

發行人：張俊盛  
 主編：曾淑娟  
 執行編輯：黃琪  
 地址：台北市研究院路二段 128 號中研院資訊所  
 劃撥帳號：19166251  
 電話：(02)2788-3799ext.:1502  
 傳真：(02)2788-1638  
 E-mail：rocling@hp.iis.sinica.edu.tw  
 網址：http://rocling.iis.sinica.edu.tw/ROCLING/  
 討論區：aclcp-rocling@listserv.nthu.edu.tw

## 本期要目

- |  |         |
|--|---------|
| 壹. ROCLING XV 報名表、暫訂議程                         | 第二~四頁   |
| 貳. The Map Task 語料庫簡介-解志強                      | 第五~八頁   |
| 參. ACL-03 參加會議心得-吳鑑城                           | 第九~十頁   |
| 肆. 學術活動預告(2 <sup>nd</sup> GWC-2004, LREC-2004) | 第十一 十六頁 |

### ROCLING XV 開始報名

「第十五屆計算語言學研討會」謹訂於 9/18、9/19 二日假新竹清華大學國際會議廳舉行，主講人則邀請到教育部國語會主委鄭良偉教授及中研院院士鄭錦全院士兩位貴賓。本次會議除了合併「資訊檢索研討會」、「語音訊號處理研討會」及「電腦輔助語言教學研討會」舉行外，亦增加 Student Workshop Papers 及 Poster Demo Papers，報名期限自即日起至 8/31 日止，歡迎有興趣之學者踴躍報名，報名訊息請參閱第二頁。

### Tcc-300 國內學術單位申請優惠

學會為了促進語音研究領域之蓬勃發展，自即日起凡屬國內學術研究單位者，申請麥克風語料庫(Tcc-300Edu)只收取工本費 1,000 元，一般民間企業則維持原價，申購說明請參閱本會網站。

(<http://rocling.iis.sinica.edu.tw/ROCLING/>)

### 第八屆會員大會開會通知

第八屆會員大會謹訂於 9/18 日假新竹清華大學國際會議廳舉行，本次會議除了例行之年度會務工作報告外，並將舉行第八屆理事及監事改選，開會通知將於八月中旬郵寄予本會會員，通知書內將夾附「委託書」及「選票」，若確定無法出席本次會議之會員者，請簽妥「委託書」及「選票」後寄回學會，以確保會員之權利。

### 歡迎加入 SIGHAN

國際計算語言學學會(ACL)所屬的 Special Interest Group(SIGHAN)，為從事中文處理研究同仁發起成立的，目前負責人為 Richard Sproat，歡迎大家踴躍加入，加入會員免入會費及年費，欲入會者只需提供個人基本資料(姓名、email)至 Mr. Tom Emerson (tree@sighan.org) 即可，SIGHAN 相關訊息請參閱 <http://www.sighan.org/constitution>。

# 第十五屆計算語言學研討會 ( ROCLING XV)

## 報 名 表 ( Registration Form )

日期：92年9月18日、19日

地點：國立清華大學國際會議廳

姓 名		E - m a i l	
單 位		職 稱	
電 話		傳 真	
通 訊 地 址		收 據 抬 頭	
參 加 費 用	註冊費：(含研討會費、論文集、晚宴) P.S：午餐自理，於網頁提供相關餐飲資料 一般人士及教師 非會員 NT\$ 2500 會員 NT\$ 1500 入會費 NT 1000 學生 非會員 NT\$ 1000 會員 NT\$ 500 入會費 NT 500 贊助單位：免費(名額五名) 非會員若同時申請加入會員，即可享有會員價，入會申請書請向學會索取 TEL：02-2788-3799 轉 1502 黃琪，或逕自上網下載 (http://rocling.iis.sinica.edu.tw/ROCLING/rocintrc.htm)。 九月一日以後報名者註冊費加收三百元。 以學生身份報名者請附上學生證影本(本會學生會員免附)。 報名前 50 者贈遠流出版“詞的學問”一本(簡介網址如下) <a href="http://www.ylib.com/search/ShowBook.asp?BookNo=LF012">http://www.ylib.com/search/ShowBook.asp?BookNo=LF012</a>		
	晚宴： 素食 非素食 三、代訂房間：國立清華大學第一招待所 住宿期間：到達日期：_____ 時間： _____ 離開日期：_____ 時間： _____ 單人房(600 元) 雙人房(800 元) 1.兩張單人床 2.一張雙人床 備註： 1. 住宿費請至第一招待所退房時自行支付。 2. 住宿停車時請將住宿收據置於擋風板前，以供校警隊辨識。 3. 房間名額有限，請先電話詢問確定房間，報名從速(詢問電話：0968098318 簡同學)。		

