# [UK Speech 2019](https://www.dropbox.com/preview/rtmp/src/org/thesis/papers/Ukspeech2019/abstractBook_UKSpeech2019_fin_noLogos.pdf?role=personal) Conference Report

## Introduction

The UK Speech is a conference for Researchers in the United Kingdom studying the processing of speech or computer speech. This research is a subset of Artificial Intelligence and Natural Language Processing. In modern times it has also been largely associated with the field of Machine Learning.

The 2019 UK Speech took place at the University of Birmingham between the 24th and 25th of June, 2019 at the University of Birmingham. The conference offered 2 keynote talks. The first by Paula Butter and Helen Yannakoudakis of the University of Cambridge Automated Language Teaching and Assessment (ALTA) Institute; was on: Exploring core technologies for automated language teaching and assessment. The second keynote presentation was on: Automated processing of pathological speech by Hiedi Christensen of the Computer Science department of the University of Sheffield.

Other highlights of the events included 32 poster presentations and 6 talks from Amazon TTS Research; University of Technology, Poland; Edinburgh University; KTH Royal Institute of Technology, Sweden; University of Maryland, USA; Toshiba Research, UK and Apple UK Inc. See corresponding sections below and attached abstract book for details.

## Poster Submission Presentation and Abstract

The conference poster submission from the University of Huddersfield was entitled: Deep Scattering End-to-end speech recognition. The abstract for the poster, presented by John Alamina, Dr. David Wilson and Prof. Andrew Crampton, is given in the following paragraphs.

End-to-end discriminative neural network speech models have now become a well-established method in Automatic Speech Recognition.

Our Bi-directional Recurrent neural network (Bi-RNN) end-to-end system, is augmented by features derived from a deep scattering network as opposed to the standard Mel Frequency Cepstral Coefficients(MFCC) features used in state of the art acoustic models. These specialised deep scattering features, consumed by the Bi-RNN, model a light-weight convolution network. This work shows that it is possible to build a speech model from a combination of deep scattering features and a Bi-RNN. There has been no record of deep scattering features being used in end-to-end bi-RNN speech models as far as we are aware.

## Oral Talks

1. Interpretable Deep Learning Models for Detection and Reconstruction of Dysarthric Speech (Danield Korzekwa, Roberto Barra-Chicote, Bozena Kostek, Thomas Drugman, Mateusz Lasjszczak)
2. Modern speech synthesis and its implication for speech sciences (Zofia Malisz, Gustav Eje Henter,m Cassia Valentini-Botinhao, Oliver Watts, Jonas Beskow, Joakim Gustafson
3. Continuous representations can support early phonetic learning (Yevgen Matusevych, Thomas Schatz, Sharon Goldwater, Naomi Feldman)
4. Using generative modelling to produce varied intonation for speech synthesis (Zack Hodari, Oliver Watts, Simon King)
5. Conversational systems: Why dialogue manager should consider context (Margarita Kotti)
6. Neural network-based modeling of phonetic durations Zizi Wei, Melvyn Hunt and Adrian Skilling

