Proceedings 11th International Society for Music Information Retrieval Conference, ISMIR 2010*, pages 589–594, Utrecht, The Netherlands. ISMIR, ISMIR. (acceptance rate: 61 %)
- 397) Brendel, M., Zaccarelli, R., Schuller, B., and Devillers, L. (2010). Towards measuring similarity between emotional corpora. In Devillers, L., Schuller, B., Cowie, R., Douglas-Cowie, E., and Batliner, A., editors, *Proceedings 3rd International Workshop on EMOTION: Corpora for Research on Emotion and Affect, satellite of LREC 2010*, pages 58–64, Valletta, Malta. ELRA, European Language Resources Association. (acceptance rate: 69 %, IF* 1.37 (2010))
- 398) Eyben, F., Batliner, A., Schuller, B., Seppi, D., and Steidl, S. (2010a). Cross-Corpus Classification of Realistic Emotions - Some Pilot Experiments. In Devillers, L., Schuller, B., Cowie, R., Douglas-Cowie, E., and Batliner, A., editors, *Proceedings 3rd International Workshop on EMOTION: Corpora for Research on Emotion and Affect, satellite of LREC 2010*, pages 77–82, Valletta, Malta. ELRA, European Language Resources Association. (acceptance rate: 69 %, IF* 1.37 (2010))
- 399) de Sevin, E., Bevacqua, E., Pammi, S., Pelachaud, C., Schröder, M., and Schuller, B. (2010). A Multimodal Listener Behaviour Driven by Audio Input. In *Proceedings International Workshop on Interacting with ECAs as Virtual Characters, satellite of AAMAS 2010*, Toronto, Canada. ACM, ACM. 4 pages (acceptance rate: 24 %, IF* 3.12 (2010))
- 400) Schröder, M., Pammi, S., Cowie, R., McKeown, G., Gunes, H., Pantic, M., Valstar, M., Heylen, D., ter Maat, M., Eyben, F., Schuller, B., Wöllmer, M., Bevacqua, E., Pelachaud, C., and de Sevin, E. (2010b). Demo: Have a Chat with Sensitive Artificial Listeners. In *Proceedings 36th Annual Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour, AISB 2010*, Leicester, UK. AISB, AISB. Symposium Towards a Comprehensive Intelligence Test, TCIT, 1 page
- 401) Schröder, M., Cowie, R., Heylen, D., Pantic, M., Pelachaud, C., and Schuller, B. (2010a). How to build a machine that people enjoy talking to. In *Proceedings 4th International Conference on Cognitive Systems, CogSys, Zurich, Switzerland*. 1 page
- 402) Seppi, D., Batliner, A., Steidl, S., Schuller, B., and Nöth, E. (2010). Word Accent and Emotion. In *Proceedings 5th International Conference on Speech Prosody, SP 2010*, Chicago, IL. ISCA, ISCA. 4 pages
- 403) Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., and Wendemuth, A. (2009j). Acoustic Emotion Recognition: A Benchmark Comparison of Performances. In *Proceedings 11th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2009*, pages 552–557, Merano, Italy. IEEE, IEEE. (acceptance rate: 43 %, > 100 citations)
- 404) Schuller, B., Steidl, S., and Batliner, A. (2009h). The Interspeech 2009 Emotion Challenge. In *Proceedings INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association*, pages 312–315, Brighton, UK. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010), > 400 citations)
- 405) Schuller, B. and Rigoll, G. (2009). Recognising Interest in Conversational Speech – Comparing Bag of Frames and Suprasegmental Features. In *Proceedings INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association*, pages 1999–2002, Brighton, UK. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 406) Schuller, B., Schenk, J., Rigoll, G., and Knaup, T. (2009g). “The Godfather” vs. “Chaos”: Comparing Linguistic Analysis based on Online Knowledge Sources and Bags-of-N-Grams for Movie Review Valence Estimation. In *Proceedings 10th International Conference on Document Analysis and Recognition, ICDAR 2009*, pages 858–862, Barcelona, Spain. IAPR, IEEE. (acceptance rate: 64 %)
- 407) Schuller, B., Can, S., Feussner, H., Wöllmer, M., Arsić, D., and Hörnler, B. (2009c). Speech Control in Surgery: a Field Analysis and Strategies. In *Proceedings 10th IEEE International Conference on Multimedia and Expo, ICME 2009*, pages 1214–1217, New York, NY. IEEE, IEEE. (acceptance rate: about 30 %, IF* 0.88 (2010))
- 408) Schuller, B., Hörnler, B., Arsić, D., and Rigoll, G. (2009d). Audio Chord Labeling by Musiological Modeling and Beat-Synchronization. In *Proceedings 10th IEEE International Conference on Multimedia and Expo, ICME 2009*, pages 526–529, New York, NY. IEEE, IEEE. (acceptance rate: about 30 %, IF* 0.88 (2010))
- 409) Schuller, B. (2009). Traits Prosodiques dans la Modélisation Acoustique à Base de Segment. In Hancil, S., editor, *Proceedings Conférence Internationale sur Prosodie et Iconicité, Proscico 2009*, pages 24–26, Rouen, France
- 410) Schuller, B., Batliner, A., Steidl, S., and Seppi, D. (2009a). Emotion Recognition from Speech: Putting ASR in the Loop. In *Proceedings 34th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2009*, pages 4585–4588, Taipei, Taiwan. IEEE, IEEE. (acceptance rate: 43 %, IF* 1.16 (2010))
- 411) Wöllmer, M., Eyben, F., Schuller, B., and Rigoll, G. (2009e). Robust Vocabulary Independent Keyword Spotting with Graphical Models. In *Proceedings 11th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2009*, pages 349–353, Merano, Italy. IEEE, IEEE. (acceptance rate: 43 %)
- 412) Eyben, F., Wöllmer, M., Schuller, B., and Graves, A. (2009b). From Speech to Letters – Using a novel Neural Network Architecture for Grapheme Based ASR. In *Proceedings 11th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2009*, pages 376–380, Merano, Italy. IEEE, IEEE. (acceptance rate: 43 %)
- 413) Wöllmer, M., Eyben, F., Schuller, B., Douglas-Cowie, E., and Cowie, R. (2009d). Data-driven Clustering in Emotional Space for Affect Recognition Using Discriminatively Trained LSTM Networks. In *Proceedings INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association*, pages 1595–1598, Brighton, UK. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 414) Wöllmer, M., Eyben, F., Schuller, B., Sun, Y., Moosmayr, T., and Nguyen-Thien, N. (2009f). Robust In-Car Spelling Recognition – A Tandem BLSTM-HMM Approach. In *Proceedings INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association*, pages 1990–1972, Brighton, UK. ISCA, ISCA. (acceptance rate: 58 %, IF* 1.05 (2010))
- 415) Schenk, J., Hörnler, B., Schuller, B., Braun, A., and Rigoll,

