



## Interspeech 2018 - Plenary Speaker-2

### Hervé Bourlard

*Idiap Research Institute and EPFL, Switzerland*

### Webpage

<http://people.idiap.ch/bourlard>



**Title:** Evolution of Neural Network Architectures for Speech Recognition

### Abstract

Over these last few years, the use of Artificial Neural Networks (ANNs), now often referred to as deep learning or Deep Neural Networks (DNNs), has significantly reshaped research and development in a variety of signal and information processing tasks. While further boosting the state-of-the-art in Automatic Speech Recognition (ASR), recent progresses in the field have also allowed for more flexible and faster developments in emerging markets and multilingual societies (e.g., under-resourced languages).

In this talk, we will provide a historical account of ANN architectures used for ASR since the mid-1980's, and now used in most ASR and spoken language understanding applications. We will start by recalling/revisiting key links between ANNs and statistical inference, discriminant analysis, and linear/nonlinear algebra. Finally, we will briefly discuss more recent trends towards novel DNN-based ASR approaches, including complex hierarchical systems, sparse recovery modeling, and “end-to-end systems.”

However, and in spite of the recent progress in the area, we still lack basic understanding of the problems in hands. Although more and more tools are now available, in association with basically “unlimited” processing and data resources, we still fail in building principled ASR models and theories. Alternatively, we are still relying on “ignorance-based” models, often exposing limitations of our understanding, rather than enriching the field of ASR. Discussion of these limitations will underpin all of our overview.

### Biography

Hervé Bourlard is Director of the Idiap Research Institute, Full Professor at the Swiss Federal Institute of Technology Lausanne (EPFL), and Founding Director of the Swiss NSF National Centre of Competence in Research on “Interactive Multimodal Information Management (IM2)” (2001-2013). He is also an External Fellow of the International Computer Science Institute (ICSI), Berkeley, CA.

His research interests mainly include statistical pattern classification, signal processing, multi-channel processing, artificial neural networks, and applied mathematics, with applications to a wide range of Information and Communication Technologies, including spoken language processing, speech and speaker recognition, language modeling, multimodal interaction, and augmented multi-party interaction.

H. Bourlard is the author/co-author/editor of 8 books, and over 330 reviewed papers (including one IEEE paper award). He is a Fellow of IEEE and ISCA, and a Senior Member and Member of the European Council of ACM. He is the recipient of several scientific and entrepreneurship awards.