

# Association for Computational Linguistics

## 2002 Annual Reports

### Contents

#### **1. Reports from ACL management**

- 1.1 President's report (John Nerbonne) –p.3
- 1.2 Secretary's report (Sandra Carberry) –p.5
- 1.3 Treasurer's report (Kathleen F. McCoy) –p.6
- 1.4 ACL European Chapter Chair (Donia Scott) –p.9
- 1.5 ACL European Chapter Treasurer (Mike Rosner) –p.11
- 1.6 ACL North American Chapter Chair (Diane Litman) –p.14
- 1.7 ACL North American Chapter Treasurer (Dragomir Radev) –p.17
- 1.8 Special Liason with Asia (Jun'Ichi Tsujii) –p.18
- 1.9 Nominating Committee (Philip Cohen, Eduard Hovy , Wolfgang Wahlster) –p.19

#### **2. ACL-2002**

- 2.1 General Conference Chair (Pierre Isabelle) –p.20
- 2.2 Program Co-Chairs (Eugene Charniak, Dekang Lin ) –p.23
- 2.3 Local Arrangements (Martha Palmer) –p.25
- 2.4 Tutorial Chairs (Hwee Tou Ng) –p.27
- 2.5 Publication Chair (Dragomir Radev) –p.28
- 2.6 Student Session Chairs (Alexander Koller, Gideon Mann) –p.29
- 2.7 Workshop Chair (Walter Daelemans) –p.34
- 2.8 Exhibits/Sponsorship Chair (Deborah Dahl) –p.35
- 2.9 Demos Chair (Ronnie Smith) –p.37

#### **3. Journal and Publications**

- 3.1 Journal (Julia Hirschberg) –p.37
- 3.2 Book Reviews (Graeme Hirst) –p.39
- 3.3 Squibs and Discussions (Pierre Isabelle) –p.40
- 3.4 Book Series (Bran Boguraev, Steven Bird) –p.41

3.5 ACL Anthology (Steven Bird) –p.42

#### **4. Recent Conferences**

4.1 HLT'02 (Kathleen McCoy) –p.46

#### **5. Future Conferences**

5.1 ACL-03 local arrangements (Kenji Araki) –p.47

5.2 ACL-04 site selection (Mark Johnson) –p.51

#### **6. Special Interest Groups**

6.1 General (Mark Johnson) –p.51

6.2 SIGDAT (Kenneth W. Church, David Yarowsky) –p.52

6.3 SIGDIAL (Laila Dybkjaer) –p.52

6.4 SIGGEN (Ehud Reiter) –p.55

6.5 SIGHAN (Richard Sproat) –p.55

6.6 SIGLEX (Adam Kilgarriff) –p.57

6.7 SIGMEDIA (Elisabeth Andre) –p.58

6.8 SIGMOL (Dick Oerhle) –p.59

6.9 SIGNLL (Walter Daelemans) –p.60

6.10 SIGPARSE (Harry Bunt) –p.61

6.11 SIGPHON (Jason Eisner) –p.62

6.12 SIGSEM (Patrick Blackburn) –p.63

#### **7. Organizations/Initiatives**

7.1 ACL Website (Dragomir Radev) –p.64

7.2 ACL Archives (Richard Sproat) –p.65

7.3 ACL Course Survey (Robert Dale and Mary Taffet) –p.65

#### **8. Other Reports**

8.1 NLSR (Thierry Declerck) –p.66

8.2 OLAC (Steven Bird) –p.67

## 1.1 President's Report for 2002

John Nerbonne

Just a little over forty years ago, on June 13, 1962, Victor Yngve, David Hayes, Harry Josselson, Ida Rhodes, Paul Garvin, Winfred Lehman, Richard See, Anthony Oettinger and Sydney Lamb met to establish the Association for Machine Translation and Computational Linguistics (AMTCL), which held its first Annual Meeting the following year. These pioneers wished to "encourage high standards by sponsoring meetings, publication and other exchange of information." The Association changed its name, but not its goals or organization, in 1968 to the Association for Computational Linguistics.

The year July, 2001 through June, 2002 is thus the fortieth anniversary year of the ACL, and two activities were initiated to celebrate this appropriately.

- The ACL Anthology is a collection of all the publications associated with ACL and COLING in the forty years of their work, including all the ACL conference and workshop proceedings, back issues of journals, COLING proceedings. Steven Bird proposed and then organized this enormous undertaking, which will make the results of computational linguistics research much more generally available. There is a special section in the journal and publications section of this report (below) on this project.