- G. (2009). GMs in On-Line Handwritten Whiteboard Note Recognition: the Influence of Implementation and Modeling. In *Proceedings 10th International Conference on Document Analysis and Recognition, ICDAR 2009*, pages 877–880, Barcelona, Spain. IAPR, IEEE. (acceptance rate: 64 %)
- 416) Hörnler, B., Arsić, D., Schuller, B., and Rigoll, G. (2009a). Boosting Multi-modal Camera Selection with Semantic Features. In *Proceedings 10th IEEE International Conference on Multimedia and Expo, ICME 2009*, pages 1298–1301, New York, NY. IEEE, IEEE. (acceptance rate: about 30 %, IF* 0.88 (2010))
- 417) Wöllmer, M., Eyben, F., Keshet, J., Graves, A., Schuller, B., and Rigoll, G. (2009c). Robust Discriminative Keyword Spotting for Emotionally Colored Spontaneous Speech Using Bidirectional LSTM Networks. In *Proceedings 34th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2009*, pages 3949–3952, Taipei, Taiwan. IEEE, IEEE. (acceptance rate: 43 %, IF* 1.16 (2010))
- 418) Arsić, D., Lyutskanov, A., Schuller, B., and Rigoll, G. (2009b). Applying Bayes Markov Chains for the Detection of ATM Related Scenarios. In *Proceedings 10th IEEE Workshop on Applications of Computer Vision, WACV 2009*, pages 464–471, Snowbird, UT. IEEE, IEEE
- 419) Schröder, M., Bevacqua, E., Eyben, F., Gunes, H., Heylen, D., ter Maat, M., Pammi, S., Pantic, M., Pelachaud, C., Schuller, B., de Sevin, E., Valstar, M., and Wöllmer, M. (2009). A Demonstration of Audiovisual Sensitive Artificial Listeners. In *Proceedings 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*, volume I, pages 263–264, Amsterdam, The Netherlands. HUMAINE Association, IEEE. Best Technical Demonstration Award
- 420) Eyben, F., Wöllmer, M., and Schuller, B. (2009a). openEAR – Introducing the Munich Open-Source Emotion and Affect Recognition Toolkit. In *Proceedings 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*, volume I, pages 576–581, Amsterdam, The Netherlands. HUMAINE Association, IEEE. (> 200 citations)
- 421) Wöllmer, M., Eyben, F., Graves, A., Schuller, B., and Rigoll, G. (2009b). A Tandem BLSTM-DBN Architecture for Keyword Spotting with Enhanced Context Modeling. In *Proceedings ISCA Tutorial and Research Workshop on Non-Linear Speech Processing, NOLISP 2009*, Vic, Spain. ISCA, ISCA. 9 pages
- 422) Lehment, N., Arsić, D., Lyutskanov, A., Schuller, B., and Rigoll, G. (2009). Supporting Multi Camera Tracking by Monocular Deformable Graph Tracking. In *Proceedings 11th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, PETS 2009, in Conjunction with the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009*, pages 87–94, Miami, FL. IEEE, IEEE. (acceptance rate: 26 %, IF* 5.97 (2010))
- 423) Arsić, D., Schuller, B., Hörnler, B., and Rigoll, G. (2009c). A Hierarchical Approach for Visual Suspicious Behavior Detection in Aircrafts. In *Proceedings 16th International Conference on Digital Signal Processing, DSP 2009*, Santorini, Greece. IEEE, IEEE. 7 pages (acceptance rate oral: 38 %)
- 424) Arsić, D., Hörnler, B., Schuller, B., and Rigoll, G. (2009a). Resolving Partial Occlusions in Crowded Environments Utilizing Range Data and Video Cameras. In *Proceedings 16th International Conference on Digital Signal Processing, DSP 2009*, Santorini, Greece. IEEE, IEEE. 6 pages (acceptance rate oral: 38 %)
- 425) Hörnler, B., Arsić, D., Schuller, B., and Rigoll, G. (2009b). Graphical Models for Multi-Modal Automatic Video Editing in Meetings. In *Proceedings 16th International Conference on Digital Signal Processing, DSP 2009*, Santorini, Greece. IEEE, IEEE. 8 pages (acceptance rate oral: 38 %)
- 426) Batliner, A., Steidl, S., Eyben, F., and Schuller, B. (2009). Laughter in Child-Robot Interaction. In *Proceedings Interdisciplinary Workshop on Laughter and other Interactional Vocalisations in Speech, Laughter 2009*, Berlin, Germany
- 427) Schuller, B., Batliner, A., Steidl, S., and Seppi, D. (2008a). Does Affect Affect Automatic Recognition of Children’s Speech? In *Proceedings 1st Workshop on Child, Computer and Interaction, WOCCI 2008, ACM ICMI 2008 post-conference workshop*, Chania, Greece. ISCA, ISCA. 4 pages (acceptance rate: 44 %, IF* 1.77 (2010))
- 428) Seppi, D., Gerosa, M., Schuller, B., Batliner, A., and Steidl, S. (2008b). Detecting Problems in Spoken Child-Computer Interaction. In *Proceedings 1st Workshop on Child, Computer and Interaction, WOCCI 2008, ACM ICMI 2008 post-conference workshop*, Chania, Greece. ISCA, ISCA. 4 pages (acceptance rate: 44 %, IF* 1.77 (2010))
- 429) Schuller, B., Wöllmer, M., Moosmayr, T., and Rigoll, G. (2008l). Speech Recognition in Noisy Environments using a Switching Linear Dynamic Model for Feature Enhancement. In *Proceedings INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, incorporating 12th Australasian International Conference on Speech Science and Technology, SST 2008*, pages 1789–1792, Brisbane, Australia. ISCA/ASSTA, ISCA. Special Session Human-Machine Comparisons of Consonant Recognition in Noise (Consonant Challenge) (acceptance rate: 59 %, IF* 1.05 (2010))
- 430) Schuller, B., Zhang, X., and Rigoll, G. (2008n). Prosodic and Spectral Features within Segment-based Acoustic Modeling. In *Proceedings INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, incorporating 12th Australasian International Conference on Speech Science and Technology, SST 2008*, pages 2370–2373, Brisbane, Australia. ISCA/ASSTA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 431) Schuller, B., Wimmer, M., Arsić, D., Moosmayr, T., and Rigoll, G. (2008i). Detection of Security Related Affect and Behaviour in Passenger Transport. In *Proceedings INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, incorporating 12th Australasian International Conference on Speech Science and Technology, SST 2008*, pages 265–268, Brisbane, Australia. ISCA/ASSTA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 432) Schuller, B., Dibiasi, F., Eyben, F., and Rigoll, G. (2008c). One Day in Half an Hour: Music Thumbnailing Incorporating Harmony- and Rhythm Structure. In *Proceedings 6th Workshop on Adaptive Multimedia Retrieval, AMR 2008*, Berlin, Germany. 10 pages
- 433) Schuller, B., Rigoll, G., Can, S., and Feussner, H. (2008g). Emotion Sensitive Speech Control for Human-Robot Interaction in Minimal Invasive Surgery. In *Proceedings 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2008*, pages 453–458, Munich, Germany. IEEE, IEEE
- 434) Schuller, B., Vlasenko, B., Arsić, D., Rigoll, G., and Wendenmuth, A. (2008h). Combining Speech Recognition and Acoustic Word Emotion Models for Robust Text-Independent Emotion Recognition. In *Proceedings 9th IEEE International Conference on Multimedia and Expo, ICME 2008*, pages 1333–1336, Hannover, Germany. IEEE, IEEE. (acceptance rate: 50 %, IF* 0.88 (2010))
- 435) Schuller, B., Wimmer, M., Mösenlechner, L., Kern, C., Arsić, D., and Rigoll, G. (2008j). Brute-Forcing Hierarchical Functionals for Paralinguistics: a Waste of Feature Space? In *Proceedings 33rd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2008*, pages 4501–4504, Las Vegas, NV. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 436) Wöllmer, M., Eyben, F., Reiter, S., Schuller, B., Cox, C., Douglas-Cowie, E., and Cowie, R. (2008). Abandoning Emotion Classes – Towards Continuous Emotion Recognition with

- Modelling of Long-Range Dependencies. In *Proceedings INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, incorporating 12th Australasian International Conference on Speech Science and Technology, SST 2008*, pages 597–600, Brisbane, Australia. ISCA/ASSTA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010), > 100 citations)
- 437) Seppi, D., Batliner, A., Schuller, B., Steidl, S., Vogt, T., Wagner, J., Devillers, L., Vidrascu, L., Amir, N., and Aharonson, V. (2008a). Patterns, Prototypes, Performance: Classifying Emotional User States. In *Proceedings INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, incorporating 12th Australasian International Conference on Speech Science and Technology, SST 2008*, pages 601–604, Brisbane, Australia. ISCA/ASSTA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 438) Vlasenko, B., Schuller, B., Mengistu, K. T., Rigoll, G., and Wendemuth, A. (2008a). Balancing Spoken Content Adaptation and Unit Length in the Recognition of Emotion and Interest. In *Proceedings INTERSPEECH 2008, 9th Annual Conference of the International Speech Communication Association, incorporating 12th Australasian International Conference on Speech Science and Technology, SST 2008*, pages 805–808, Brisbane, Australia. ISCA/ASSTA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 439) Batliner, A., Schuller, B., Schaeffler, S., and Steidl, S. (2008a). Mothers, Adults, Children, Pets – Towards the Acoustics of Intimacy. In *Proceedings 33rd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2008*, pages 4497–4500, Las Vegas, NV. IEEE, IEEE. (acceptance rate: 50 %, IF* 1.16 (2010))
- 440) Arsić, D., Schuller, B., and Rigoll, G. (2008b). Multiple Camera Person Tracking in multiple layers combining 2D and 3D information. In *Proceedings Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications, M2SFA2 2008, in conjunction with 10th European Conference on Computer Vision, ECCV 2008*, pages 1–12, Marseille, France. (acceptance rate: about 23 %, IF* 5.91 (2010))
- 441) Schröder, M., Cowie, R., Heylen, D., Pantic, M., Pelachaud, C., and Schuller, B. (2008). Towards responsive Sensitive Artificial Listeners. In *Proceedings 4th International Workshop on Human-Computer Conversation*, Bellagio, Italy. 6 pages
- 442) Arsić, D., Lehment, N., Hristov, E., Schuller, B., and Rigoll, G. (2008a). Applying Multi Layer Homography for Multi Camera tracking. In *Proceedings Workshop on Activity Monitoring by Multi-Camera Surveillance Systems, AMMCSS 2008, in conjunction with 2nd ACM/IEEE International Conference on Distributed Smart Cameras, ICDSC 2008*, Stanford, CA. ACM/IEEE, IEEE. 9 pages
- 443) Wimmer, M., Schuller, B., Arsić, D., Radig, B., and Rigoll, G. (2008). Low-Level Fusion of Audio and Video Features For Multi-Modal Emotion Recognition. In *Proceedings 3rd International Conference on Computer Vision Theory and Applications, VISAPP 2008*, Funchal, Portugal. 7 pages
- 444) Schuller, B., Vlasenko, B., Minguez, R., Rigoll, G., and Wendemuth, A. (2007f). Comparing One and Two-Stage Acoustic Modeling in the Recognition of Emotion in Speech. In *Proceedings 10th Biannual IEEE Automatic Speech Recognition and Understanding Workshop, ASRU 2007*, pages 596–600, Kyoto, Japan. IEEE, IEEE. (acceptance rate: 43 %)
- 445) Schuller, B., Batliner, A., Seppi, D., Steidl, S., Vogt, T., Wagner, J., Devillers, L., Vidrascu, L., Amir, N., Kessous, L., and Aharonson, V. (2007a). The Relevance of Feature Type for the Automatic Classification of Emotional User States: Low Level Descriptors and Functionals. In *Proceedings INTERSPEECH 2007, 8th Annual Conference of the International Speech Communication Association*, pages 2253–2256, Antwerp, Belgium. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010), > 100 citations)
- 446) Schuller, B., Seppi, D., Batliner, A., Maier, A., and Steidl, S. (2007e). Towards More Reality in the Recognition of Emotional Speech. In *Proceedings 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2007*, volume IV, pages 941–944, Honolulu, HI. IEEE, IEEE. (acceptance rate: 46 %, IF* 1.16 (2010), > 100 citations)
- 447) Schuller, B., Wimmer, M., Arsić, D., Rigoll, G., and Radig, B. (2007g). Audiovisual Behavior Modeling by Combined Feature Spaces. In *Proceedings 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2007*, volume II, pages 733–736, Honolulu, HI. IEEE, IEEE. (acceptance rate: 46 %, IF* 1.16 (2010))
- 448) Schuller, B., Eyben, F., and Rigoll, G. (2007b). Fast and Robust Meter and Tempo Recognition for the Automatic Discrimination of Ballroom Dance Styles. In *Proceedings 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2007*, volume I, pages 217–220, Honolulu, HI. IEEE, IEEE. (acceptance rate: 46 %, IF* 1.16 (2010))
- 449) Vlasenko, B., Schuller, B., Wendemuth, A., and Rigoll, G. (2007a). Combining Frame and Turn-Level Information for Robust Recognition of Emotions within Speech. In *Proceedings INTERSPEECH 2007, 8th Annual Conference of the International Speech Communication Association*, pages 2249–2252, Antwerp, Belgium. ISCA, ISCA. (acceptance rate: 59 %, IF* 1.05 (2010))
- 450) Arsić, D., Hofmann, M., Schuller, B., and Rigoll, G. (2007a). Multi-Camera Person Tracking and Left Luggage Detection Applying Homographic Transformation. In Ferryman, J. M., editor, *Proceedings 10th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, PETS 2007, in association with ICCV 2007*, pages 55–62, Rio de Janeiro, Brazil. IEEE, IEEE. (acceptance rate: 24 %, IF* 5.05 (2010))
- 451) Batliner, A., Steidl, S., Schuller, B., Seppi, D., Vogt, T., Devillers, L., Vidrascu, L., Amir, N., Kessous, L., and Aharonson, V. (2007). The Impact of F0 Extraction Errors on the Classification of Prominence and Emotion. In *Proceedings 16th International Congress of Phonetic Sciences, ICPHS 2007*, pages 2201–2204, Saarbrücken, Germany. (acceptance rate: 66 %)
- 452) Eyben, F., Schuller, B., Reiter, S., and Rigoll, G. (2007). Wearable Assistance for the Ballroom-Dance Hobbyist – Holistic Rhythm Analysis and Dance-Style Classification. In *Proceedings 8th IEEE International Conference on Multimedia and Expo, ICME 2007*, pages 92–95, Beijing, China. IEEE, IEEE. (acceptance rate: 45 %, IF* 0.88 (2010))
- 453) Reiter, S., Schuller, B., and Rigoll, G. (2007). Hidden Conditional Random Fields for Meeting Segmentation. In *Proceedings 8th IEEE International Conference on Multimedia and Expo, ICME 2007*, pages 639–642, Beijing, China. IEEE, IEEE. (acceptance rate: 45 %, IF* 0.88 (2010))
- 454) Arsić, D., Schuller, B., and Rigoll, G. (2007b). Suspicious Behavior Detection In Public Transport by Fusion of Low-Level Video Descriptors. In *Proceedings 8th IEEE International Conference on Multimedia and Expo, ICME 2007*, pages 2018–2021, Beijing, China. IEEE, IEEE. (acceptance rate: 45 %, IF* 0.88 (2010))
- 455) Schuller, B. and Rigoll, G. (2006b). Timing Levels in Segment-Based Speech Emotion Recognition. In *Proceedings INTERSPEECH 2006, 9th International Conference on Spoken Language Processing, ICSLP*, pages 1818–1821, Pittsburgh, PA. ISCA, ISCA
- 456) Schuller, B., Köhler, N., Müller, R., and Rigoll, G. (2006c). Recognition of Interest in Human Conversational Speech. In *Proceedings INTERSPEECH 2006, 9th International Conference on Spoken Language Processing, ICSLP*, pages 793–796, Pittsburgh, PA. ISCA, ISCA
- 457) Schuller, B., Reiter, S., and Rigoll, G. (2006e). Evolutionary Feature Generation in Speech Emotion Recognition. In *Proceedings 7th IEEE International Conference on Multimedia and Expo, ICME 2006*, pages 5–8, Toronto, Canada. IEEE, IEEE.

