



## Machine Listening in Multisource Environments (CHiME 2011)

CHiME 2011 was an ISCA-approved satellite workshop of Interspeech 2011 that considered the challenge of developing machine listening applications for operation in multisource environments, i.e. real-world conditions with acoustic clutter, where the number and nature of the sound sources is unknown and changing over time. CHiME brings together researchers from a broad range of disciplines (computational hearing, blind source separation, speech recognition, machine learning) to discuss novel and established approaches to this problem. The cross-fertilisation of ideas will foster fresh approaches that efficiently combine the complementary strengths of each research field.

Relevant research topics include (but are not limited to),

- automatic speech recognition in multisource environments,
- acoustic event detection in multisource environments,
- sound source detection and tracking in multisource environments,
- music information retrieval in multisource environments,
- sound source separation or enhancement in multisource environments,
- robust feature extraction and classification in multisource environments,
- scene analysis and understanding for multisource environments.

### The PASCAL CHiME Challenge

As a focus for discussion during the day, the workshop acted as host to the PASCAL CHiME Speech Separation and Recognition Challenge. This is a binaural, multisource speech separation and recognition challenge supported by the EU PASCAL network and the UK EPSRC. To find out more please visit the [Challenge Website](#).

### CHiME 2011, September 1

*Florence, Italy*

#### Organising Committee

- Jon Barker (UK)
- Emmanuel Vincent (FR)
- Walter Kellermann (DE)
- Dan Ellis (US)
- John Hershey (US)
- Hiroshi Okuno (JP)
- Phil Green (UK)

#### Sponsors

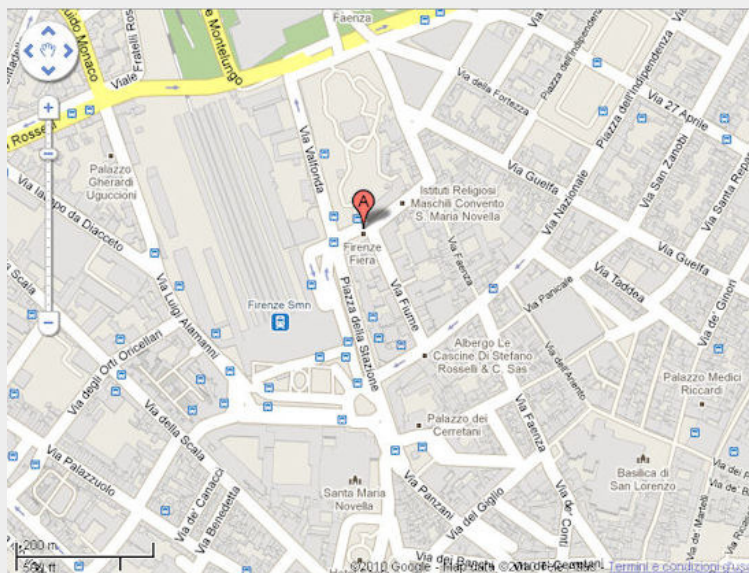
- [The PASCAL network](#)
- [EPSRC](#)



## Venue

The workshop took place at the Firenze Fiera congress centre in Florence. Sessions were held on the ground floor of the [Palazzo dei Congressi](#) which is an 18th century villa at the centre of the congress site.

Click on the map below to open a link to Google Maps in a separate window.



## Workshop Chairs

- Dr Jon Barker,  
University of Sheffield, UK
- Dr Emmanuel Vincent,  
INRIA Rennes, France

## Organising Committee

- Prof. Walter Kellermann,  
University of Erlangen-Nuremberg, Germany
- Prof. Dan Ellis,  
Columbia University, USA
- Dr. John Hershey,  
Mitsubishi Electric Research Laboratories, USA
- Prof. Hiroshi Okuno,  
Kyoto University, Japan
- Prof. Phil Green,  
University of Sheffield, UK

## Scientific Committee

- Dr Jon Barker,  
University of Sheffield, UK
- Dr Heidi Christensen,  
University of Sheffield, UK
- Prof. Reinhold Häb-Umbach,  
University of Paderborn, Germany
- Prof. Walter Kellermann,  
University of Erlangen-Nuremberg, Germany
- Dr Ning Ma,  
University of Sheffield, UK,
- Dr Atsushi Nakamura,  
NTT Communication Science Labs, Kyoto, Japan
- Dr Francesco Nesta,  
FBK-IRST, Italy
- Prof. Hiroshi Okuno,  
Kyoto University, Japan
- Dr Alexey Ozerov,  
INRIA Rennes, France
- Dr.-Ing. Armin Sehr,  
University of Erlangen-Nuremberg, Germany
- Dr Emmanuel Vincent,  
INRIA Rennes, France







## Workshop Programme

The [detailed programme](#) follows below with links to papers and presentations (where available).