1. 報名及繳費方式：

郵政劃撥：帳號：19166251，中華民國計算語言學學會(請註明匯款目的)

信用卡：Card type: ( Visa Master-card JCB AE Card )

姓名：\_\_\_\_\_ (請以正楷書寫) 發卡銀行：\_\_\_\_\_

卡號： - - - 有效日期：\_\_\_\_\_

金額：\_\_\_\_\_ 持卡人簽名：\_\_\_\_\_

支票(請加註抬頭：中華民國計算語言學學會)

2. 報名表一人一份(可影印使用)

請將此報名表郵寄或傳真至：

300 新竹市光復路二段 101 號 清華大學資訊系統與應用研究所 黃淑真小姐收

電話：(03)5715131 轉-3530 傳真：(03)572-3694

E-mail：G904374@oz.nthu.edu.tw WWW: http://140.114.75.14/rocling/

第十五屆計算語言學研討會  
資訊檢索研討會  
語音訊號處理研討會  
電腦輔助語言教學研討會

暫訂議程

9月18日

- 09:00~10:00 邀請演講 (鄭良偉教授 教育部國語會主委)
- 10:00~10:30 休息
- 10:30~12:00 Technical Paper Session 1
- 12:00~13:00 午餐
- 13:00~14:00 計算語言學學會會員大會
- 14:00~15:30 Technical Paper Session 2    14:00~17:30 Poster/Interactive Demo
- 15:30~16:00 休息
- 16:00~17:30 Technical Paper Session 3
- 18:30~20:30 晚宴

資訊檢索研討會

10:00~17:30

9月19日

- 09:00~10:00 邀請演講 (中央研究院 鄭錦全院士)
- 10:00~10:30 休息
- 10:30~12:00 Technical Paper Session 4    10:30~12:00 Student Workshop
- 12:00~13:00 午餐

語音訊號處理研討會

- 13:30~14:20 演講：當前語音技術研究之趨勢與展望
- 14:30~15:20 演講：車內環境之對話系統
- 15:40~16:20 座談會：電信學門語音處理領域發展與國際合作

13:30~17:00 國科會研究計畫成果發表(壁報)

電腦輔助語言教學研討會

- 13:30~14:20 演講：電腦輔助語言教學的趨勢與展望
- 14:30~15:20 演講：數位學習
- 14:30~15:20 演講：國家型數位學習計畫 - CANDLE

13:30~17:00 國科會研究計畫成果發表(壁報)

附註：報名「第十五屆計算語言學研討會」可同時免費參加「資訊檢索研討會」、「語音訊號處理研討會」及「電腦輔助語言教學研討會」，若只單獨參加上列三個研討會者，請直接至「第十五屆計算語言學研討會」網站上報名，報名費現場繳交。

## 其他特約飯店住宿介紹

下列為第十五屆計算語言學研討會的特約飯店，表列中的價格為參加第十五屆計算語言學研討會的優惠價，訂房時請事先說明是參加清華大學第十五屆計算語言學研討會的來賓，退房時請出示研討會收據結帳。

元首大飯店 地址：新竹市建功一路 29 號 電話：03-5711111 <a href="http://www.h2ocity.com.tw/shop.asp?shop_id=5986">http://www.h2ocity.com.tw/shop.asp?shop_id=5986</a>
價目表
1000 NT/2 人 1400 NT/3 人 1800 NT/4 人
*附早餐*含服務費

柏克萊商務 地址：新竹市光復路二段 272 號 電話：03-5728035 <a href="http://98.to/bbhotel8888/">http://98.to/bbhotel8888/</a>
價目表
標準房 2000 NT/人 2270 NT/2 人
*附早餐*含服務費*房間為十多坪的大小

新苑庭院大飯店 地址：新竹市大同路 11 號 電話：03-5260522 <a href="http://www.shineyou.com.tw/shinyuan/">http://www.shineyou.com.tw/shinyuan/</a>
價目表
高級單人房 NT\$2300
豪華單人房 NT\$2700
新苑套房 NT\$3200
新苑豪華套房 NT\$3900
新苑總統套房 NT\$7900
以上房間價目為原價，住宿可享六折優待
*附早餐*含服務費