- The ACL Lifetime Achievement Award was suggested to me and Phil Cohen by Erhard Hinrichs. The executive committee resolved to present this award at most annually in recognition of scientific and technical achievement. The recipient will be invited to speak on a topic of his choice at the Annual Meeting. While the primary motivation for establishing this award is to recognize the same "high standards" our founders sought to encourage, we also hope that the award will publicize our work in larger circles.

Naturally, several ongoing projects continued to occupy the association throughout this period.

- The relation of the international organization to its regional chapters has begun crystallizing even while it continues to occupy us. Regional chapters host regional meetings and co-host (international) ACL meetings held in their regions, and they share in the financial risks and benefits of all those conferences. This financial structure is now in place. In addition to hosting meetings, chapters will also take (regional) education into their portfolios, including the support of summer schools and the like.

In this vein the EACL has supported European Summer School in Logic, Language and Information (ESSLLI) and NAACL cooperates with the Johns Hopkins summer school.

- Asian researchers have introduced a new umbrella federation of organizations in response to ACL's internationalization push. This differs from the sort of chapter structure now in place in Europe and North America, but Asian researchers do wish to host international conferences regularly and to participate in the international organization.

- We continue to investigate greater financial autonomy for special interest groups.

- The executive committee has begun investigating a fee structure that would accommodate people that wished to attend only a single satellite workshop at one of the annual meetings or chapter meetings.

- The conference handbook, initiated by Ed Hovy, is being extended and maintained by the Association's secretary, Sandee Carberry, and it is already streamlining many issues of planning, responsibility and organization.

- The North American chapter has embarked on a course of collaboration with the Human Language Conferences, originally sponsored by DARPA in the 1980's. The report of the chapter chair, Diane Litman, contains further details.

- We established mutual reductions membership rates for members of ISCA who wish to join ACL (and vice versa) in an experiment to try to involve more speech researchers in ACL.

Finally, I note issues of importance which were not foci of attention in 2001-2002.

- Publicity and public awareness of Computational Linguistics. Although recognized as valuable, public relations efforts inevitably need to showcase individuals and their specific accomplishments. ACL support for proposals has foundered in the past on the reluctance of the executive committee to be involved with these decisions, including the reluctance to delegate them to others. This is clearly suboptimal, since it ends up leaving the ACL with no systematic influence on publicity about the field.

- Industrial involvement. While the current conference guidelines always foresee an applications track,

designed to supersede the old "Applied Natural Language Processing", industrial involvement seems disproportionately low in the organization, meaning in particular membership, participation at meetings, active support in the organization.

The organization continues to be healthy in its membership numbers and its balance sheet, and for this we should be grateful to Treasurer Kathy McCoy, and Office Manager Priscilla Rasmussen, who are now complemented essentially by Association Secretary Sandee Carberry.

John Nerbonne, Groningen  
ACL President, 2002

## **1.2 Secretary's Report for 2002**

The ACL Board met on February 2, 2002 in Brighton, England. Minutes are posted on the web site in the ACL archive.

A Constitutional amendment to change the timing of officer elections was submitted to the membership and approved. The revised section of the Constitution now reads:

7. Elections shall be conducted as follows:
  - A. The Nominating Committee shall nominate (normally two) people who are willing to serve in each position to be filled at an annual election.
  - B. The Secretary shall announce these nominees at the annual meeting (normally held during the summer), post their names on the Association web site, and invite the membership to submit additional nominations.
  - C. Additional nominations will be accepted by email, starting immediately after the conference and ending one month later, provided that two members of the association second each such nomination and provided that evidence is presented that such a nominee will serve if elected.
  - D. Nominees will be given at least two weeks to provide a statement of their vision for ACL; these vision statements will be posted on the web site for viewing by the membership.
  - E. After the nomination and posting of vision statements, the Secretary will mail a ballot containing the final list of nominees for each position, along with information about the URL containing the vision statements, to the membership. Voting will be performed in a manner specified in the ballot, which manner will be open to all members of the Association, secret, verifiable, and overseen by

an executive committee member (normally the Secretary) who is not running for office, and who is acceptable to all nominees. Any nominee receiving the most votes cast will be declared elected and the results will be posted on the Association's web site. Should no candidate receive a majority of the votes cast, subsequent ballots will be held until one candidate receives a majority. Those elected shall take office on the first of January following the vote and serve until succeeded.