- 9:00 Opening/Welcome
- 9:10 Overview of CHiME Challenge including summary of results
- 9:50 [Oral session 1](#): challenge papers
- 10:40 Break
- 11:00 [Oral session 2](#): challenge papers
- 12:15 Lunch break
- 13:45 [Poster session](#)
- 15:45 Break
- 16:00 [Oral session 3](#): multisource event detection and classification
- 16:50 Plenary discussion: results and future evaluations
- 17:50 Closing

## Detailed Programme

### Introduction

- Opening Remarks [**View Slides:** [pdf](#)]  
Jon Barker (*University of Sheffield, UK*), Emmanuel Vincent (*INRIA, Rennes, France*)
- Overview of the PASCAL CHiME Speech Separation and Recognition Challenge [**View Slides:** [pdf](#)]  
Jon Barker (*University of Sheffield, UK*), Emmanuel Vincent (*INRIA, Rennes, France*), Ning Ma (*University of Sheffield, UK*), Heidi Christensen (*University of Sheffield, UK*) and Phil Green (*University of Sheffield, UK*),

### Oral session 1

*Session Chair: Phil Green, University of Sheffield*

- [Exemplar-Based Recognition of Speech in Highly Variable Noise](#) [**View Slides:** [pdf](#)]  
Antti Hurmalainen, Katariina Mahkonen (*Tampere University of Technology, Finland*), Jort F. Gemmeke (*Katholieke Universiteit Leuven, Belgium*), and Tuomas Virtanen (*Tampere University of Technology, Finland*)
- [CHiME Challenge: Approaches to Robustness using Beamforming and Uncertainty-of-Observation Techniques](#) [**View Slides:** [pdf](#), [ppt](#)]  
Dorothea Kolossa (*Ruhr University Bochum, Germany*), Ramon Fernandez Astudillo, Alberto Abad (*INESC-ID, Portugal*), Steffen Zeiler (*Ruhr University Bochum, Germany*), Rahim Saeidi (*University of Eastern Finland*), Pejman Mowlaei (*Ruhr University Bochum, Germany*), Joao Paulo da Silva Neto (*INESC-ID, Portugal*), and Rainer Martin (*Ruhr University Bochum, Germany*)

### Oral session 2

*Session Chair: Reinhold Häb-Umbach, University of Paderborn*

- [Speech Recognition in the Presence of Highly Non-Stationary Noise Based on Spatial, Spectral and Temporal Speech/Noise Modeling Combined with Dynamic Variance Adaptation](#) [**View Slides:** [pdf](#)]  
Marc Delcroix, Keisuke Kinoshita, Tomohiro Nakatani, Shoko Araki, Atsunori Ogawa, Takaaki Hori, Shinji Watanabe, Masakiyo Fujimoto, Takuya Yoshioka, Takanobu Oba, Yotaro Kubo, Mehrez Souden,

- Seong-Jun Hahm, and Atsushi Nakamura (*NTT Communication Science Labs, Japan*)
- [Robust Automatic Speech Recognition through On-line Semi-Blind Source Extraction](#) [**View Slides:** pdf]  
Francesco Nesta and Marco Matassoni (*Fondazione Bruno Kessler-irst, Italy*)
- [The Munich 2011 CHiME Challenge Contribution: NMF-BLSTM Speech Enhancement and Recognition for Reverberated Multisource Environments](#) [**View Slides:** pdf]  
Felix Weninger, Jürgen Geiger, Martin Wöllmer, Björn Schuller, and Gerhard Rigoll (*Technische Universität München, Germany*)

### Oral session 3

Session Chair: Hiroshi Okuno, Kyoto University

- [GMM-Based Classification from Noisy Features](#) [**View Slides:** pdf]  
Alexey Ozerov (*INRIA, Rennes, France*), Mathieu Lagrange (*STMS Lab IRCAM - CNRS - UPMC, France*), and Emmanuel Vincent (*INRIA, Rennes, France*)
- [Sound Event Detection in Multisource Environments Using Source Separation](#) [**View Slides:** pdf]  
Toni Heittola, Annamaria Mesaros, Tuomas Virtanen (*Tampere University of Technology, Finland*), and Antti Eronen (*Nokia Research Center, Finland*)