# The Map Task 語料庫簡介

解志強

清華大學外語系

在計算語言學通訊第十四卷第二期當中，曾淑娟教授在介紹口語語料庫時，曾經提到 The Map Task Corpus 語料庫。Map Task 的發明者 Professor Gillian Brown 在英國學術界極受敬重，她是我在劍橋大學讀 MPhil 學位時的所長。後來我到愛丁堡大學，又接觸到 HCRC 這個建構 Map Task 語料庫的研究中心

(語料庫參考網頁：<http://www.hcrc.ed.ac.uk/dialogue/maptask.html>)。因此，我對於這個語料庫很有興趣。

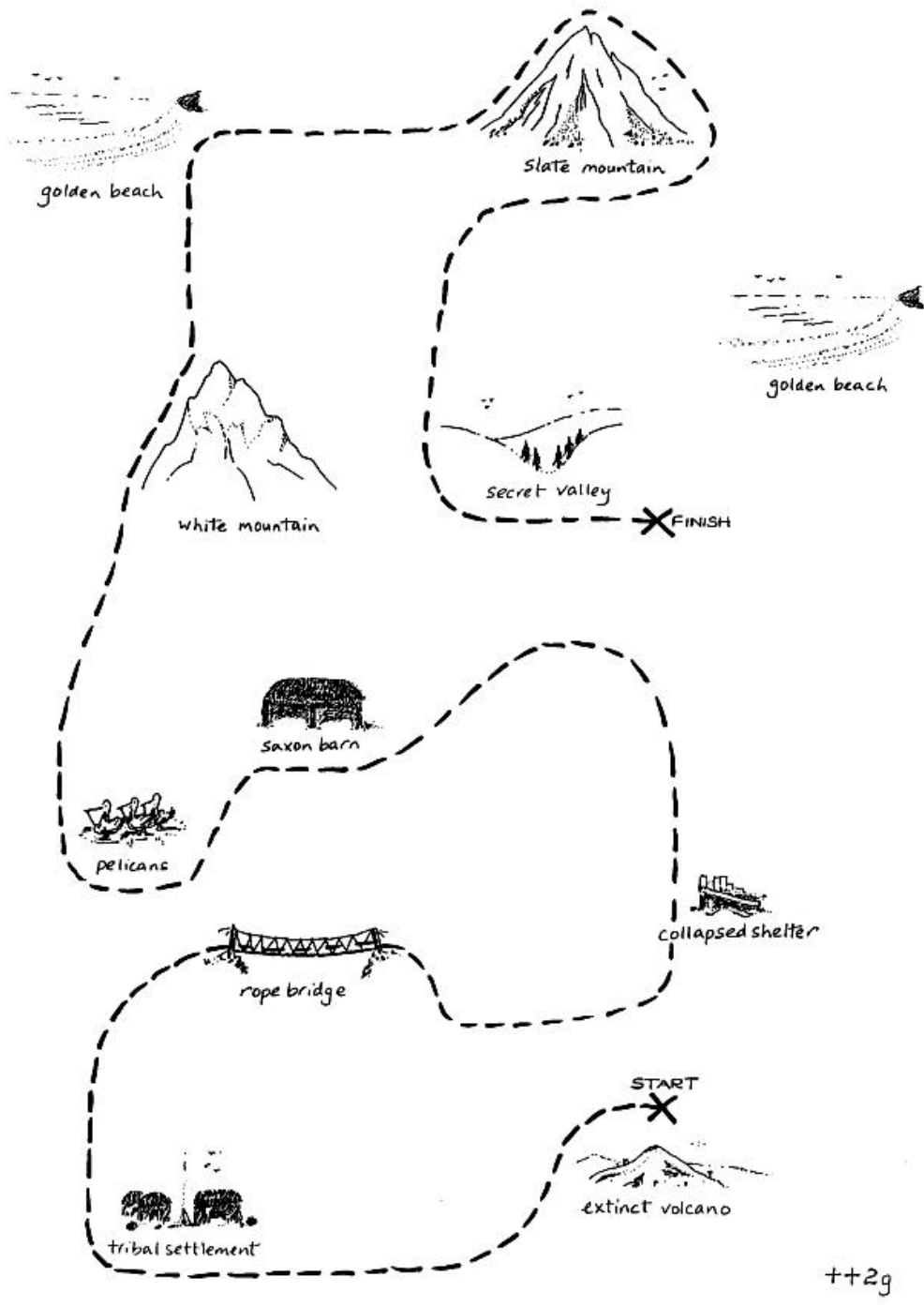
Map Task 基本上要求兩個學生各自持一份簡單的地圖進行對話。其中一個叫做 Giver (給予指示的人)，另一個叫做 Follower (接受指示的人)。Giver 手中所握有的地圖(如附件一)，除了火山，吊橋，農場，石頭等等約十個左右的地標外，還有繞行這些地標的一個路線圖。但是 Follower 手中地圖沒有路線(如附件二)，而且還有一些‘陷阱’ – 他的地表有些和 Giver 相同，有些不同，有些則缺乏。於是 Giver 便在整個實驗過程中，在互相看不見對方地圖的情況下，引導 Follower 儘量畫出正確的路線圖(對話樣本如附件三)。