- F. In the event that the annual meeting is not held during the summer, the Executive Board may vote to alter the timing of the election for that year.

A newsletter containing news about the organization was mailed to all members in January, 2002.

In order to facilitate better documentation of ACL policies, practices, and procedures, a policies web page has been established at [www.cis.udel.edu/~carberry/ACL](http://www.cis.udel.edu/~carberry/ACL). In addition, the following efforts are underway:

1. Conference Handbook: The conference handbook is being updated to include suggestions and practices from ACL-02. It will also include a link to software for electronic submission that is being used at this year's conference.
2. ACL officer duties: To help new ACL officers benefit from the insight of former officers, a template is being constructed to help each officer compile a list of his or her duties, along with strategic dates, standard procedures, and recommendations.

We are investigating how best to facilitate the dissemination of information through the ACL web site, electronic conference registration, and electronic elections. Results will be implemented in fall, 2002.

### 1.3 Treasurer's Report for 2002

Presented at ACL 2002

Kathleen F. McCoy

-----  
Account Balances

Converted to base currency: US Dollar

As of 6/29/02

ASSETS	TOTAL USD
Morgan Stanley Checking (old DWR)	\$391,121.96
Swiss Bonds	\$129,483.05

First Union CAP Account	\$105,824.88
Morgan Stanley – Walker	\$30,731.37
Swiss Checking	\$19,025.42
French Bank	\$9,951.98
All Others	\$124.50
<b>TOTAL ASSETS</b>	<b>\$686,263.16</b>

-----  
Income vs. Spending -- 1/1/01-6/29/02

Converted to base currency: US Dollar

1/1/01 Through 6/29/02

Category	1/1/01 - 12/31/01	1/1/02 - 6/29/02	Total
<b>INCOME CATEGORIES</b>			
NAACL-01	247,210.00		247,210.00
ACL-02		236,801.00	236,801.00
Membership	56,559.32	55,892.00	112,451.32
HLT-02		84,006.88	84,006.88
ACL-2001	34,872.72	36,485.00	71,357.72
Publications Income	49,354.56	20,132.34	69,486.90
ACL Offices	13,624.16	19,619.70	33,243.86
INLG-02		29,235.00	29,235.00
SigDIAL 2001	10,461.00	133.00	10,594.00
Income – Unassigned	7,266.56	0.00	7,266.56
ACL Anthology Income		5,300.00	5,300.00
Investment Income	2,501.69	366.40	2,868.09
Funds	1,581.00	1,052.00	2,633.00
ACL-99	1,822.00		1,822.00
Bank Error Fix		675.00	675.00
ACL-00	640.00		640.00
Expense Reimbursement		506.50	506.50
<b>TOTAL INCOME CATEGORIES</b>	<b>425,893.01</b>	<b>490,204.82</b>	<b>916,097.83</b>

EXPENSE CATEGORIES			
NAACL-2001	53,066.30	97,032.66	150,098.96
CL Journal	9,038.34	100,405.07	109,443.41
ACL-2002	603.20	84,262.53	84,865.73
HLT-2002	54.04	84,609.84	84,663.88
ACL Office	48,657.13	25,835.42	74,492.55
ACL/EACL-2001	38,300.53	25,900.00	64,200.53
ACL-2000	23,101.36	27,769.38	50,870.74
Publication	24,636.20	13,669.92	38,306.12
Anthology		21,836.56	21,836.56
Memberships	12,538.79	8,181.12	20,719.91
NAACL Administration		20,134.10	20,134.10
INLG-2002		18,388.60	18,388.60
Exec Meeting	3,481.92	4,410.11	7,892.03
Bank Charges	2,732.51	871.38	3,603.89
SigDIAL2001	3,551.64		3,551.64
EACL	3,169.59		3,169.59
ANLP/NAACL-2000	2,905.00		2,905.00
Bank Error	943.25	6.75	950.00
ACL-1999	945.00		945.00
Priscilla Rasmussen	645.76		645.76
Donations		415.00	415.00
SigDial	365.50		365.50
NAACL-Executive	304.00		304.00
SigLex	96.35		196.35
European Charges	2.34		2.34
TOTAL EXPENSE CATEGORIES	229,238.75	533,728.44	762,967.19
GRAND TOTAL	196,654.26	(43,523.62)	153,130.64

-----  
Net Worth over Time

Converted to base currency: US Dollar

1/1/98 Through 6/25/02

	1/1/98 - 12/31/98	1/1/99 - 12/31/99	1/1/00 - 12/31/00	1/1/01 - 12/31/01	1/1/02 - 6/25/02
NET WORTH					
	427,759.83	395,698.78	534,132.52	732,186.78	687,663.16

## 1.4 ACL European Chapter Chair Report for 2002

Donia Scott, Chair

John Carroll, Secretary

### 1. Constitution

The ACL executive ratified the new EACL constitution at their meeting in Toulouse, July 2002.