### Poster session

1. [A Two-Channel Acoustic Front-End for Robust Automatic Speech Recognition in Noisy and Reverberant Environments](#)  
Roland Maas, Andreas Schwarz, Yuanhang Zheng, Klaus Reindl, Stefan Meier, Armin Sehr, and Walter Kellermann (*University of Erlangen-Nuremberg, Germany*)
2. [CHiME Data Separation Based on Target Signal Cancellation and Noise Masking](#)  
Zbynek Koldovsky, Jiri Malek, Jan Nouza (*TU Liberec, Czech Republic*), and Miroslav Balik (*VUT Brno, Czech Republic*)
3. [Designing Multimodal Acoustic Environment Corpus to Improve Speech Interaction in Living Room](#) Kenichi Shibata, Kengo Ikeya, Yuki Deguchi, Yoichi Takebayashi, Shigeyoshi Kitazawa, Shinya Kiriya (*Shizuoka University, Japan*)
4. [Exemplar-Based Speech Enhancement and its Application to Noise-Robust Automatic Speech Recognition](#)  
Jort F. Gemmeke (*Katholieke Universiteit Leuven, Belgium*), Tuomas Virtanen, and Antti Hurmalainen (*Tampere University of Technology, Finland*)
5. [Mask Estimation and Sparse Imputation for Missing Data Speech Recognition in Multisource Reverberant Environments](#)  
Heikki Kallasjoki (*Aalto SCI, Finland*), Sami Keronen (*Aalto University School of Science, Finland*), Guy J. Brown (*University of Sheffield, UK*), Jort F. Gemmeke (*Katholieke Universiteit Leuven, Belgium*), Ulpu Remes, and Kalle J. Palomäki (*Aalto University School of Science, Finland*)
6. [Multi-layer Collaborative Microphone Array Beamforming in Presence of Nonstationary Interfering Signals](#)  
Danilo Comminiello, Michele Scarpiniti, Raffaele Parisi (*Sapienza University of Rome, Italy*), Albenzio Cirillo, Mauro Falcone (*Fondazione Ugo Bordoni, Italy*), and Aurelio Uncini (*Sapienza University of Rome, Italy*)
7. [Recent Advances in Fragment-Based Speech Recognition in Reverberant Multisource Environments](#)  
Ning Ma, Jon Barker, Heidi Christensen, Phil Green (*University of Sheffield, UK*)
8. [Robust Speech Recognition in Multi-Source Noise Environments using Convolutional Non-Negative Matrix Factorization](#)  
Ravichander Vippera, Simon Bozonnet, Dong Wang, and Nicholas Evans (*EURECOM, France*)
9. [Source Separation using the Spectral Flatness Measure](#)  
Rolf Bardeli (*Fraunhofer IAIS, Germany*)
10. [Using the FASST Source Separation Toolbox for Noise Robust Speech Recognition](#)  
Alexey Ozerov and Emmanuel Vincent (*INRIA, Rennes, France*)



---

[CHiME Workshop Home](#)

▪ [CHiME Challenge Home](#)

▪ [EPSRC CHiME Project Home](#)

▪ [PASCAL Network](#)



# The PASCAL 'CHiME' Speech Separation and Recognition Challenge

[Home](#)   [Background](#)   [Instructions](#)   [Data](#)   [Tools](#)   [Results](#)   [Workshop](#)

(N.B the PASCAL challenge is now officially closed and the results are available [here](#). The challenge instruction and data can be found on this site for the benefit of groups wishing to compare their algorithms with those that have been submitted.)

## Welcome to the PASCAL 'CHiME' Challenge

In 2006 the PASCAL network funded the [1st Speech Separation challenge](#) which addressed the problem of separating and recognising speech mixed with speech. We are now launching a successor to this challenge that aims to tackle speech separation and recognition in more typical everyday listening conditions.

**The challenge employs noise background that has been collected from a real family living room using binaural microphones.** Target speech commands have been mixed into the environment at a fixed position using genuine room impulse responses. The task is to separate the speech and recognise the commands being spoken using systems that have been trained on noise-free commands and room noise recordings.

**On this web site** you will find everything you need to get started. The [background](#) section explains the general motivation for the challenge. The [instructions](#) section describes the separation/recognition task in more detail and what you need to do in order to take part. The datasets are available for [download](#) already, and the evaluation tools will become available by the end of October. Further important dates are listed in the [schedule](#).

**This is a multidisciplinary challenge.** We hope to encourage participants from the machine learning, source separation and speech recognition communities. Although the ultimate evaluation will be through speech recognition scores, participants may submit either separated signals, robust feature extractors or complete recognition systems. All entrants will be invited to submit papers describing their work to a dedicated [satellite workshop hosted at Interspeech 2011](#). Participants will also be invited to submitted longer versions of their work for a **Special Issue of Computer Speech and Language** on the theme of Machine Listening in Multisource Environments has been organised.

If you have any questions regarding the challenge please do not hesitate to [contact](#) us.

### News

- **New! RESULTS** of the PASCAL CHiME challenge are now available
- Proceedings of the **CHiME 2011 Workshop** are now online.
- 16 Mar 2011. Complete final test set available for [download](#)

### Organisers

University of Sheffield, UK,

- [Jon Barker](#)
- Heidi Christensen
- Ning Ma
- Phil Green

INRIA Rennes, France,

- [Emmanuel Vincent](#)

### Sponsors

- [The PASCAL network](#)
- [EPSRC](#)