Gillian Brown 教授在她的 1995 年著作 *Speakers, Listeners and Communication* (Cambridge University Press) 裡面，詳細描述了作 Map Task 的人，如何使用指稱(reference)，搜索範圍(search field)等觀念來引導對方，互動描繪路線圖通常如何進行，以及溝通能力強弱的人各有那些談話特色等等。是有關 Map Task 的言談分析裡面最詳盡有用的一份參考資料。

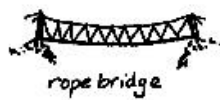
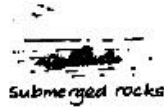
因為是一個口語語料庫，所以 The Map Task 語料庫當然在 phonology(音韻學)和 disfluency(不順暢)的研究上會是個良好資源。尤其是該語料庫在 disfluency coding(不順暢現象標記)上做的很完整，舉凡說話時的重複，改口，停頓，追加等各種言談繞口現象，都可以從這裡汲取資源來研究。

我個人認為 Map Task 可以用來作為建立一個互動式的英語口語教學模式。我理想中的一個系統，就是電腦可以扮演 Giver 或是 Follower 的角色，來與學英語的學生共同進行 Map Task。電腦具有自然語言處理(natural language processing)、自然語言了解(natural language understanding)，與自然語言產生(natural language generation)的能力，在經過 Map Task 實際語料的訓練後，它可以擔任 Giver，指示學生操作電腦螢幕上的地圖，和學生對話，使學生可以畫出理想中的路線圖。另一方面，電腦也可以擔任 Follower，接受學生的命令，按照學生的指示畫出學生所掌握的路線圖。我相信這個系統如果能夠研發出來，將是一個非常有趣的系統，可以和學生一起進行非常有價值的溝通式教學活動。

我個人才疏學淺，到目前還無法進行我想像中的這個系統的研發，僅能在此作一個簡單說明，或許有其他有興趣及有能力的老師或學生，願意從事此一研究。以下網址為 HCRC 有關 The Map Task 的一些著作的網頁：<http://www.hcrc.ed.ac.uk/maptask/maptask-papers.html>



++2g



++2f

附件三: Map Task 其中一段的開始幾句對話 (從  
<http://www.ltg.ed.ac.uk/~amyi/maptask/demo.html> 編號 q4nc1)

GIVER: right, you got a map with an extinct volcano?  
-----

FOLLOWER: right yes i have, i'm just in front of that.  
-----

GIVER: right.  
-----

FOLLOWER: with the start.  
-----

GIVER: right, you've got a cross marked start?  
-----

FOLLOWER: yes.  
-----

GIVER: right, if you just want to come ... .. like down past the extinct volcano ... down to like to  
towards the bottom of the page.  
-----

FOLLOWER: right okay, just straight down directly south?  
-----

GIVER: uh-huh ... just straight down, uh south.  
-----

FOLLOWER: how far?  
-----

GIVER: south down to about ... equal to where the tribal settlement is ... further on.  
-----

FOLLOWER: equal to where the tribal settlement is.  
-----

GIVER: just a wee bit below it because you're going to ... .. go start going ehm west you.

# ACL03 參加會議心得

吳鑑城

清華大學資訊工程研究所博士班

## 壹、前言

國際計算語言學學會(Association for Computational Linguistics (ACL))為計算機語言學學界中，最重要的一個國際性組織，除在歐洲設有分會外，亦成立了美洲分會，而亞洲部份因為受制於日本，目前尚未成立分會。學會重要的常態活動，一是出版 Computational Linguistic 期刊，另外則是每年所舉辦的國際學術會議。ACL 會議 ( Annual Meeting of the Association for Computational Linguistics ) 歷史悠久，從 1998 年起，改變了長年在北美洲舉辦會議的慣例，而輪流在美、歐、亞三洲舉辦。