- (acceptance rate: 51 %, IF* 0.88 (2010))
- 458) Schuller, B., Wallhoff, F., Arsić, D., and Rigoll, G. (2006g). Musical Signal Type Discrimination Based on Large Open Feature Sets. In *Proceedings 7th IEEE International Conference on Multimedia and Expo, ICME 2006*, pages 1089–1092, Toronto, Canada. IEEE, IEEE. (acceptance rate: 51 %, IF* 0.88 (2010))
- 459) Schuller, B., Arsić, D., Wallhoff, F., and Rigoll, G. (2006b). Emotion Recognition in the Noise Applying Large Acoustic Feature Sets. In *Proceedings 3rd International Conference on Speech Prosody, SP 2006*, pages 276–289, Dresden, Germany. ISCA, ISCA
- 460) Al-Hames, M., Zettl, S., Wallhoff, F., Reiter, S., Schuller, B., and Rigoll, G. (2006). A Two-Layer Graphical Model for Combined Video Shot And Scene Boundary Detection. In *Proceedings 7th IEEE International Conference on Multimedia and Expo, ICME 2006*, pages 261–264, Toronto, Canada. IEEE, IEEE. (acceptance rate: 51 %, IF* 0.88 (2010))
- 461) Arsić, D., Schenk, J., Schuller, B., Wallhoff, F., and Rigoll, G. (2006). Submotions for Hidden Markov Model Based Dynamic Facial Action Recognition. In *Proceedings 13th IEEE International Conference on Image Processing, ICIP 2006*, pages 673–676, Atlanta, GA. IEEE, IEEE. (acceptance rate: 41 %, IF* 0.65 (2010))
- 462) Batliner, A., Steidl, S., Schuller, B., Seppi, D., Laskowski, K., Vogt, T., Devillers, L., Vidrascu, L., Amir, N., Kessous, L., and Aharonson, V. (2006). Combining Efforts for Improving Automatic Classification of Emotional User States. In *Proceedings 5th Slovenian and 1st International Language Technologies Conference, ISLTC 2006*, pages 240–245, Ljubljana, Slovenia. Slovenian Language Technologies Society. (> 100 citations)
- 463) Reiter, S., Schuller, B., and Rigoll, G. (2006a). A combined LSTM-RNN-HMM-Approach for Meeting Event Segmentation and Recognition. In *Proceedings 31st IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2006*, volume 2, pages 393–396, Toulouse, France. IEEE, IEEE. (acceptance rate: 49 %, IF* 1.16 (2010))
- 464) Reiter, S., Schuller, B., and Rigoll, G. (2006b). Segmentation and Recognition of Meeting Events Using a Two-Layered HMM and a Combined MLP-HMM Approach. In *Proceedings 7th IEEE International Conference on Multimedia and Expo, ICME 2006*, pages 953–956, Toronto, Canada. IEEE, IEEE. (acceptance rate: 51 %, IF* 0.88 (2010))
- 465) Wallhoff, F., Schuller, B., Hawellek, M., and Rigoll, G. (2006). Efficient Recognition of Authentic Dynamic Facial Expressions on the FEEDTUM Database. In *Proceedings 7th IEEE International Conference on Multimedia and Expo, ICME 2006*, pages 493–496, Toronto, Canada. IEEE, IEEE. (acceptance rate: 51 %, IF* 0.88 (2010))
- 466) Schuller, B., Arsić, D., Wallhoff, F., Lang, M., and Rigoll, G. (2005a). Bioanalog Acoustic Emotion Recognition by Genetic Feature Generation Based on Low-Level-Descriptors. In *Proceedings International Conference on Computer as a Tool, EUROCON 2005*, volume 2, pages 1292–1295, Belgrade, Serbia and Montenegro. IEEE, IEEE
- 467) Schuller, B., Schmitt, B. J. B., Arsić, D., Reiter, S., Lang, M., and Rigoll, G. (2005f). Feature Selection and Stacking for Robust Discrimination of Speech, Monophonic Singing, and Polyphonic Music. In *Proceedings 6th IEEE International Conference on Multimedia and Expo, ICME 2005*, pages 840–843, Amsterdam, The Netherlands. IEEE, IEEE. (acceptance rate: 23 %, IF* 0.88 (2010))
- 468) Schuller, B., Reiter, S., Müller, R., Al-Hames, M., Lang, M., and Rigoll, G. (2005d). Speaker Independent Speech Emotion Recognition by Ensemble Classification. In *Proceedings 6th IEEE International Conference on Multimedia and Expo, ICME 2005*, pages 864–867, Amsterdam, The Netherlands. IEEE, IEEE. (acceptance rate: 23 %, IF* 0.88 (2010), > 100 citations)
- 469) Schuller, B., Villar, R. J., Rigoll, G., and Lang, M. (2005g). Meta-Classifiers in Acoustic and Linguistic Feature Fusion-Based Affect Recognition. In *Proceedings 30th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2005*, volume I, pages 325–328, Philadelphia, PA. IEEE, IEEE. (acceptance rate: 52 %, IF* 1.16 (2010))
- 470) Arsić, D., Wallhoff, F., Schuller, B., and Rigoll, G. (2005a). Bayesian Network Based Multi Stream Fusion for Automated Online Video Surveillance. In *Proceedings International Conference on Computer as a Tool, EUROCON 2005*, volume 2, pages 995–998, Belgrade, Serbia and Montenegro. IEEE, IEEE
- 471) Arsić, D., Wallhoff, F., Schuller, B., and Rigoll, G. (2005c). Vision-Based Online Multi-Stream Behavior Detection Applying Bayesian Networks. In *Proceedings 6th IEEE International Conference on Multimedia and Expo, ICME 2005*, pages 1354–1357, Amsterdam, The Netherlands. IEEE, IEEE. (acceptance rate: 23 %, IF* 0.88 (2010))
- 472) Arsić, D., Wallhoff, F., Schuller, B., and Rigoll, G. (2005b). Video Based Online Behavior Detection Using Probabilistic Multi-Stream Fusion. In *Proceedings 12th IEEE International Conference on Image Processing, ICIP 2005*, volume 2, pages 606–609, Genova, Italy. IEEE, IEEE. (acceptance rate: about 45 %, IF* 0.65 (2010))
- 473) Müller, R., Schreiber, S., Schuller, B., and Rigoll, G. (2005). A System Structure for Multimodal Emotion Recognition in Meeting Environments. In Renals, S. and Bengio, S., editors, *Proceedings 2nd International Workshop on Machine Learning for Multimodal Interaction, MLMI 2005*, Edinburgh, UK. 2 pages
- 474) Wallhoff, F., Schuller, B., and Rigoll, G. (2005b). Speaker Identification – Comparing Linear Regression Based Adaptation and Acoustic High-Level Features. In *Proceedings 31. Jahrestagung für Akustik, DAGA 2005*, pages 221–222, Munich, Germany. DEGA, DEGA
- 475) Wallhoff, F., Arsić, D., Schuller, B., Stadermann, J., Störmer, A., and Rigoll, G. (2005a). Hybrid Profile Recognition on the Mugshot Database. In *Proceedings International Conference on Computer as a Tool, EUROCON 2005*, volume 2, pages 1405–1408, Belgrade, Serbia and Montenegro. IEEE, IEEE
- 476) Schuller, B., Rigoll, G., and Lang, M. (2004d). Multimodal Music Retrieval for Large Databases. In *Proceedings 5th IEEE International Conference on Multimedia and Expo, ICME 2004*, volume 2, pages 755–758, Taipei, Taiwan. IEEE, IEEE. Special Session Novel Techniques for Browsing in Large Multimedia Collections (acceptance rate: 30 %, IF* 0.88 (2010))
- 477) Schuller, B., Rigoll, G., and Lang, M. (2004b). Discrimination of Speech and Monophonic Singing in Continuous Audio Streams Applying Multi-Layer Support Vector Machines. In *Proceedings 5th IEEE International Conference on Multimedia and Expo, ICME 2004*, volume 3, pages 1655–1658, Taipei, Taiwan. IEEE, IEEE. (acceptance rate: 30 %, IF* 0.88 (2010))
- 478) Schuller, B., Rigoll, G., and Lang, M. (2004c). Emotion Recognition in the Manual Interaction with Graphical User Interfaces. In *Proceedings 5th IEEE International Conference on Multimedia and Expo, ICME 2004*, volume 2, pages 1215–1218, Taipei, Taiwan. IEEE, IEEE. (acceptance rate: 30 %, IF* 0.88 (2010))
- 479) Schuller, B., Müller, R., Rigoll, G., and Lang, M. (2004a). Applying Bayesian Belief Networks in Approximate String Matching for Robust Keyword-based Retrieval. In *Proceedings 5th IEEE International Conference on Multimedia and Expo, ICME 2004*, volume 3, pages 1999–2002, Taipei, Taiwan. IEEE, IEEE. (acceptance rate: 30 %, IF* 0.88 (2010))
- 480) Schuller, B., Rigoll, G., and Lang, M. (2004e). Speech Emotion Recognition Combining Acoustic Features and Linguistic Information in a Hybrid Support Vector Machine-Belief Network Architecture. In *Proceedings 29th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2004*, volume I, pages 577–580, Montreal, Canada. IEEE, IEEE. (acceptance rate: 54 %, IF* 1.16 (2010), > 200 citations)
- 481) Müller, R., Schuller, B., and Rigoll, G. (2004b). Enhanced Ro-