## Poster sessions

1. Data Science system for the quality assessment and monitoring of Neural Text-To-Speech on a large scale (A. Gabrys, D. Korzekwa, J. Rohnke, A. Ezzerg, R. Srikanth, G. Czachor and K. Viacheslav)
2. Computational cognitive assessment: investigating the use of an Intelligent Virtual Agent for the detection of early signs of dementia Bahman Mirheidari, Daniel Blackburn, Ronan O'Malley, Traci Walker, Annalena Venneri, Markus Reuber and Heidi Christensen
3. Disentangling Style Factors from Speaker Representations (Jennifer Williams and Simon King)
4. Automatic Grammatical Error Detection of Non-Native Spoken Learner English (Kate Knill(1), Mark J.F. Gales(1), Potsawee Manakul(1) and Andrew Caines(2))
5. Exploring how phone classification neural networks learn phonetic information by visualising and interpreting bottleneck features (Linxue Bai(1), Philip Weber(2), Peter Jancovic(1) and Martin Russell(1))
6. End-to-end speaker recognition using CNN-LSTM-TDNN (Xiaoxiao Miao(1,2) and Ian McLoughlin(1))
7. Singing Voice Conversion with Generative Adversarial Networks (Berrak Sisman(1,2) and Haizhou Li(1))
8. Towards the Understanding of Communicating Emotions for People with Dysarthria (Lubna Alhinti, Heidi Christensen and Stuart Cunningham)
9. Lip-Reading with Limited-Data Network (Adriana Fernandez-Lopez and Federico M. Sukno)
10. Developing Coherent Fallback Strategies for Open-domain Conversational Agents (Ioannis Papaioannou and Oliver Lemon)
11. Spontaneous conversational TTS from found data (Eva Szekely, Gustav Eje Henter, Jonas Beskow and Joakim Gustafson)
12. Hierarchical RNNS for Waveform Level Speech Synthesis (Qingyun Dou, Moquan Wan, Gilles Degottex, Zhiyi Ma and Mark J.F. Gales)
13. Exploring the Trade-offs between Acoustic and Language Modelling Constraints for Dysarthric Speech Recognition (Zhengjun Yue, Feifei Xiong, Heidi Christensen and Jon Barker)
14. On the Usefulness of Statistical Normalisation of Bottleneck Features for Speech Recognition (Erfan Loweimi, Peter Bell and Steve Renals)
15. Identification of geographical origin from accented speech (Wen Wu(1) and Martin Russell(2))
16. Multitasking with Alexa: How Using Intelligent Personal Assistants Impacts Language-based Primary Task Performance (Justin Edwards(1), He Liu(1), Tianyu Zhou(1), Sandy Gould(2), Leigh Clark(1), Phillip Doyle(1) and Benjamin Cowan(1))
17. On Learning Interpretable CNNs with Parametric Modulated Kernel-based Filters (Erfan Loweimi, Peter Bell and Steve Renals)
18. Non-native Speaker Verication for Spoken Language Assessment: Malpractice Detection in Speaking Tests (Linlin Wang, Yu Wang and Mark J. F. Gales)
19. Mapping Perceptions of Humanness in Intelligent Personal Assistant Interactions (Philip R Doyle, Justin Edwards, Odile Dumbleton, Leigh Clark and Benjamin R Cowan)
20. Natural Language Processing Applied to Empathy Agent for People with Mental Health Problem (Feifei Xiong(1), Fuschia Sirois(2), Katherine Easton(3,7), Abigail Millings(2), Matthew Bennion(3,7), Paul Radin(4), Ian Tucker(5), Rafaela Ganga(6) and Heidi Christensen(1,7))
21. Speech Synthesis and Dramatic Performance: You have to Suffer Darling (Matthew P. Aylett(1), Benjamin R. Cowan(2) and Leigh Clark(2))
22. Deep Scattering End-to-End Architectures for Speech Recognition (Iyalla John Alamina, David Wilson and Andrew Crampton)
23. Improving the intelligibility of speech playback in everyday scenarios (Carol Chermaz(1), Cassia Valentini-Botinhao(1), Henning Schepker(2) and Simon King(1))
24. Disfluency Detection for Spoken Learner English (Yiting Lu, Mark Gales, Kate Knill, Potsawee Manakul and Yu Wang)
25. Lattice inspired semisupervised training of end to end speech recognition (Andrea Carmantini, Peter Bell and Steve Renals)
26. EFFECT OF DATA REDUCTION ON SEQUENCE-TO-SEQUENCE NEURAL TTS Javier Latorre, Jakub Lachowicz, Jaime Lorenzo-Trueba, Thomas Merritt, Thomas Drugman, Srikanth Ronanki, Klimkov Viacheslav
27. Diligently Delete Entry: Determining Errors in Non-Native Spontaneous Speech John Sloan, Emma O'Neill and Julie Carson-Berndsen
28. The University of Birmingham 2019 Spoken CALL Shared Task Systems: Exploring the importance of word order in text processing (Mengjie Qian(1), Peter Jancovic(1) and Martin Russell(2))
29. Using Video Information to Improve Automatic Speech Recognition in the Distant Microphone Scenario Jack Deadman and Jon Barker
30. Exploring Generalizability of Automatic Phoneme Recognition Models (Emir Demirel(1), Sven Ahlback(2) and Simon Dixon(1))
31. An investigation of auditory models to objectively analyze speech synthesis (Sebastien Le Maguer, Marie-Caroline Villedieu and Naomi Harte)
32. The effects of expressional feature transplant on singing synthesis (Christopher G. Buchanan, Matthew P. Aylett, and David A. Braude)