

# Mohamed Bouallegue

☎ +33 615 473 786 • ✉ mohamed.bouallegue@univ-lemans.fr

## Ph.D. in Computer Science

### Personel Data

---

- Date of birth : August, 19 1984.
- Nationality : French
- Marital status : single.

### Research Interests

---

- Automatic speech recognition
- Natural Language Processing
- Keywords/Topics Extraction

### Education

---

#### Ph.D. in computer science

*University of Avignon, LIA, France*  
Speech processing

**Avignon**  
*sep 2009 - 2013*

#### Master of Science

*University of Grenoble, France*  
Specialized in Natural Language Processing

**Avignon**  
*2008*

### Employment Experience (Research)

---

★ **Since Sept. 2014:** Postdoctoral researcher

LIUM - Computer Science Department , University of Le Mans (France)

★ **2012-2014:** Teaching and research assistant (ATER)

CERI, University of Avignon (France)

★ **2009-2012:** Research Engineer

Easy Top Transcription (France)

★ **2009-2013:** Ph.D. in Computer Science

CERI, University of Avignon (France)

★ **2008-2009:** Master of Science

LIDILEM, University of Grenoble (France)

## Employment Experience (Teaching)

---

★ *Various courses in Computer Science:*

- C++ Programming
- Data base engineering
- Internet engineering : HTML, PHP, CSS

★ **2012-2014:** Teaching and research assistant (ATER)

CERI, University of Avignon (France)

Total number of hours: 285.5 hours.

★ **Oct. 2010-June 2012:** Ph.D. in Computer Science

CERI, University of Avignon (France)

Total number of hours: 111 hours.

## Computer skills

---

**Programming:** C++, Java, C#, Perl, Prolog

**Modeling:** Oriented Object Programming, UML

**Research:** Maple, MatLab, R

## Professional societies

---

**EEE:** The Institute of Electrical and Electronics Engineers, Student Member.

**ISCA:** International Speech Communication Association, Student Member.

**AFCP:** Association Francophone de la Communication Parlée, Student Member.

## Languages

---

**French:** Written, spoken and read.

**English:** Written, spoken and read.

**Arabic:** Native language.

## Master in computer science

---

**Title:** *The expanding request technique*

**Supervisor:** Olivier Kraif

**Abstract:** The expanding request technique is harnessed in the search for documents so as find out more relevant document. Similar techniques are used to uncover linguistically phrases by analyzing the syntactical and semantic levels in order to increase the number of relevant expressions in the course of a research. This work is based on the utilization of this research technique to find out illustrations of polylexical structures that are expressed in a "canonical" form. After analyzing various sorts of canonical expressions that are likely to be used for teaching purposes whether handbooks or dictionaries, we have chosen to focus on the paradigmatic instance of verb phrases and verb-based structures.

**Key Words:** Expression seeking, Specific grammar, expanding request, consistent expression

## Ph.D. thesis

---

**Title:** *Factor analysis for acoustic modeling for systems speech recognition*

**Supervisor:** Driss Matrouf, Georges Linarès,

**Abstract:** In this thesis, we propose to use techniques based on factor analysis to build acoustic models for automatic speech processing, especially Automatic Speech Recognition (ASR). Firstly, we were interested in reducing the footprint memory of acoustic models. Our factor analysis-based method demonstrated that it is possible to pool the parameters of acoustic models and still maintain performance similar to the one obtained with the baseline models. The proposed modeling leads us to deconstruct the ensemble of the acoustic model parameters into independent parameter sub-sets, which allow a great flexibility for particular adaptations (speakers, genre, new tasks etc.). With current modeling techniques, the state of a Hidden Markov Model (HMM) is represented by a combination of Gaussians (GMM: Gaussian Mixture Model). We propose as an alternative a vector representation of states: the factors of states. These factors of states enable us to accurately measure the similarity between the states of the HMM by means of an euclidean distance for example. Using this vector representation, we propose a simple and effective method for building acoustic models with shared states. This procedure is even more effective when applied to under-resourced languages. Finally, we concentrated our efforts on the robustness of the speech recognition systems to acoustic variabilities, particularly those generated by the environment. In our various experiments, we examined speaker variability, channel variability and additive noise. Through our factor analysis-based approach, we demonstrated the possibility of modeling these different types of acoustic variability as an additive component in the cepstral domain. By compensation of this component from the cepstral vectors, we are able to cancel out the harmful effect it has on speech recognition.

**Key Words:** Automatic speech recognition, factor analysis, compact acoustic modeling, phonetic classification, acoustic variability.

## Selected Publications

---

Yannick Estève, Mohamed Bouallegue, Carole Lailler, Mohamed Morchid, Richard Dufour, Georges Linarès, Driss Matrouf and Renato De Mori, "Integration of Word and Semantic Features for Theme Identification in Telephone Conversations", *IWSDS 2015*.

Mohamed Morchid, Mohamed Bouallegue, Richard Dufour, Georges Linarès, Driss Matrouf and Renato De Mori, "An I-vector Based Approach to Compact Multi-Granularity Topic Spaces Representation of Textual Documents", *EMNLP 2014* - long paper.

Mohamed Bouallegue, Mohamed Morchid, Richard Dufour, Driss Matrouf, Georges Linarès and Renato De Mori "Subspace Gaussian Mixture Models for Dialogues Classification", *ISCA INTERSPEECH 2014*.

Mohamed Bouallegue, Mohamed Morchid, Richard Dufour, Driss Matrouf, Georges Linarès and Renato De Mori, "Factor Analysis based Semantic Variability Compensation for Automatic Conversation Representation", *ISCA INTERSPEECH 2014*.

Mohamed Bouallegue, Emmanuel Ferreire, Driss Matrouf, Georges Linarès, Maria Goudi and Pascal Nocera, "Acoustic modeling for under-resourced languages based on vectorial HMM-states representation using Subspace Gaussian Mixture Models", *IEEE SLT 2012*.

Mohamed Bouallegue, Driss Matrouf, Mickael Rouvier and Georges Linarès, "Subspace Gaussian Mixture Models for vectorial HMM-states representation", *IEEE ASRU 2011*.

**Full list of publications is available [here](#)**

## References

---

**Driss Matrouf**

Associate Professor (HDR)  
University of Avignon  
339 chemin des Meinajariès  
84911 Avignon, France  
Phone: +33 490 843 526  
*driss.matrouf@univ-avignon.fr*

**Professor Georges Linares**

Head of the LIA  
University of Avignon  
339 chemin des Meinajariès  
84911 Avignon, France  
Phone: +33 490 843 520  
*georges.linares@univ-avignon.fr*

**Emeritus Professor Renato De Mori**

School of Computer Science  
McGill University McConnell Eng. Bldg.  
204N 3480 University Street  
Montreal, Canada  
Phone: +1 514 398 7071  
*rdemori@cs.mcgill.ca*

## Interests

---

**Reading:** Historical and political books

**Sports:** Foot-ball, chess sets