- bustness in Speech Emotion Recognition Combining Acoustic and Semantic Analysis. In *Proceedings HUMAINE Workshop From Signals to Signs of Emotion and Vice Versa*, page 2 pages, Santorini, Greece. HUMAINE
- 482) Müller, R., Schuller, B., and Rigoll, G. (2004a). Belief Networks in Natural Language Processing for Improved Speech Emotion Recognition. In Bengio, S. and Bourlard, H., editors, *Proceedings 1st International Workshop on Machine Learning for Multimodal Interaction, MLMI 2004*, Martigny, Switzerland. 1 page
- 483) Schuller, B., Rigoll, G., and Lang, M. (2003d). Sprachliche Emotionserkennung im Fahrzeug. In *Proc. 45. Fachausschusssitzung Anthropotechnik, Entscheidungsunterstützung für die Fahrzeug- und Prozessführung*, volume DGLR Bericht 2003-04, pages 227–240, Neubiberg, Germany. Deutsche Gesellschaft für Luft- und Raumfahrt, Deutsche Gesellschaft für Luft- und Raumfahrt
- 484) Schuller, B., Rigoll, G., and Lang, M. (2003a). Hidden Markov Model-based Speech Emotion Recognition. In *Proceedings 4th IEEE International Conference on Multimedia and Expo, ICME 2003*, volume I, pages 401–404, Baltimore, MD. IEEE, IEEE. (acceptance rate: 58 %, IF* 0.88 (2010))
- 485) Schuller, B., Zobl, M., Rigoll, G., and Lang, M. (2003e). A Hybrid Music Retrieval System using Belief Networks to Integrate Queries and Contextual Knowledge. In *Proceedings 4th IEEE International Conference on Multimedia and Expo, ICME 2003*, volume I, pages 57–60, Baltimore, MD. IEEE, IEEE. (acceptance rate: 58 %, IF* 0.88 (2010))
- 486) Schuller, B., Rigoll, G., and Lang, M. (2003c). HMM-Based Music Retrieval Using Stereophonic Feature Information and Framelength Adaptation. In *Proceedings 4th IEEE International Conference on Multimedia and Expo, ICME 2003*, volume II, pages 713–716, Baltimore, MD. IEEE, IEEE. (acceptance rate: 58 %, IF* 0.88 (2010))
- 487) Schuller, B., Rigoll, G., and Lang, M. (2003b). Hidden Markov Model-based Speech Emotion Recognition. In *Proceedings 28th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2003*, volume II, pages 1–4, Hong Kong, China. IEEE, IEEE. (acceptance rate: 54 %, IF* 1.16 (2010), > 300 citations)
- 488) Zobl, M., Geiger, M., Schuller, B., Rigoll, G., and Lang, M. (2003). A Realtime System for Hand-Gesture Controlled Operation of In-Car Devices. In *Proceedings 4th IEEE International Conference on Multimedia and Expo, ICME 2003*, volume III, pages 541–544, Baltimore, MD. IEEE, IEEE. (acceptance rate: 58 %, IF* 0.88 (2010))
- 489) Schuller, B. (2002). Towards intuitive speech interaction by the integration of emotional aspects. In *Proceedings IEEE International Conference on Systems, Man and Cybernetics, SMC 2002*, volume 6, Yasmine Hammamet, Tunisia. IEEE, IEEE. 6 pages
- 490) Schuller, B., Althoff, F., McGlaun, G., Lang, M., and Rigoll, G. (2002a). Towards Automation of Usability Studies. In *Proceedings IEEE International Conference on Systems, Man and Cybernetics, SMC 2002*, volume 5, Yasmine Hammamet, Tunisia. IEEE, IEEE. 6 pages
- 491) Schuller, B., Lang, M., and Rigoll, G. (2002c). Multimodal Emotion Recognition in Audiovisual Communication. In *Proceedings 3rd IEEE International Conference on Multimedia and Expo, ICME 2002*, volume 1, pages 745–748, Lausanne, Switzerland. IEEE, IEEE. (acceptance rate: 50 %, IF* 0.88 (2010))
- 492) Schuller, B., Lang, M., and Rigoll, G. (2002b). Automatic Emotion Recognition by the Speech Signal. In *Proceedings 6th World Multiconference on Systemics, Cybernetics and Informatics, SCI 2002*, volume IX, pages 367–372, Orlando, FL. SCI, SCI
- 493) Schuller, B. and Lang, M. (2002). Integrative rapid-prototyping for multimodal user interfaces. In *Proceedings USEWARE 2002, Mensch – Maschine – Kommunikation /Design*, volume VDI report #1678, pages 279–284, Darmstadt, Germany. VDI, VDI-Verlag
- 494) Althoff, F., Geiss, K., McGlaun, G., Schuller, B., and Lang, M. (2002a). Experimental Evaluation of User Errors at the Skill-Based Level in an Automotive Environment. In *Proceedings International Conference on Human Factors in Computing Systems, CHI 2002*, pages 782–783, Minneapolis, MN. ACM, ACM. (acceptance rate: 15 %, IF* 6.37 (2010))
- 495) Althoff, F., McGlaun, G., Schuller, B., Lang, M., and Rigoll, G. (2002b). Evaluating Misinterpretations during Human-Machine Communication in Automotive Environments. In *Proceedings 6th World Multiconference on Systemics, Cybernetics and Informatics, SCI 2002*, volume VII, Orlando, FL. SCI, SCI. 5 pages
- 496) McGlaun, G., Althoff, F., Schuller, B., and Lang, M. (2002). A new technique for adjusting distraction moments in multitasking non-field usability tests. In *Proceedings International Conference on Human Factors in Computing Systems, CHI 2002*, pages 666–667, Minneapolis, MN. ACM, ACM. (acceptance rate: 15 %, IF* 6.37 (2010))
- 497) Schuller, B., Althoff, F., McGlaun, G., and Lang, M. (2001). Navigation in virtual worlds via natural speech. In Stephanidis, C., editor, *Proceedings 9th International Conference on Human-Computer Interaction, HCI 2001*, pages 19–21, New Orleans, LA. Lawrence Erlbaum
- 498) Althoff, F., McGlaun, G., Schuller, B., Morguet, P., and Lang, M. (2001). Using Multimodal Interaction to Navigate in Arbitrary Virtual VRML Worlds. In *Proceedings 3rd International Workshop on Perceptive User Interfaces, PUI 2001*, Orlando, FL. ACM, ACM. 8 pages (acceptance rate: 29 %)

D) PATENTS (4)

- 499) Schuller, B., Weninger, F., Kirst, C., and Grosche, P. (2013k). Apparatus and Method for Improving a Perception of a Sound Signal. WO2015070918 A1, Huawei Technologies Co Ltd, Technische Universität München, international patent, pending
- 500) Joder, C., Weninger, F., Schuller, B., and Virette, D. (2012c). Method for Determining a Dictionary of Base Components from an Audio Signal. EP73149, Huawei Technologies Co Ltd, Technische Universität München, European/US patent, pending
- 501) Joder, C., Weninger, F., Schuller, B., and Virette, D. (2012b). Method and Device for Reconstructing a Target Signal from a Noisy Input Signal. US9536538 (B2), Huawei Technologies Co Ltd, Technische Universität München, US patent, granted
- 502) Burkhardt, F. and Schuller, B. (2009). Method and system for training speech processing devices. EP2325836, Deutsche Telekom AG, Technische Universität München, European patent, pending

E) OTHER

Theses (3):

- 503) Schuller, B. (2012a). *Intelligent Audio Analysis – Speech, Music, and Sound Recognition in Real-Life Conditions*. Habilitation thesis, Technische Universität München, Munich, Germany. 313 pages
- 504) Schuller, B. (2006). *Automatische Emotionserkennung aus sprachlicher und manueller Interaktion*. Doctoral thesis, Technische Universität München, Munich, Germany. 244 pages
- 505) Schuller, B. (1999). *Automatisches Verstehen gesprochener mathematischer Formeln*. Diploma thesis, Technische Universität München, Munich, Germany

Editorials / Edited Volumes (50):

- 506) Schuller, B. (2017c). Editorial: Transactions on Affective Computing - Challenges and Chances. *IEEE Transactions on*