## 貳、大會記實

今年的七月五日到七月十二日，ACL 就在日本北海道札幌市的 Sapporo Convention Center 召開第四十一屆大會。此次大會，並沒有受到 SARS 影響，參與的學者高達六百多人，足見語言學界對此大會的重視。我和葉至程同學，一起隨著指導教授張俊盛博士，發表我們獲選的研究成果，並且做系統之展示。

ACL 主要的研究課題為計算語言學及其相關領域。此會議之主要目的是希望能夠集合世界各地的研究人員，交換其在此領域之研究心得，意見，知識，以及相關技術。今年所發表的文章主要領域包含眾多，諸如：

- Question Answering
- Machine Translation
- Machine Learning
- Parsing and Grammar Formalisms
- Conference in Discourse and Dialogue
- Information Retrieval
- Summarization
- Chunking
- Named Entities and Bootstrapping
- Word Sense Disambiguation
- Phonology and Morphology

等等領域，都有不錯的技術論文發表。

本次大會共錄取了約一百篇論文。令人惋惜的是，絕大部分的論文是由美洲國家所發表，亞洲部份的論文明顯少了許多。希望在不久的將來，我國亦能夠有更多優秀的論文在此發表。在此次會期之中共計有四個 tutorials、二十四個技術分場(main session conference program)、九個研討會(workshop)、兩場海報/展示 (poster/demo)以及三個學生研討會(student workshop)。每篇獲選於 main conference 和 student workshop 的論文，以二十五分鐘的時間向與會人員報告心得。而 poster/demo 則是在七月九日、十日除了於會場張貼相關的成果海報、介紹外，亦需以五分鐘的時間做簡短的報告。

大會今年的終生成就獎頒發給第二十三任的日本京都大學校長 Makoto Nagao 博士，表彰他提出的範例式的機器翻譯模型的學術貢獻。最佳論文獎由兩組研究人員共同獲得，分別是：

Dan Klein 和 Christopher D. Manning 教授發表的 Accurate Unlexicalized Parsing，以及 Yukiko Nakano, Gabe Reinstein, Tom Stocky, 與 Justine Cassell 教授發表的 Towards a Model of Face-to-Face Grounding。

此次我們所獲選的 Poster/demo 論文是 “TotalRecall : A Bilingual Concordance for Computer Assisted Translation and Language Learning ”。相對於 main conference，Poster/demo 能夠直接面對有興趣的來賓，將自己的研究成果對其做詳盡的解說和展示。除了有機會訓練自我英文的聽說能力外，還能夠與各地的學者專家交換彼此的心得，並且進一步的提升自己的研究視野，也對他人的研究成果有進一步的認識，著實是一個不錯的體驗。

TotalRecall 為數位學習國家型計畫 前瞻性數位英文學習中心 CANDLE 的一部分。其主要功能是利用目前難得的大量雙語語料庫 光華雜誌，與自然語言處理技術，來幫助英語學習。這份語料在經由近年在清華研發出來的高準確性之句子，以及詞彙搭配對應的處理後，可提供單語，雙語的辭彙查詢。TotalRecall 利用網頁的方式，提供查詢，以及結果之呈現之介面。

當在 TotalRecall 輸入中文或英文的 Query 時，除了在單語模式可以查詢到許多的包含此關鍵字的對應文句外，在雙語模式，更能夠自動的將輸入之關鍵字，以及其在另一種語言中的相對翻譯自動地將其標記，讓查詢者能夠輕易的取得查詢的相關資訊，並且獲得大量的對應文句的資料。在示範過程中，有許多有著濃烈興趣的學者來聽取我們的解說，並且都給予 TotalRecall 相當高的評價。我們目前更持續地整合 collocations 方面的資訊到此系統中。相信 TotalRecall，必可有效地幫助電腦輔助翻譯與英語學習。因研究尚在進行中，且有語料版權上的授權問題，所以目前此系統僅限於研究團隊中的人員做測試及修改。未來可望成為全民學習英語的輔助學習工具。

### 參、其他心得

在此次會議中，舉凡詳盡的住宿資訊、完善的地圖資訊、午餐接送專用巴士、甚至乎地鐵一日票販售的提供，都讓我們深刻的體驗主辦單位安排的細心。唯一美中不足的，是大會工作人員並非全部都精通英文，因此有時會有些溝通上的問題。