### 2. EACL Chair-elect

The new constitution introduces a new EACL post: Chair-elect. The EACL Nominating Committee proposed Claire Gardent. The EACL Secretary emailed out a call to members for any further nominations; none were received. Claire Gardent was therefore elected unopposed, as Chair-elect 2002. She will become Chair in January 2003, for an initial term of 2 years.

### 3. EACL'03

Three bids to hold EACL'03 were received, from Budapest, Heidelberg and Trento. The EACL Board at its meeting in Toulouse, July 2002, shortlisted Budapest and Trento, and asked for further specific details in revised bids. When these were received, the Board discussed them and chose Budapest. Major considerations were: encouragement of East-West links in Europe, and the fact that Trento is hosting a major CL event this year (ESSLLI'02). Claire Gardent (as Chair-elect ) will be in overall charge of conference arrangements, assisted by Philippe Blache (Board member responsible for conferences), both reporting to the rest of the Board. The local arrangements chair is Ferenc Kiefer.

The conference will take place 12-17 April 2003, at the Hotel Agro situated in the Buda hills. The Board is seeking to widen participation, and there will be dedicated tracks for project notes and demos; the Board is also encouraging national CL groups in Europe to organize events in conjunction with the conference. It is hoped that there will also be official participation from EU projects.

Conference chairs have been appointed and have held one meeting, at LREC in Las Palmas. The list of

chairs is:

Programme Co-Chairs: Ann Copestake (UK), Jan Hajic (Czech Republic)

Demos and Poster Chair: Alberto Lavelli (Italy)

Tutorial Chair: Dan Cristea (Romania)

Publication Chair: Patrick Paroubek (France)

Workshop Chair: Steven Krauwer (The Netherlands)

Student Workshop Chair: EACL Student Board (M. Gabsdil, J. Hockenmaier, J. Herring)

Planning for the conference is going well, and the first call for papers will be sent out soon. The conference website is <http://www.conferences.hu/EACL03/>.

#### 4. Newsletter

There has been much positive feedback about the first EACL newsletter.

The second EACL newsletter was sent out to members 18 January 2002, and the third is in preparation and will be sent out to members in early July.

#### 5. Sponsorships

In 2002, EACL will sponsor the ESSLLI foundational course on Shallow

Text Processing to be given by Chris Brew, with a contribution of 750 euro. In return for sponsorship, members of EACL will be eligible for a reduced registration fee.

EACL will also sponsor the computational lexicography training workshop LEXICOM, to be held at the University of Brighton, July 2002. The contribution will be 750 euro, to be administered by LEXICOM to support the participation of a student from Eastern Europe.

#### 6. European Masters in Language and Speech

EACL (jointly with ISCA) continues to endorse EMasters awards to successful students. Walter Daelemans, the EMasters board member appointed by EACL reports that the initiative seems to be thriving. The summer school in Leuven this summer (8-12 July) will attract 50 students and 10 staff. The summer school structure in Brno, which included student presentations, will be kept in future schools. The 14 partners in the EMasters have now graduated 24 students, with 52 students ongoing. The Universities of Muenster and Muenchen have expressed an interest in becoming partners in the EMasters. The EMasters website (<http://www.cstr.ed.ac.uk/Euromasters/>) is informative and

well-designed. There are plans to try to solve the problem of lack of financial support by having departments pay for their own costs and in addition apply for an Erasmus intensive programme as a group. Finally, Gerrit Bloothoofst has announced that he will retire as coordinator of the EMasters; a new coordinator will probably be appointed by the EMasters board at the next meeting in Leuven on July 13.

#### 7. New EACL Website

The EACL Student Board have devised an attractive and informative new website, which went online in May 2002, at <http://www.eacl.org>. The site contains links to EACL news, the main ACL site, conferences, CL research and educational resources, and archival material. The EACL Board congratulate the students on a very professional job.