- Affective Computing*, 8(1):1–2. (IF: 3.466, 5-year IF: 3.871 (2013))
- 507) Schuller, B. W., Paletta, L., Robinson, P., Sabouret, N., and Yannakakis, G. N. (2017e). Guest Editorial: Computational Intelligence in Serious Digital Games. *IEEE Transactions on Computational Intelligence and AI in Games, Special Issue on Computational Intelligence in Serious Digital Games*. (IF: 1.481 (2014))
- 508) Squartini, S., Schuller, B., Uncini, A., and Ting, C.-K. (2017). Guest Editorial: Computational Intelligence for End-to-End Audio Processing. *IEEE Transaction on Emerging Topics in Computational Intelligence, Special Issue of Computational Intelligence for End-to-End Audio Processing*. to appear
- 509) Veselkov, K. and Schuller, B. (2017). Guest Editorial: Translational data analytics and health informatics. *Methods, Special Issue on on Translational data analytics and health informatics*. to appear (IF: 3.503, 5-year IF: 3.789 (2015)) xxx
- 510) Schuller, B. (2016d). Editorial: Transactions on Affective Computing - Changes and Continuance. *IEEE Transactions on Affective Computing*, 7(1):1–2. (IF: 3.466, 5-year IF: 3.871 (2013))
- 511) Devillers, L., Schuller, B., Mower Provost, E., Robinson, P., Mariani, J., and Delaborde, A., editors (2016). *Proceedings of the 1st International Workshop on ETHics In Corpus Collection, Annotation and Application (ETHI-CA² 2016)*, Portoroz, Slovenia. ELRA, ELRA. Satellite of the 10th Language Resources and Evaluation Conference (LREC 2016)
- 512) Ringeval, F., Schuller, B., Valstar, M., Gratch, J., Cowie, R., and Pantic, M. (2016b). Introduction to the Special Section on Multimedia Computing and Applications of Socioaffective Behaviors in the Wild. *ACM Transactions on Multimedia Computing, Communications and Applications*. editorial, to appear (IF: 0.904 (2013))
- 513) Sánchez-Rada, J. and Schuller, B., editors (2016). *Proceedings of the 6th International Workshop on Emotion and Sentiment Analysis (ESA 2016)*, Portoroz, Slovenia. ELRA, ELRA. Satellite of the 10th Language Resources and Evaluation Conference (LREC 2016)
- 514) Soleymani, M., Schuller, B., and Chang, S.-F. (2016). Introduction to the Special Issue on Multimodal Sentiment Analysis and Mining in the Wild. *Image and Vision Computing, Special Issue on Multimodal Sentiment Analysis and Mining in the Wild*, 34. to appear (IF: 1.587, 5-year IF: 2.384 (2014))
- 515) Valstar, M., Gratch, J., Schuller, B., Ringeval, F., Cowie, R., and Pantic, M., editors (2016a). *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge, AVEC'16, co-located with MM 2016*, Amsterdam, The Netherlands. ACM, ACM
- 516) Hartmann, K., Siegert, I., Schuller, B., Morency, L.-P., Salah, A. A., and Böck, R., editors (2015b). *Proceedings of the Workshop on Emotion Representations and Modelling for Companion Systems, ERM4CT 2015*, Seattle, WA. ACM, ACM. held in conjunction with the 17th ACM International Conference on Multimodal Interaction, ICMI 2015
- 517) Hartmann, K., Siegert, I., Schuller, B., Morency, L.-P., Salah, A. A., and Böck, R. (2015a). ERM4CT 2015 - Workshop on Emotion Representations and Modelling for Companion Systems. In Hartmann, K., Siegert, I., Schuller, B., Morency, L.-P., Salah, A. A., and Böck, R., editors, *Proceedings of the Workshop on Emotion Representations and Modelling for Companion Systems, ERM4CT 2015*, Seattle, WA. ACM, ACM. 2 pages, held in conjunction with the 17th ACM International Conference on Multimodal Interaction, ICMI 2015
- 518) Cambria, E., Schuller, B., Xia, Y., and White, B. (2016b). New Avenues in Knowledge Bases for Natural Language Processing. *Knowledge-Based Systems, Special Issue on New Avenues in Knowledge Bases for Natural Language Processing*, C(108):1–4. editorial (IF: 3.058, 5-year IF: 2.920 (2013))
- 519) Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Burkhardt, F., and van Son, R. (2015f). Introduction to the Special Issue on Next Generation Computational Paralinguistics. *Computer Speech and Language, Special Issue on Next Generation Computational Paralinguistics*, 29(1):98–99. editorial (acceptance rate: 23 %, IF: 1.812, 5-year IF: 1.776 (2013))
- 520) Paletta, L., Schuller, B. W., Robinson, P., and Sabouret, N. (2015b). IDGEI 2015: 3rd international workshop on intelligent digital games for empowerment and inclusion. In *Proceedings of the 20th ACM International Conference on Intelligent User Interfaces, IUI 2015*, pages 450–452, Atlanta, GA. ACM, ACM
- 521) Paletta, L., Schuller, B., Robinson, P., and Sabouret, N., editors (2015a). *IDGEI 2015 - 3rd International Workshop on Intelligent Digital Games for Empowerment and Inclusion*, Atlanta, GA. ACM, ACM. held in conjunction with the 20th International Conference on Intelligent User Interfaces, IUI 2015
- 522) Ringeval, F., Schuller, B., Valstar, M., Cowie, R., and Pantic, M., editors (2015e). *Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge, AVEC'15, co-located with MM 2015*, Brisbane, Australia. ACM, ACM
- 523) Ringeval, F., Schuller, B., Valstar, M., Cowie, R., and Pantic, M. (2015c). AVEC 2015 Chairs' Welcome. In Ringeval, F., Schuller, B., Valstar, M., Cowie, R., and Pantic, M., editors, *Proceedings of the 5th International Workshop on Audio/Visual Emotion Challenge, AVEC'15, co-located with the 23rd ACM International Conference on Multimedia, MM 2015*, page iii, Brisbane, Australia. ACM, ACM
- 524) Schuller, B., Steidl, S., Batliner, A., Vinciarelli, A., Burkhardt, F., and van Son, R. (2015f). Introduction to the Special Issue on Next Generation Computational Paralinguistics. *Computer Speech and Language, Special Issue on Next Generation Computational Paralinguistics*, 29(1):98–99. editorial (acceptance rate: 23 %, IF: 1.812, 5-year IF: 1.776 (2013))
- 525) Schuller, B., Falk, T. H., Parsa, V., and Nöth, E. (2015a). Introduction to the Special Issue on Atypical Speech & Voices: Corpora, Classification, Coaching & Conversion. *EURASIP Journal on Audio, Speech, and Music Processing, Special Issue on Atypical Speech & Voices: Corpora, Classification, Coaching & Conversion*. to appear (IF: 0.630 (2012))
- 526) Schuller, B., Buitelaar, P., Devillers, L., Pelachaud, C., Declerck, T., Batliner, A., Rosso, P., and Gaines, S., editors (2014a). *Proceedings of the 5th International Workshop on Emotion Social Signals, Sentiment & Linked Open Data (ES³ LOD 2014)*, Reykjavik, Iceland. ELRA, ELRA. Satellite of the 9th Language Resources and Evaluation Conference (LREC 2014), (acceptance rate: 72 %)
- 527) Gunes, H., Schuller, B., Celiktutan, O., Sariyanidi, E., and Eyben, F., editors (2014b). *Proceedings of the Personality Mapping Challenge & Workshop (MAPTRAITS 2014)*, Istanbul, Turkey. ACM, ACM. Satellite of the 16th ACM International Conference on Multimodal Interaction (ICMI 2014)
- 528) Gunes, H., Schuller, B., Celiktutan, O., Sariyanidi, E., and Eyben, F. (2014a). MAPTRAITS'14 Foreword. In *Proceedings of the Personality Mapping Challenge & Workshop (MAPTRAITS 2014), Satellite of the 16th ACM International Conference on Multimodal Interaction (ICMI 2014)*, page iii, Istanbul, Turkey. ACM, ACM
- 529) Hartmann, K., Böck, R., Schuller, B., and Scherer, K. R., editors (2013c). *Proceedings of the 2nd Workshop on Emotion representation and modelling in Human-Computer-Interaction-Systems, ERM4HCI 2014*, Istanbul, Turkey. ACM, ACM. held in conjunction with the 16th ACM International Conference on Multimodal Interaction, ICMI 2014
- 530) Hartmann, K., Scherer, K. R., Schuller, B., and Böck, R. (2014b). Welcome to the ERM4HCI 2014! In Hartmann, K., Böck, R., Schuller, B., and Scherer, K., editors, *Proceedings of the 2nd Workshop on Emotion representation and modelling in Human-Computer-Interaction-Systems, ERM4HCI 2014*, page iii, Istanbul, Turkey. ACM, ACM. held in conjunction with the