台灣的計算語言學界，在這十幾年來的辛苦耕耘後，也有了相當的成就。去年(2002年)，更向國際組織 ICCL (International Committee on Computational Linguistics) 爭取到 Coling 國際會議在台灣主辦的主辦權，並且邀請到國內外知名的學者來台與會。今年由中華民國計算機語言學學會主辦的 Rocling 會議(Research on Computational Linguistics Conference)，亦將於清華大學舉辦第十五屆大會。今年大會特別參考了國際研討會的形式，舉辦了第一屆的學生研討會(Student workshop)。特別新增此學生研討會之目的，主要在於鼓勵國內的學生，使其能有機會發表自己的研究。而從事相關領域研究的學者也能夠給予建議，如此一來，必可使國內學生所發表之論文水準，能夠更上一層樓。(會議及投稿等相關訊息，歡迎至學會網頁 <http://rocling.iis.sinica.edu.tw/ROCLING/> 查詢。)未來，若能獲得教育部以及國科會等機構的鼎力支持，相信台灣必能夠在語言學界中，發揮出更大的影響力，並且建立起更佳的形象。

**2nd International Wordnet Conference**  
**Masaryk University, Brno (Czech Republic)**  
**<http://www.fi.muni.cz/gwc2004/>**  
**20-23 January 2004**

Details about the conference can be found on the conference website:

<http://www.fi.muni.cz/gwc2004/>

Details about the Association can be found on the GWA website:

<http://www.globalwordnet.org/>

The presentations will be categorized as:

oral presentation

poster presentation (in special Poster session)

Submissions will have to state the preferred presentation category. Acceptance may be subject to changes in the presentation format following the recommendation of the reviewers.

Papers submitted to GWC 2004 must not be under review by any other conference or publication during the GWC review cycle, and must not be previously published or accepted for publication elsewhere.

Final papers should be submitted in electronic form (Word or LaTeX in the LNCS style).

The conference program will include oral presentations and two special sessions -- Poster session (for parallel poster presentations) and Demonstration session, where authors are invited to present actual projects, developed software or interesting material relevant to the topics of the conference. The demonstrations presented will not appear in the printed version of the Proceedings of GWC'2004. The authors of the demonstrations should provide an abstract, which will be included in the CD version of the Proceedings.

The deadline for submissions is September 1st, 2003. Proposals for demonstrations are accepted until the registration deadline (December 5). Decisions regarding acceptance will be announced to the authors in early October, the deadline for the final papers is 1st of November.

The Proceedings will be published by Masaryk University in both paper and CD format.

**Early Registration information:**

Early registration: until October 20th;

Final registration (bank or card transfers should have been received): December 5th;

Fees and prices cannot be guaranteed for on site registration;

**Important dates:**

01-Sep-2003 Deadline for paper submission

20-Oct-2003 Deadline for early registration

01-Nov-2003 Final Papers

05-Dec-2003 Deadline for Registration (bank payments should have been received)

20/23-Jan-2004 Conference

# **4th INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION**

The fourth international conference on Language Resources and Evaluation, LREC 2004, is organised by ELRA in cooperation with other Associations and consortia, national and international organisations.

Centro Cultural de Belem, Lisbon, Portugal

Pre-conference workshops: 24-25 May 2004

Main conference: 26-27-28 May 2004

Post-conference workshops: 29-30 May 2004

Conference web site: <http://www.lrec-conf.org/>

## **CONFERENCE AIMS**

In the Information Society, the pervasive character of Human Language Technologies (HLT) and their relevance to practically all fields of Information Society Technologies (IST) has been widely recognised.

Two issues are particularly relevant: the availability of Language Resources (LRs) and the methods for the evaluation of resources, technologies, products and applications.

Substantial mutual benefits are achieved by addressing these issues through international collaboration.

The term "language resources" (LRs) refers to sets of language data and descriptions in machine readable form, used in many types of areas/components/systems/applications:

- creation and evaluation of natural language, speech and multimodal algorithms and systems,
- software localisation and language services,
- language enabled information and communication services,
- knowledge management,
- e-commerce, e-publishing, e-learning, e-government,
- cultural heritage,
- linguistic studies,
- etc.

This large range of uses makes the LRs infrastructure a strategic part of the e-society, where the creation of a basic set of LRs for all languages must be ensured in order to bring all languages to the same level of usability and availability.

Examples of LRs are written or spoken corpora and lexica, which may be annotated or not, multimodal resources, grammars, terminology or domain specific databases and dictionaries, ontologies, multimedia databases, etc. LRs also cover basic software tools for the acquisition, preparation, collection, management, customisation and use of the above mentioned examples.

The relevance of evaluation for language technologies development is increasingly recognised. This involves assessing the state-of-the-art for a given technology, measuring the progress achieved within a programme, comparing different approaches to a given problem, assessing the availability of technologies for a given application, benchmarking, and assessing system usability and user satisfaction.