#### 8. New officers

The biennial cycle for electing new EACL officers is coming around again, and the EACL Nominating Committee will shortly be preparing the slate for the term 2003-5.

### 1.5 ACL European Chapter Treasurer Report for 2002

ACL EUROPEAN ACCOUNT EUR/CHF 0.679521

SWISS SAVINGS ACCOUNT AS AT 31/12/2001 29.01.02

Date		CHF	Euro
01/01/2001		32,232.03	21,902.34
<b>INCOME</b>			
dues	2794.88		
Pubs	2526.50		
Mail	254.20		
Late	87.81		
Interest	108.40		
total income	5771.79	5771.79	3922.053
<b>EXPENSES</b>			
--EACL EXPENSES--			
eacl exec trav	-4021.10		
eacl.org	-291.85		

Subtotal	-4312.95	-4312.95	-2930.74
--ACL EXPENSES--			
Toulouse	-3680.25		
Reimbursemen	-893.15		
Siglex.org	-290.65		
bank charges	-663.50		
Subtotal	-5527.55	-5527.55	-3756.86
total expenses		-9840.50	
31/12/2001		28,163.32	19,137.57

Mike Rosner 29.01.2002

=====

ACL EUROPEAN ACCOUNT

FRENCH CURRENT ACCOUNT AS AT 31.12.2001

Date		FF	Euro
31/12/2000		60,617.81	9,241.13
INCOME			
Dues	5,251.02		
Interest	1,721.54		
total income	6,972.56	6,972.56	1,062.96
EXPENSES			
EACL exec meeting	-1,694.77		
Charges	-15.48		
total expenses	-1,710.25	-1,710.25	-260.73
31/12/2001		65,880.12	10,043.36

Mike Rosner 26.01.2002

=====

ACL EUROPEAN CHAPTER EU/USD 1.16332

SHADOW ACCOUNT

EUR

01.01.2000		0.00
INCOME		

ACL (= USD 5000)	5,816.60	
EACL99 (may include acl subscriptions)	3,354.00	
total income 2000		9,170.60
EXPENSES		
sponsorship NATO conference	-1,945.00	
sponsorship ESSLI	-2,006.00	
total expenses 2000		-3,951.00
01.01.2001		5,219.86
INCOME nil		
EXPENSES		
Exec Meeting Brighton	-2,993.07	
eacl.org	-291.85	
total expenses 2001		-3,284.92
01.01.2002		1,934.94

=====

ACL European Account EU/ FF EU/ CHF

Statement of Assets 31.12.2001

Date	A - CHF	A - EUR	B - FF	B - EUR	C - EUR	EUR/FF	EUR/CHF
31.12.95	164,340	102,427	54,683				
31.12.96	183,317	114,255	59,740				
31.12.97	189,427	118,063	61,850				
31.12.98	201,725	125,728	58,987				
31.12.99	198,877	123,952	66,388	10,113	134,065	0.152332	0.62326
31.12.00	217,923	143,238	60,617	9,241	152,479	0.152451	0.657287
31.12.01	219,837	149,384	65,880	10,043	159,427	0.152449	0.679521

A = ALL SWISS = Savings + Bond Fund + any outstanding interest

B = ALL FRENCH

C = TOTAL ASSETS IN EUR = A + B

Mike Rosner, Malta, Jan 2002

## 1.6 ACL North American Chapter Chair Report for 2002

Report from NAACL

Diane Litman, Chair

<http://www.naacl.org>

### 1. Election

The NAACL election in the fall of 2001 (held electronically) produced the following results:

Chair: Diane Litman (renewed)

Secretary: Claire Cardie (renewed)

Treasurer: Dragomir Radev

New Chapter Board members: Lynette Hirschman, Owen Rambow

All Chapter Board members are elected for 2 year terms. This fall there will be an election for the two remaining Chapter Board positions (currently held by Graeme Hirst and Daniel Jurafsky). Any North American member of ACL can run for these positions, and can be nominated by either the nominating committee (which will be chaired by Lori Levin), or directly by a NAACL member. For all other positions, the next elections will be in the fall of 2003.

### 2. 2002 Conference: ACL

Since ACL is in North America this year, there is no separate meeting of NAACL. NAACL is co-hosting ACL.

