Closing Session

By the Program Committee

© ELRA

General issues

- Infrastructures
- Metadata & Standards
- Linguistic Web Services of any type
- Sustainability
    - We start to converge?
    - Are we going open?

© ELRA

Written (1)

- evident consolidation of know-how and best practices in multilevel corpus processing/annotation and variable corpora building
- extension to new national & regional languages not sufficiently covered so far
- semantic resources, annotation and ontological engineering almost at cruising speed now (ca 60 papers at this LREC)
- considerable progress in multilingual information processing and machine translation (ca 70 papers at this LREC)

© ELRA

Written (2)

- new areas receiving considerable attention
  - spatiotemporal annotation
  - sentiment analysis and emotion
- future challenges
  - robustness of language technology wrt user generated content in e.g. Web2.0
  - intelligent integration/interfacing of language with other media and modalities in content processing and man-machine interfaces
  - language resources and technology in cognition/perception analysis

© ELRA

Written (3)

- Multiword Expressions
  - Proper identification and treatment essential for MT, QA, IR, etc.
  - around 20 papers at LREC
  - Workshop on Sunday
  - Nevertheless, more work in this area is desirable and needed

"Traditional" (already in LREC-04 and 06)

	2008	2006
Dialogue Corpora	30%	24%
Emotional Speech	10%	6%
Noisy Conditions	10%	4%
Broadcast News	8%	14%
Telephone Speech	6%	7%
Text to Speech	6%	20%
Non-native	4%	14%

"Specialization of Databases"

	(% papers)
Age Groups	16%
Verbal + non Verbal	14%
Medical/Legal/Lectures	12%
Sign Languages	6%
Biometric DBs for Speaker Identif/verif.	4%

Tools, Platforms and Methodologies

- Important effort in the development of annotation methodologies, platforms and tools for speech and multimodal resources.
- Some thinking and some actions should be taken to standardize annotation schemes, information structure, methodologies, etc. to easily reuse and improve the already available and future resources and to guarantee the quality of the resources.

© ELRA

Spoken Language Evaluation

lrec 2008 marrakech

- **Basic reflections on evaluation principles and evaluation metrics**
- **Increasing number of languages**
  - Broadcast News
    - Czech, Khmer, Thai, Norwegian, Arabic…
    - 1 hour of speech needs 1000 working hours for phonemic transcription
- **Evaluation of more and more complex systems including SLP**
  - Speech to speech translation (in dialogs)
    - ASR + MT + TTS (+SLD)
    - Still no use of prosodics
    - Automated evaluation vs human judgment
  - Question answering on speech in meetings
    - Conversational SR + Q&A
  - Multy-party dialogs
    - Multy-party, spontaneous SR + SLD
    - Evaluation of the detection of language disfluencies
  - Historic speech collections retrieval
    - ASR + IR + UI
    - From Broadcast News to Cultural Heritage :
    - From 30% to 70 % WER : New metrics RIA (Rank Index Accuracy)
- **Evaluation campaigns or within projects**
  - Transtac, GALE, CLEF, Technolangue, STEVIN, CENSREC…
  - TC-Star, CHIL, Quaero…

© ELRA

Evaluation - Written

- **Evaluation is very important at LREC – and mirrors the field**
- Evaluation of simple components and of applications
- Semantics coming up: evaluation of semantics, senses, time etc.
- MT evaluation is developingInformation extraction, IR, QA, etc. is strong and evolving
- Classification, summarisation etc.