- 16th ACM International Conference on Multimodal Interaction, ICMI 2014
- 531) Paletta, L., Schuller, B. W., Robinson, P., and Sabouret, N. (2014c). IDGEI 2014: 2nd international workshop on intelligent digital games for empowerment and inclusion. In *Proceedings of the companion publication of the 19th international conference on Intelligent User Interfaces, IUI Companion 2014*, pages 49–50, Haifa, Israel. ACM, ACM
- 532) Paletta, L., Schuller, B., Robinson, P., and Sabouret, N., editors (2014a). *IDGEI 2014 - 2nd International Workshop on Intelligent Digital Games for Empowerment and Inclusion*, Haifa, Israel. ACM, ACM. held in conjunction with the 19th International Conference on Intelligent User Interfaces, IUI 2014
- 533) Paletta, L., Schuller, B., Robinson, P., and Sabouret, N., editors (2014b). *Proceedings of the 2nd International Workshop on Digital Games for Empowerment and Inclusion, IDGEI 2014*, Haifa, Israel. ACM, ACM. held in conjunction with the 19th International Conference on Intelligent User Interfaces, IUI 2014
- 534) Valstar, M., Schuller, B., Krajewski, J., Cowie, R., and Pantic, M. (2014a). AVEC 2014: the 4th International Audio/Visual Emotion Challenge and Workshop. In *Proceedings of the 22nd ACM International Conference on Multimedia, MM 2014*, pages 1243–1244, Orlando, FL. ACM, ACM
- 535) Salah, A. A., Cohn, J., Schuller, B., Aran, O., Morency, L.-P., and Cohen, P. R. (2014). ICMI 2014 Chairs' Welcome. In *Proceedings of the 16th ACM International Conference on Multimodal Interaction, ICMI*, pages iii–v, Istanbul, Turkey. ACM, ACM. editorial. (acceptance rate: 40 %, IF* 1.77 (2010))
- 536) Schuller, B., Paletta, L., and Sabouret, N. (2013d). Intelligent Digital Games for Empowerment and Inclusion – An Introduction. In Schuller, B., Paletta, L., and Sabouret, N., editors, *Proceedings 1st International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI 2013) held in conjunction with the 8th Foundations of Digital Games 2013 (FDG)*, Chania, Greece. ACM, SASDG. 2 pages
- 537) Schuller, B., Valstar, M., Cowie, R., Krajewski, J., and Pantic, M., editors (2013j). *Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge*, Barcelona, Spain. ACM, ACM. held in conjunction with the 21st ACM international conference on Multimedia, ACM MM 2013
- 538) Epps, J., Chen, F., Oviatt, S., Mase, K., Sears, A., Jokinen, K., and Schuller, B. (2013). ICMI 2013 Chairs' Welcome. In *Proceedings of the 15th ACM International Conference on Multimodal Interaction, ICMI*, pages 3–4, Sydney, Australia. ACM, ACM. editorial. (acceptance rate: 37 %, IF* 1.77 (2010))
- 539) Hartmann, K., Böck, R., Becker-Asano, C., Gratch, J., Schuller, B., and Scherer, K. R., editors (2013b). *Proceedings of the Workshop on Emotion representation and modelling in Human-Computer-Interaction-Systems, ERM4HCI 2013*, Sydney, Australia. ACM, ACM. held in conjunction with the 15th ACM International Conference on Multimodal Interaction, ICMI 2013
- 540) Hartmann, K., Böck, R., Becker-Asano, C., Gratch, J., Schuller, B., and Scherer, K. R. (2013a). ERM4HCI 2013 - The 1st Workshop on Emotion Representation and Modelling in Human-Computer-Interaction-Systems. In Hartmann, K., Böck, R., Becker-Asano, C., Gratch, J., Schuller, B., and Scherer, K. R., editors, *Proceedings of the Workshop on Emotion representation and modelling in Human-Computer-Interaction-Systems, ERM4HCI 2013*, Sydney, Australia. ACM, ACM. held in conjunction with the 15th ACM International Conference on Multimodal Interaction, ICMI 2013
- 541) Müller, M., Narayanan, S. S., and Schuller, B., editors (2013). *Report from Dagstuhl Seminar 13451 – Computational Audio Analysis*, volume 3 of *Dagstuhl Reports*, Dagstuhl, Germany. Schloss Dagstuhl, Leibniz-Zentrum fuer Informatik, Dagstuhl Publishing. 28 pages
- 542) Paletta, L., Itti, L., Schuller, B., and Fang, F., editors (2013). *Proceedings of the 6th International Symposium on Attention in Cognitive Systems 2013, ISACS 2013*, LNAI, Beijing, P.R. China. arxiv.org, Springer. held in conjunction with the 23rd International Joint Conference on Artificial Intelligence, IJCAI 2013
- 543) Schuller, B., Valstar, M., Cowie, R., and Pantic, M. (2012f). AVEC 2012: the continuous audio/visual emotion challenge – an introduction. In Morency, L.-P., Bohus, D., Aghajan, H. K., Cassell, J., Nijholt, A., and Epps, J., editors, *Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI*, pages 361–362, Santa Monica, CA. ACM, ACM. (acceptance rate: 36 %, IF* 1.77 (2010))
- 544) Schuller, B., Steidl, S., and Batliner, A. (2013f). Introduction to the Special Issue on Paralinguistics in Naturalistic Speech and Language. *Computer Speech and Language, Special Issue on Paralinguistics in Naturalistic Speech and Language*, 27(1):1–3. editorial (acceptance rate: 36 %, IF: 1.812, 5-year IF: 1.776 (2013))
- 545) Gunes, H. and Schuller, B. (2013b). Introduction to the Special Issue on Affect Analysis in Continuous Input. *Image and Vision Computing, Special Issue on Affect Analysis in Continuous Input*, 31(2):118–119. (IF: 1.723, 5-Year IF: 1.743 (2011))
- 546) Schuller, B., Douglas-Cowie, E., and Batliner, A. (2012a). Guest Editorial: Special Section on Naturalistic Affect Resources for System Building and Evaluation. *IEEE Transactions on Affective Computing, Special Issue on Naturalistic Affect Resources for System Building and Evaluation*, 3(1):3–4. (IF: 3.466, 5-year IF: 3.871 (2013))
- 547) Cambria, E., Hussain, A., Schuller, B., and Howard, N. (2014). Guest Editorial Introduction Affective Neural Networks and Cognitive Learning Systems for Big Data Analysis. *Neural Networks, Special Issue on Affective Neural Networks and Cognitive Learning Systems for Big Data Analysis*, 58(10):1–3. editorial (IF: 2.182, 5-year IF: 2.477 (2012))
- 548) Cambria, E., Schuller, B., Liu, B., Wang, H., and Havasi, C. (2013a). Guest Editor's Introduction: Knowledge-based Approaches to Concept-Level Sentiment Analysis. *IEEE Intelligent Systems Magazine, Special Issue on Statistical Approaches to Concept-Level Sentiment Analysis*, 28(2):12–14. (IF: 2.570, 5-year IF: 2.632 (2010))
- 549) Cambria, E., Schuller, B., Liu, B., Wang, H., and Havasi, C. (2013b). Guest Editor's Introduction: Statistical Approaches to Concept-Level Sentiment Analysis. *IEEE Intelligent Systems Magazine, Special Issue on Statistical Approaches to Concept-Level Sentiment Analysis*, 28(3):6–9. (IF: 2.570, 5-year IF: 2.632 (2010))
- 550) Devillers, L., Schuller, B., Batliner, A., Rosso, P., Douglas-Cowie, E., Cowie, R., and Pelachaud, C., editors (2012). *Proceedings of the 4th International Workshop on EMOTION SENTIMENT & SOCIAL SIGNALS 2012 (ES 2012) – Corpora for Research on Emotion, Sentiment & Social Signals*, Istanbul, Turkey. ELRA, ELRA. held in conjunction with LREC 2012
- 551) Epps, J., Cowie, R., Narayanan, S., Schuller, B., and Tao, J. (2012). Editorial Emotion and Mental State Recognition from Speech. *EURASIP Journal on Advances in Signal Processing, Special Issue on Emotion and Mental State Recognition from Speech*, 2012(15). 2 pages (acceptance rate: 38 %, IF: 1.012, 5-Year IF: 1.136 (2010))
- 552) Squartini, S., Schuller, B., and Hussain, A. (2012). Cognitive and Emotional Information Processing for Human-Machine Interaction. *Cognitive Computation, Special Issue on Cognitive and Emotional Information Processing for Human-Machine Interaction*, 4(4):383–385. (IF: 1.000 (2011))
- 553) Schuller, B., Steidl, S., and Batliner, A. (2011c). Introduction to the Special Issue on Sensing Emotion and Affect – Facing Realism in Speech Processing. *Speech Communication, Special Issue Sensing Emotion and Affect - Facing Realism in Speech Processing*, 53(9/10):1059–1061. (acceptance rate: 38 %, IF: 1.267, 5-year IF: 1.526 (2011))

- 554) Schuller, B., Valstar, M., Cowie, R., and Pantic, M., editors (2011d). *Proceedings of the First International Audio/Visual Emotion Challenge and Workshop, AVEC 2011*, volume 6975, Part II of *Lecture Notes on Computer Science (LNCS)*, Memphis, TN. HUMAINE Association, Springer. held in conjunction with the International HUMAINE Association Conference on Affective Computing and Intelligent Interaction 2011, ACII 2011
- 555) Devillers, L., Schuller, B., Cowie, R., Douglas-Cowie, E., and Batliner, A., editors (2010). *Proceedings of the 3rd International Workshop on EMOTION: Corpora for Research on Emotion and Affect*, Valletta, Malta. ELRA, ELRA. Satellite of 7th International Conference on Language Resources and Evaluation (LREC 2010) (acceptance rate: 69%, IF* 1.37 (2010))

Abstracts in Journals (6):

- 556) Pokorny, F. B., Schuller, B. W., Tomantschger, I., Zhang, D., Einspieler, C., and Marschik, P. B. (2016f). Intelligent pre-linguistic vocalisation analysis: a promising novel approach for the earlier identification of Rett syndrome. *Wiener Medizinische Wochenschrift*, 166(11–12):382–383. Rett Syndrome – RTT50.1 (IF: 0.56 (2015))
- 557) Schuller, B. (2013a). Approaching Cross-Audio Computer Audition. *Dagstuhl Reports*, 3(11):22–22
- 558) Metze, F., Anguera, X., Ewert, S., Gemmeke, J., Kolossa, D., Provost, E. M., Schuller, B., and Serrà, J. (2013). Learning of Units and Knowledge Representation. *Dagstuhl Reports*, 3(11):13–13
- 559) Schuller, B., Fridenzon, S., Tal, S., Marchi, E., Batliner, A., and Golan, O. (2012b). Learning the Acoustics of Autism-Spectrum Emotional Expressions – A Children’s Game? *Neuropsychiatrie de l’Enfance et de l’Adolescence*, 60(5S):32. invited contribution
- 560) Schuller, B. (2011b). Next Gen Music Analysis: Some Inspirations from Speech. *Dagstuhl Reports*, 1(1):93–93
- 561) Ewert, S., Goto, M., Grosche, P., Kaiser, F., Yoshii, K., Kurth, F., Mauch, M., Müller, M., Peeters, G., Richard, G., and Schuller, B. (2011). Signal Models for and Fusion of Multimodal Information. *Dagstuhl Reports*, 1(1):97–97

Abstracts in Conference/Challenge Proceedings (22)

- 562) Cummins, N., Hantke, S., Schnieder, S., Krajewski, J., and Schuller, B. (2017a). Classifying the Context and Emotion of Dog Barks: A Comparison of Acoustic Feature Representations. In *Proceedings Pre-Conference on Affective Computing 2017 SAS Annual Conference*, pages 14–15, Boston, MA. Society for Affective Science (SAS)
- 563) Guo, J., Qian, K., Schuller, B., and Matsuoaka, S. (2017a). GPU Processing Accelerates Training Autoencoders for Bird Sounds Data. In *Proceedings 2017 GPU Technology Conference (GTC)*, San Jose, CA. NVIDIA. 1 page, to appear
- 564) Pokorny, F. B., Schuller, B. W., Bartl-Pokorny, K. D., Zhang, D., Einspieler, C., and Marschik, P. B. (2017). In a bad mood? Automatic audio-based recognition of infant fussing and crying in video-taped vocalisations. In *Proceedings 2017 13th International Infant Cry Workshop*, Castel Noarna-Rovereto, Italy. 1 page
- 565) Schuller, B. (2016e). Engage to Empower: Emotionally Intelligent Computer Games & Robots for Autistic Children. In Rynkiewicz, A. and Grabowski, K., editors, *Proceedings “The world innovations combining medicine, engineering and technology in autism diagnosis and therapy”*, pages 65–67, Rzeszow, Poland. SOLIS RADIUS
- 566) Schuller, B. (2016g). Zaangazowanie aby wzmacniac kompetencje: emocjonalnie inteligentne roboty i gry komputerow dla dzieci z autyzmem. In Rynkiewicz, A. and Grabowski,