The aim of this conference is to provide an overview of the state-of-the-art, discuss problems and opportunities, exchange information regarding LRs, their applications, ongoing and planned activi-

ties, industrial uses and needs, requirements coming from the new e-society, both with respect to policy issues and to technological and organisational ones. LREC will also elaborate on evaluation methodologies and tools, explore the different trends and promote initiatives for international collaboration in the areas mentioned above.

### **CONFERENCE TOPICS**

Examples of the topics which may be addressed by papers submitted to the conference are given below.

#### **\*\*\* Issues in the design, construction and use of Language Resources (LRs)\*\*\***

Guidelines, standards, specifications, models and best practices for LRs,  
Methods, tools and procedures for the acquisition, creation, management, access, distribution and use of LRs,  
Methods for the extraction and acquisition of knowledge (e.g. terms, lexical information, language modelling) from LRs,  
Organisational and legal issues in the construction, distribution, access and use of LRs,  
Availability and use of generic vs. task/domain specific LRs,  
Definition and requirements for a Basic and Extended Language Resource Kit (BLARK, ELARK) for all languages,  
Monolingual and multilingual LRs,  
Multimedia and multimodal LRs. - Integration of various media and modalities in LRs (speech, vision, language),  
Documentation and archiving of languages, including minority and endangered languages,  
Ontologies and knowledge representation,  
Terminology, term extraction, domain-specific dictionaries,  
LRs for linguistic research in human-machine communication,  
Exploitation of LRs in different types of applications (information extraction, information retrieval, speech dictation, translation, summarisation, web services, semantic web, etc.),  
Exploitation of LRs in different types of interfaces (dialog systems, natural language and multimodal/multisensorial interactions, etc.)  
Industrial LRs requirements, user needs and community's response,  
Industrial production of LRs,  
Industrial use of LRs,  
Metadata descriptions of LRs.

#### **\*\*\* Issues in Human Language Technologies (HLT) evaluation \*\*\***

Evaluation, validation, quality assurance of LRs,  
Evaluation methodologies, protocols and measures,  
Benchmarking of systems and products, resources for benchmarking and evaluation, blackbox, glassbox and diagnostic evaluation of systems,  
Usability and user experience evaluation, qualitative and perceptive evaluation,  
Evaluation in written language processing (document production and management, text retrieval, terminology extraction, message understanding, text alignment, machine translation, morphosyntactic tagging, parsing, semantic tagging, word sense disambiguation, text understanding, summarisation, question answering, localisation, etc.),  
Evaluation in spoken language processing (speech recognition and understanding, voice dictation, oral dialog, speech synthesis, speech coding, speaker and language recognition, spoken translation, etc.),  
Evaluation of multimedia document retrieval and search systems (including detection, indexing, filtering, alert, question answering, etc),  
Evaluation of multimodal systems,  
From evaluation to standardisation.

\*\*\* General issues \*\*\*

National and international activities and projects,  
LRs and the needs/opportunities of the emerging industries,  
LRs and contributions to societal needs (e.g. e-society),  
Priorities, perspectives, strategies in national and international policies for LRs,  
Needs, possibilities, forms, initiatives of/for international cooperation, and their organisational and technological implications,  
Open architectures for LRs.

The Conference targets the integration of different types of LRs (spoken, written and other modalities) and of the respective communities. To this end, LREC encourages submissions covering issues which are common to different types of Language technologies, such as dialog strategy, written and spoken translation, domain-specific data, multimodal communication or multimedia document processing, and will organise, in addition to the usual tracks, common sessions encompassing the different areas of LRs.

### **PROGRAMME**

The Scientific Programme will include invited talks, oral presentations, poster presentations, referenced demonstrations and panels. There is no difference in quality between oral presentations and poster presentations.

Only the appropriateness of the type of communication to the content of the paper will be considered.

### **FORMAT FOR ABSTRACT SUBMISSION**

Submitted abstracts of papers for oral and poster presentations should consist of about 800 words. Demonstrations of LRs and related tools will be reviewed as well. You should send an outline of about 400 words. If a demo is connected to a paper, please attach the outline to the paper abstract. A limited number of panels and workshops is foreseen, proposals will be reviewed by the Programme Committee.

For panels, please send a brief description, including an outline of the intended structure (topic, organiser, panel moderator, tentative list of panelists).

For workshops, see the dedicated section below.

Only electronic submissions will be considered. Further details about submission will be circulated in the 2nd call for papers and posted on the LREC web site ([www.lrec-conf.org](http://www.lrec-conf.org)).