- **The field is maturing, metrics become more standardized**
- **Wanted: more on metrics and methodologies – NB NIST-MATR**

Conference statistics

Submissions 2006			Submissions 2008		
Track	**Number**	**%**	**Track**	**Number**	**%**
E	165	21,02	**E**	180	20,00
G	42	5,35	**G**	163	18,11
MM	38	4,84	**MM**	49	5,44
S	101	12,87	**S**	114	12,67
T	69	8,79	**T**	72	8,00
W	370	47,13	**W**	322	35,78
Total	785	100,00	**Total**	900	100,00

Participation statistics

lrec 2008 marrakech

Total number of Participants (Conference & workshops)	1111
Status	**Number of Participants**
Full	597
ELRA Member	136
Student	305
Others (staff, org. committees, invited)	73

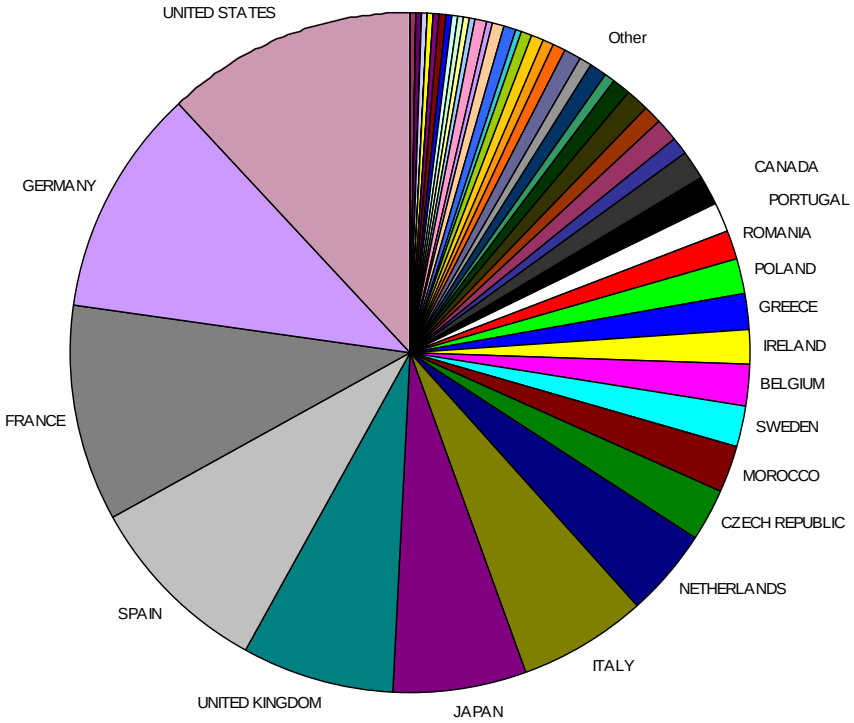
© ELRA

Participation statistics

lrec 2008 marrakech

Total number of participants (Conference & workshops)	1111
Status	**Number of Participants**
Academic Institutions	933
Industry	111
Others	66

Participation statistics



Tutorials and workshops

6 tutorials

25 workshops

- Corpora for Research on Emotion and Affect
- Sentiment Analysis: Emotion, Metaphor, Ontology and Terminology

- 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora – 72 registered participants!!

© ELRA

Acknowledgments

- **Program Committee**
- **ELRA and the ELRA Board**
- **Scientific Committee**
- **International Advisory Board**
- **Local Advisory Board**
- **Local Organising Committee**, in particular **Abdelhak Mouradi** & **Mme S. Kerfi**

Acknowledgement Sponsors

Acknowledgments

- **Authorities, associations, organisations, committees, agencies, companies** that have supported LREC
- **Workshop, tutorial, and panel organisers**
- **Authors**

Acknowledgments

- **ELDA & ILC**

Behind the scene

lrec 2008 marrakech

- **Antonio**
- **Cecile**
- **Chomicha**
- **Djamel**
- **Eva**
- **Giulia**
- **Mathieu**
- **Nicolas**
- **Niklas**
- **Olivier**
- **Riccardo**
- **Tommaso**
- **Valerie**
- **Valeria**
- **Victoria**

3 Key Men

- **Roberto**
- **Sergio**
- **Vincenzo**

© ELRA

- **Hélène and Sara**

Moroccan Summer Time

START SATURDAY NIGHT
02:00 → 03:00
Make sure not to miss your flight !

GALA - BANQUET

Shuttles in front of the conf Center

At 20.15

© ELRA