- K., editors, *Proceedings “Swiatowe innowacje laczące medycyne, inzynierie oraz technologie w diagnozowaniu i terapii autyzmu”*, pages 62–64, Rzeszow, Poland. SOLIS RADIUS
- 567) Pokorny, F. B., Schuller, B. W., Bartl-Pokorny, K. D., Einspieler, C., and Marschik, P. B. (2016c). Contributing to the early identification of Rett syndrome: Automated analysis of vocalisations from the pre-regression period. In *Proceedings Symposium of the Austrian Physiological Society 2016*, Graz, Austria. ÖPG, ÖPG. 1 page, Best Poster award 3rd place
- 568) Pokorny, F. B., Schuller, B. W., Peharz, R., Pernkopf, F., Bartl-Pokorny, K. D., Einspieler, C., and Marschik, P. B. (2016d). Contributing to the early identification of neurodevelopmental disorders: The retrospective analysis of pre-linguistic vocalisations in home video material. In *Proceedings IX Congreso Internacional y XIV Nacional de Psicología Clínica*, Santander, Spain. 1 page
- 569) Pokorny, F. B., Schuller, B. W., Peharz, R., Pernkopf, F., Bartl-Pokorny, K. D., Einspieler, C., and Marschik, P. B. (2016e). Retrospektive Analyse frühkindlicher Lautäußerungen in „Home-Videos“: Ein signalanalytischer Ansatz zur Früherkennung von Entwicklungsstörungen. In *Proceedings 42. Österreichische Linguistiktagung, ÖLT*, Graz, Austria. 1 page
- 570) Rudovic, O., Evers, V., Pantic, M., Schuller, B., and Petrovic, S. (2016a). DE-ENIGMA Robots: Playfully Empowering Children with Autism. In *Proceedings XI Autism-Europe International Congress*, Edinburgh, Scotland. Autism Europe, The National Autistic Society. 1 page
- 571) Rudovic, O., Lee, J., Schuller, B., and Picard, R. (2016b). Automated Measurement of Engagement of Children with Autism Spectrum Conditions during Human-Robot Interaction (HRI). In *Proceedings XI Autism-Europe International Congress*, Edinburgh, Scotland. Autism Europe, The National Autistic Society. 1 page
- 572) Schuller, B. (2015f). Modelling User Affect and Sentiment in Intelligent User Interfaces: a Tutorial Overview. In *Proceedings of the 20th ACM International Conference on Intelligent User Interfaces, IUI 2015*, pages 443–446, Atlanta, GA. ACM, ACM
- 573) Pokorny, F. B., Einspieler, C., Zhang, D., Kimmerle, A., Bartl-Pokorny, K. D., Schuller, B. W., and Bölte, S. (2014). The Voice of Autism: An Acoustic Analysis of Early Vocalisations. In *COST ESSEA Conference 2014 Book of Abstracts*, Toulouse, France. 1 page
- 574) Schuller, B. (2013c). Interfaces Seeing and Hearing the User. In *Proceedings The Rank Prize Funds Symposium on Natural User Interfaces, Augmented Reality and Beyond: Challenges at the Intersection of HCI and Computer Vision*, Grasmere, UK. The Rank Prize Funds, The Rank Prize Funds. invited contribution, 1 page
- 575) Ferroni, G., Marchi, E., Eyben, F., Squartini, S., and Schuller, B. (2013b). Onset Detection Exploiting Wavelet Transform with Bidirectional Long Short-Term Memory Neural Networks. In *Proceedings Annual Meeting of the MIREX 2013 community as part of the 14th International Conference on Music Information Retrieval*, Curitiba, Brazil. ISMIR, ISMIR. 3 pages
- 576) Ferroni, G., Marchi, E., Eyben, F., Gabrielli, L., Squartini, S., and Schuller, B. (2013a). Onset Detection Exploiting Adaptive Linear Prediction Filtering in DWT Domain with Bidirectional Long Short-Term Memory Neural Networks. In *Proceedings Annual Meeting of the MIREX 2013 community as part of the 14th International Conference on Music Information Retrieval*, Curitiba, Brazil. ISMIR, ISMIR. 4 pages
- 577) Gunes, H. and Schuller, B. (2012). Dimensional and Continuous Analysis of Emotions for Multimedia Applications: a Tutorial Overview. In *Proceedings of the 20th ACM International Conference on Multimedia, MM 2012*, Nara, Japan. ACM, ACM. 2 pages
- 578) Schröder, M., Pammi, S., Gunes, H., Pantic, M., Valstar, M., Cowie, R., McKeown, G., Heylen, D., ter Maat, M., Eyben, F., Schuller, B., Wöllmer, M., Bevacqua, E., Pelachaud, C.,

- and de Sevin, E. (2011). Come and Have an Emotional Workout with Sensitive Artificial Listeners! In *Proceedings 9th IEEE International Conference on Automatic Face & Gesture Recognition and Workshops, FG 2011*, pages 646–646, Santa Barbara, CA. IEEE, IEEE
- 579) Eyben, F. and Schuller, B. (2010a). Music Classification with the Munich openSMILE Toolkit. In *Proceedings Annual Meeting of the MIREX 2010 community as part of the 11th International Conference on Music Information Retrieval*, Utrecht, Netherlands. ISMIR, ISMIR. 2 pages (acceptance rate: 61 %)
- 580) Eyben, F. and Schuller, B. (2010b). Tempo Estimation from Tatum and Meter Vectors. In *Proceedings Annual Meeting of the MIREX 2010 community as part of the 11th International Conference on Music Information Retrieval*, Utrecht, Netherlands. ISMIR, ISMIR. 1 page (acceptance rate: 61 %)
- 581) Böck, S., Eyben, F., and Schuller, B. (2010a). Beat Detection with Bidirectional Long Short-Term Memory Neural Networks. In *Proceedings Annual Meeting of the MIREX 2010 community as part of the 11th International Conference on Music Information Retrieval*, Utrecht, Netherlands. ISMIR, ISMIR. 2 pages (acceptance rate: 61 %)
- 582) Böck, S., Eyben, F., and Schuller, B. (2010b). Onset Detection with Bidirectional Long Short-Term Memory Neural Networks. In *Proceedings Annual Meeting of the MIREX 2010 community as part of the 11th International Conference on Music Information Retrieval*, Utrecht, Netherlands. ISMIR, ISMIR. 2 pages (acceptance rate: 61 %)
- 583) Böck, S., Eyben, F., and Schuller, B. (2010c). Tempo Detection with Bidirectional Long Short-Term Memory Neural Networks. In *Proceedings Annual Meeting of the MIREX 2010 community as part of the 11th International Conference on Music Information Retrieval*, Utrecht, Netherlands. ISMIR, ISMIR. 3 pages (acceptance rate: 61 %)
- Invited Papers (30):*
- 584) Schuller, B., Wöllmer, M., Eyben, F., Rigoll, G., and Arsić, D. (2011h). Semantic Speech Tagging: Towards Combined Analysis of Speaker Traits. In Brandenburg, K. and Sandler, M., editors, *Proceedings AES 42nd International Conference*, pages 89–97, Ilmenau, Germany. AES, Audio Engineering Society. invited contribution
- 585) Schuller, B., Can, S., and Feussner, H. (2009b). Robust Keyword Spotting in Field Noise for Open-Microphone Surgeon-Robot Interaction. In *Proceedings 5th Russian-Bavarian Conference on Bio-Medical Engineering, RBC-BME 2009*, pages 121–123, Munich, Germany. invited contribution
- 586) Schuller, B., Lehmann, A., Weninger, F., Eyben, F., and Rigoll, G. (2009e). Blind Enhancement of the Rhythmic and Harmonic Sections by NMF: Does it help? In *Proceedings International Conference on Acoustics including the 35th German Annual Conference on Acoustics, NAG/DAGA 2009*, pages 361–364, Rotterdam, The Netherlands. Acoustical Society of the Netherlands, DEGA, DEGA. invited contribution
- 587) Schuller, B. (2008). Speaker, Noise, and Acoustic Space Adaptation for Emotion Recognition in the Automotive Environment. In *Proceedings 8th ITG Conference on Speech Communication*, volume 211 of *ITG-Fachbericht*, Aachen, Germany. ITG, VDE-Verlag. invited contribution, 4 pages
- 588) Schuller, B., Wöllmer, M., Moosmayr, T., and Rigoll, G. (2008k). Robust Spelling and Digit Recognition in the Car: Switching Models and Their Like. In *Proceedings 34. Jahrestagung für Akustik, DAGA 2008*, pages 847–848, Dresden, Germany. DEGA, DEGA. invited contribution, Structured Session Sprachakustik im Kraftfahrzeug
- 589) Schuller, B., Eyben, F., and Rigoll, G. (2008d). Beat-Synchronous Data-driven Automatic Chord Labeling. In *Proceedings 34. Jahrestagung für Akustik, DAGA 2008*, pages 555–556, Dresden, Germany. DEGA, DEGA. invited contribution, Structured Session Music Processing
- 590) Schuller, B., Can, S., Scheuermann, C., Feussner, H., and Rigoll, G. (2008b). Robust Speech Recognition for Human-Robot Interaction in Minimal Invasive Surgery. In *Proceedings 4th Russian-Bavarian Conference on Bio-Medical Engineering, RBC-BME 2008*, pages 197–201, Zelenograd, Russia. invited contribution
- 591) Schuller, B., Müller, R., Hörnler, B., Höthker, A., Konosu, H., and Rigoll, G. (2007c). Audiovisual Recognition of Spontaneous Interest within Conversations. In *Proceedings 9th ACM International Conference on Multimodal Interfaces, ICMI 2007*, pages 30–37, Nagoya, Japan. ACM, ACM. invited contribution, Special Session on Multimodal Analysis of Human Spontaneous Behaviour (acceptance rate: 56 %, IF* 1.77 (2010))
- 592) Schuller, B., Rigoll, G., Grimm, M., Kroschel, K., Moosmayr, T., and Ruske, G. (2007d). Effects of In-Car Noise-Conditions on the Recognition of Emotion within Speech. In *Proceedings 33. Jahrestagung für Akustik, DAGA 2007*, pages 305–306, Stuttgart, Germany. DEGA, DEGA. invited contribution, Structured Session Sprachakustik im Kraftfahrzeug
- 593) Schuller, B., Stadermann, J., and Rigoll, G. (2006f). Affect-Robust Speech Recognition by Dynamic Emotional Adaptation. In *Proceedings 3rd International Conference on Speech Prosody, SP 2006*, Dresden, Germany. ISCA, ISCA. invited contribution, Special Session Prosody in Automatic Speech Recognition, 4 pages
- 594) Schuller, B., Lang, M., and Rigoll, G. (2006d). Recognition of Spontaneous Emotions by Speech within Automotive Environment. In *Proceedings 32. Jahrestagung für Akustik, DAGA 2006*, pages 57–58, Braunschweig, Germany. DEGA, DEGA. invited contribution, Structured Session Sprachakustik im Kraftfahrzeug
- 595) Schuller, B. and Rigoll, G. (2006a). Self-learning Acoustic Feature Generation and Selection for the Discrimination of Musical Signals. In *Proceedings 32. Jahrestagung für Akustik, DAGA 2006*, pages 285–286, Braunschweig, Germany. DEGA, DEGA. invited contribution, Structured Session Music Processing
- 596) Schuller, B., Müller, R., Lang, M., and Rigoll, G. (2005c). Speaker Independent Emotion Recognition by Early Fusion of Acoustic and Linguistic Features within Ensembles. In *Proceedings Interspeech 2005, Eurospeech*, pages 805–809, Lisbon, Portugal. ISCA, ISCA. invited contribution, Special Session Emotional Speech Analysis and Synthesis: Towards a Multimodal Approach (acceptance rate: 61 %, IF* 1.05 (2010), > 100 citations)
- 597) Schuller, B., Lang, M., and Rigoll, G. (2005b). Robust Acoustic Speech Emotion Recognition by Ensembles of Classifiers. In *Proceedings 31. Jahrestagung für Akustik, DAGA 2005*, pages 329–330, Munich, Germany. DEGA, DEGA. invited contribution, Structured Session Automatische Spracherkennung in gestörter Umgebung
- 598) Schuller, B., Rigoll, G., and Lang, M. (2005e). Matching Monophonic Audio Clips to Polyphonic Recordings. In *Proceedings 31. Jahrestagung für Akustik, DAGA 2005*, pages 299–300, Munich, Germany. DEGA, DEGA. invited contribution, Structured Session Music Processing
- 599) Zhang, Z., Weninger, F., and Schuller, B. (2012). Towards Automatic Intoxication Detection from Speech in Real-Life Acoustic Environments. In Fingscheidt, T. and Kellermann, W., editors, *Proceedings of Speech Communication; 10. ITG Symposium*, pages 1–4, Braunschweig, Germany. ITG, IEEE. invited contribution
- 600) Joder, C. and Schuller, B. (2012a). Exploring Nonnegative Matrix Factorisation for Audio Classification: Application to Speaker Recognition. In Fingscheidt, T. and Kellermann, W., editors, *Proceedings of Speech Communication; 10. ITG Symposium*, pages 1–4, Braunschweig, Germany. ITG, IEEE. invited contribution
- 601) Deng, J., Han, W., and Schuller, B. (2012). Confidence Measures for Speech Emotion Recognition: a Start. In Fin-