### **IMPORTANT DATES**

Submission of proposals for panels and workshops: 20th October 2003

Submission of proposals for oral and poster papers, referenced demos: 31st October 2003

Notification of acceptance of workshop and panel proposals: 14th November 2003

Notification of acceptance of oral papers, posters, referenced demos: 23rd January 2004

Final versions for the proceedings: 1st March 2004

Conference: 26th 28th May 2004

Pre-conference workshops: 24th and 25th May 2004

Post-conference workshops: 29th and 30th May 2004

The proceedings of the conference will include both oral and poster papers.

Internet connections and various computer platforms and facilities will be available at the Conference site. In addition to referenced demos concerning LRs and related tools, it will be possible to run unreferenced demos of language processing products, systems and tools. Those interested should contact the organiser of the demonstrations (details will be posted on [www.lrec-conf.org](http://www.lrec-conf.org)).

### **WORKSHOPS**

Pre-conference workshops will be organised on 24th and 25th May 2004, and post-conference work-

shops on 29th and 30th May 2004. A workshop is normally either half day or full day.

Proposals for workshops should be no longer than three pages, and include:

A brief technical description of the specific technical issues that the workshop will address.

The reasons why the workshop is of interest this time.

The names, postal addresses, phone and fax numbers and email addresses of the workshop organising committee, which should consist of at least three people knowledgeable in the field, coming from different institutions.

The name of the member of the workshop organising committee designated as the contact person.

A time schedule of the workshop and a preliminary agenda.

A summary of the intended workshop call for participation.

A list of audio-visual or technical requirements and any special room requirements.

The workshop proposers will be responsible for the organisational aspects (e.g. workshop call preparation and distribution, review of papers, notification of acceptance, assembling of the workshop proceedings, etc.).

Further details about submission will be circulated in the 2nd call for papers and posted on the LREC web site: [www.lrec-conf.org](http://www.lrec-conf.org).

Proceedings will be printed for each workshop.

### **CONSORTIA AND PROJECT MEETINGS**

Consortia or projects wishing to take this opportunity for organising meetings should contact the ELDA office,

[lrec@elda.fr](mailto:lrec@elda.fr) (further details are given at the end of the document).

### **CONFERENCE PROGRAMME COMMITTEE**

Nicoletta Calzolari, Istituto di Linguistica Computazionale del CNR, Pisa, Italy

Khalid Choukri, ELRA, Paris, France

Teresa Lino, Universidade Nova de Lisboa, Lisbon, Portugal

Bente Maegaard, CST, Copenhagen, Denmark

Joseph Mariani, LIMSI-CNRS, Orsay, France

Jan Odijk, UIL-OTS, Utrecht, the Netherlands, and ScanSoft, Merelbeke, Belgium

Daniel Tapias, Telefonica Moviles, Madrid, Spain

Antonio Zampolli, Istituto di Linguistica Computazionale del CNR, Pisa, Italy (Conference chair)

The composition of the committees as well as instructions and addresses for registration and accommodation will be detailed on the LREC web site at [www.lrec-conf.org](http://www.lrec-conf.org) and will be announced in the 2nd call for papers to be issued at the end of July.

### **ELRA**

For more information about ELRA (European Language Resources Association), please contact:

Khalid Choukri, ELRA CEO

55-57 Rue Brillat-Savarin,

75013 Paris - France

Tel: + 33 1 43 13 33 33

Fax: + 33 1 43 13 33 30

Email: [choukri@elda.fr](mailto:choukri@elda.fr)

Web: <http://www.elra.info/> or <http://www.elda.fr/>

The first LREC was organised in Granada (Spain) in 1998: 197 papers and posters were presented, with about 510 registered participants from 38 different countries from all continents. Among these, the largest group came from Spain (81 participants), followed by France (75), USA (73), Germany (47), UK (43) and Italy (41). Registered participants belonged to over 325 different organisations.

LREC 2000, in Athens, had 129 oral papers and 152 posters presented, with around 600 participants from 51 different countries from all continents. Among these, the largest group came from Greece (117), followed by USA (70), France (59), Germany (45), UK (43), Japan (35) and Italy (29). Registered participants belonged to 319 different organisations.

LREC 2002, which took place in Las Palmas de Gran Canaria (Spain), attracted over 700 representatives, coming from 38 countries around the world, who could take advantage of the numerous oral and poster presentations (around 365, covering every area in HLT).

A similar number of participants is expected at LREC 2004.