- gscheidt, T. and Kellermann, W., editors, *Proceedings of Speech Communication; 10. ITG Symposium*, pages 1–4, Braunschweig, Germany. ITG, IEEE. invited contribution
- 602) Wöllmer, M., Kaiser, M., Eyben, F., Weninger, F., Schuller, B., and Rigoll, G. (2012b). Fully Automatic Audiovisual Emotion Recognition – Voice, Words, and the Face. In Fingscheidt, T. and Kellermann, W., editors, *Proceedings of Speech Communication; 10. ITG Symposium*, pages 1–4, Braunschweig, Germany. ITG, IEEE. invited contribution
- 603) Weninger, F., Wöllmer, M., and Schuller, B. (2012h). Sparse, Hierarchical and Semi-Supervised Base Learning for Monaural Enhancement of Conversational Speech. In Fingscheidt, T. and Kellermann, W., editors, *Proceedings 10th ITG Conference on Speech Communication*, pages 1–4, Braunschweig, Germany. ITG, IEEE. invited contribution
- 604) Han, W., Zhang, Z., Deng, J., Wöllmer, M., Weninger, F., and Schuller, B. (2012c). Towards Distributed Recognition of Emotion in Speech. In *Proceedings 5th International Symposium on Communications, Control, and Signal Processing, ISCCSP 2012*, pages 1–4, Rome, Italy. IEEE, IEEE. invited contribution, Special Session Interactive Behaviour Analysis
- 605) Weninger, F. and Schuller, B. (2011b). Fusing Utterance-Level Classifiers for Robust Intoxication Recognition from Speech. In *Proceedings MMCogEmS 2011 Workshop (Inferring Cognitive and Emotional States from Multimodal Measures), held in conjunction with the 13th International Conference on Multimodal Interaction, ICMI 2011*, Alicante, Spain. ACM, ACM. invited contribution, 2 pages (acceptance rate: 39%, IF* 1.77 (2010))
- 606) Weninger, F., Schuller, B., Liem, C., Kurth, F., and Hanjalić, A. (2012e). Music Information Retrieval: An Inspirational Guide to Transfer from Related Disciplines. In Müller, M. and Goto, M., editors, *Multimodal Music Processing*, volume Seminar 11041 of *Dagstuhl Follow-Ups*, pages 195–215, Schloss Dagstuhl, Germany. invited contribution
- 607) Wöllmer, M., Klebert, N., and Schuller, B. (2010e). Switching Linear Dynamic Models for Recognition of Emotionally Colored and Noisy Speech. In *Proceedings 9th ITG Conference on Speech Communication*, volume 225 of *ITG-Fachbericht*, Bochum, Germany. ITG, VDE-Verlag. invited contribution, Special Session Bayesian Methods for Speech Enhancement and Recognition, 4 pages
- 608) Devillers, L. and Schuller, B. (2010). The Essential Role of Language Resources for the Future of Affective Computing Systems: A Recognition Perspective. In *Proceedings 2nd European Language Resources and Technologies Forum: Language Resources of the future – the future of Language Resources*, Barcelona, Spain. FLAReNet. invited contribution, 2 pages
- 609) Steidl, S., Schuller, B., Seppi, D., and Batliner, A. (2009). The Hinterland of Emotions: Facing the Open-Microphone Challenge. In *Proceedings 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*, volume I, pages 690–697, Amsterdam, The Netherlands. HUMAINE Association, IEEE. invited contribution, Special Session Recognition of Non-Prototypical Emotion from Speech – The final frontier?
- 610) Baggia, P., Burkhardt, F., Pelachaud, C., Peter, C., Schuller, B., Wilson, I., and Zovato, E. (2008). Elements of an EmotionML 1.0. Technical report, W3C
- 611) Batliner, A., Seppi, D., Schuller, B., Steidl, S., Vogt, T., Wagner, J., Devillers, L., Vidrascu, L., Amir, N., and Aharonson, V. (2008b). Patterns, Prototypes, Performance. In Hornegger, J., Höller, K., Ritt, P., Borsdorf, A., and Niedermeier, H.-P., editors, *Pattern Recognition in Medical and Health Engineering, Proceedings HSS-Cooperation Seminar Ingenieurwissenschaftliche Beiträge für ein leistungsfähigeres Gesundheitssystem*, pages 85–86, Wildbad Kreuth, Germany. invited contribution
- 612) Grimm, M., Kroschel, K., Schuller, B., Rigoll, G., and Moosmayr, T. (2007b). Acoustic Emotion Recognition in Car Environment Using a 3D Emotion Space Approach. In *Proceedings 33. Jahrestagung für Akustik, DAGA 2007*, pages 313–314, Stuttgart, Germany. DEGA, DEGA. invited contribution, Structured Session Session Sprachakustik im Kraftfahrzeug
- 613) Rigoll, G., Müller, R., and Schuller, B. (2005). Speech Emotion Recognition Exploiting Acoustic and Linguistic Information Sources. In Kokkinakis, G., editor, *Proceedings 10th International Conference Speech and Computer, SPECOM 2005*, volume 1, pages 61–67, Patras, Greece. invited contribution
- Forewords (4):*
- 614) Schuller, B. (2015a). A decade of encouraging speech processing “outside of the box” - a Foreword. In Esposito, A., Faundez-Zanuy, M., Esposito, A. M., Cordasco, G., Drugman, T., Sol-Casals, J., and Morabito, C. F., editors, *Recent Advances in Nonlinear Speech Processing – 7th International Conference, NOLISP 2015, Vietri sul Mare, Italy, May 18–19, 2015, Proceedings*, volume 48, Smart Innovation, Systems and Technologies of *Smart Innovation Systems and Technologies*, pages 3–4. Springer. invited
- 615) Cowie, R., Ji, Q., Tao, J., Gratch, J., and Schuller, B. (2015). Foreword ACII 2015 – Affective Computing and Intelligent Interaction at Xi’an. In *Proc. 6th biannual Conference on Affective Computing and Intelligent Interaction (ACII 2015)*, Xi’an, P.R. China. AAC, IEEE. 2 pages
- 616) Schuller, B. (2015e). Foreword. In Cambria, E. and Husain, A., editors, *Sentic Computing – A Common-Sense-Based Framework for Concept-Level Sentiment Analysis*, pages v–vi. Springer, 2 edition. invited foreword
- 617) Schuller, B. (2015i). Supervisor’s Foreword. In Eyben, F., editor, *Real-time Speech and Music Classification by Large Audio Feature Space Extraction*. Springer. invited foreword, 2 pages
- Reviewed Technical Newsletter Contributions (5):*
- 618) Eyben, F. and Schuller, B. (2014). openSMILE:) The Munich Open-Source Large-Scale Multimedia Feature Extractor. *ACM SIGMM Records*, 6(4)
- 619) Schuller, B., Steidl, S., and Batliner, A. (2013g). The INTERSPEECH 2013 Computational Paralinguistics Challenge – A Brief Review. *Speech and Language Processing Technical Committee (SLTC) Newsletter*
- 620) Schuller, B., Steidl, S., Batliner, A., Schiel, F., and Krajewski, J. (2012e). The INTERSPEECH 2011 Speaker State Challenge – A review. *Speech and Language Processing Technical Committee (SLTC) Newsletter*
- 621) Schuller, B. and Devillers, L. (2010a). Emotion 2010 - On Recent Corpora for Research on Emotion and Affect. *ELRA Newsletter, LREC 2010 Special Issue*, 15(1–2):18–18
- 622) Schuller, B., Steidl, S., Batliner, A., and Jurcicek, F. (2009i). The INTERSPEECH 2009 Emotion Challenge – Results and Lessons Learnt. *Speech and Language Processing Technical Committee (SLTC) Newsletter*
- Technical Reports (7):*
- 623) Keren, G. and Schuller, B. (2016b). Convolutional RNN: an Enhanced Model for Extracting Features from Sequential Data. *arxiv.org*, (1602.05875). 8 pages
- 624) Keren, G., Sabato, S., and Schuller, B. (2016b). Tunable Sensitivity to Large Errors in Neural Network Training. *arxiv.org*, (arXiv:1611.07743). 10 pages
- 625) Schmitt, M. and Schuller, B. (2016). openXBOW – Introducing the Passau Open-Source Crossmodal Bag-of-Words Toolkit. *arxiv.org*, (1605.06778). 9 pages
- 626) Abdić, I., Fridman, L., Marchi, E., Brown, D. E., Angell, W., Reimer, B., and Schuller, B. (2015). Detecting Road Surface Wetness from Audio: A Deep Learning Approach. *arxiv.org*, (1511.07035). 5 pages

- 627) Mousa, A. E.-D., Marchi, E., and Schuller, B. (2015). The ICSTM+TUM+UP Approach to the 3rd CHIME Challenge: Single-Channel LSTM Speech Enhancement with Multi-Channel Correlation Shaping Dereverberation and LSTM Language Models. *arxiv.org*, (1510.00268). 9 pages
- 628) Trigeorgis, G., Bousmalis, K., Zafeiriou, S., and Schuller, B. (2015a). A deep matrix factorization method for learning attribute representations. *arxiv.org*, (1509.03248). 15 pages
- 629) Weninger, F., Schuller, B., Eyben, F., Wöllmer, M., and Rigoll, G. (2014c). A Broadcast News Corpus for Evaluation and Tuning of German LVCSR Systems. *arxiv.org*, (1412.4616). 4 pages