### 3. 2003 Conference: HLT-NAACL 03

For 2003, the NAACL Chapter Board has decided to merge the NAACL conference series with the HLT (Human Language Technology) conference series. The goal of the merger is to combine the best of the two conferences, specifically, the high standard of full paper reviews for NAACL and the broader range of topics of HLT. We feel that NLP researchers and the broader HLT community will both benefit from increased exposure and exchange of ideas. Furthermore, we feel that the division into an "applied" conference and a main conference (presumably more theoretical in outlook) is somewhat artificial, as applications have always been featured in the main conference as well, and as the line between application and more fundamental research is not always that clear in

our discipline.

To plan the conference, an Advisory Board has been formed, whose members include three representatives from NAACL (Graeme Hirst, its chair; Lynette Hirschman; Dragomir Radev), three representatives from US funding agencies (Karen Kukich, NSF; John Prange, ARDA; Charles Wayne, DARPA), a representative from the IR community (Donna Harman, NIST), a representative from the speech community (Salim Roukos, IBM), and a co-chair of HLT2002 (David Yarowsky, JHU). The Advisory Board proceeds by consensus, and the NAACL members of the Advisory Board refer all important decisions for ratification by the entire NAACL Chapter Board. To date, the Advisory Board has appointed Eduard Hovy as General Chair, and Marti Hearst and Mari Ostendorf as PC co-chairs. Needless to say, acceptance of submissions will be, as always, entirely based on merit; the members of the Advisory Board will have no direct influence on the selection of papers to be presented at the conference. It is the intention of the NAACL Chapter Board that the merged conference will include all the research areas of previous NAACL, ACL, and HLT conferences; further details will be elaborated by the PC.

The conference will take place in late May / early June, 2003, in Edmonton, Alberta, Canada. Dekang Lin has kindly agreed to serve as Local Arrangements Chair. The conference site was selected via an open bidding process.

We are excited about HLT-NAACL 03 and we think it will provide a good opportunity for us to present our work to a larger audience. However, we recognize that this is an experiment. We are now committed only for 2003, after which time we will assess the success of the merged conference.

See <http://www.hlt-naacl03.org> for further updates.

#### 4. Support for Summer Schools

Following some suggestions from members regarding a desire for educational initiatives, the Chapter Board has decided to cooperate with Johns Hopkins's CSLP Summer School on Human Language Technology to fund 5-10 undergraduate and graduate students to attend the 10 day course that takes place before the workshop. The course covers NLP

and speech, with much hands-on work. As this year the 10-day course overlaps with ACL, the students will also attend ACL (the students will miss three days of course work, mainly related to speech and IR).

This year 9 students were selected, and were given scholarships ranging from \$475 - \$1225 (which covered JHU and ACL registration, as well as funds for transportation, lodging, and food, depending on need). The total cost to NAACL will be \$9575 dollars, paid for by part of the surplus from NAACL-01. This experiment will be repeated two more times, at which time we will evaluate the experience.

More information on the course:

<http://www.clsp.jhu.edu/ws2002/calendar/preliminary.shtml>,

and on the scholarships:

<http://www.cs.cornell.edu/home/cardie/naacl/summer-school/summer-school-home.html>

We are also considering expanding our support to include cooperation with other relevant North American summer schools. In particular, ACL has asked us to take responsibility for any future arrangements with the LSA Linguistics Institute (which before our chapter came into existence, was handled by the ACL executive committee).

#### 5. Outreach to Latin America

In the summer of 2001 the NAACL executive committee asked Ted Pedersen to act as a liaison to members of the CL/NLP research community in Latin America. His objectives were to identify researchers and organizing efforts in Latin America, and to solicit opinions as to what would increase participation in NAACL in Latin America. Ted has prepared a summary of his findings based on contact with approximately 100 researchers. The NAACL chapter board will discuss Ted's report at our 2002 executive committee meeting (during ACL).

#### 6. Newsletter

We have started a newsletter series, which is sent electronically to all NAACL members. The first issue went out in September 2001, while the second issue was sent in May 2002.

## **7. Executive Committee Meeting**

The NAACL Board will meet for its yearly meeting on Saturday, July 6, 2002, at the ACL conference in Philadelphia. Minutes will be posted on the chapter website.

### **1.7 ACL North American Chapter Treasurer Report for 2002**

REPORT by Dragomir Radev

My term started in January 2002 so there is not too much to report.

Breaking news: as of today (June 28!), the NAACL exec finally has check writing privileges!

#### 1. NAACL-01 conference finances.

Alon Lavie had done an excellent job keeping the conference books so even that task was quite easy for me to approve his numbers. After consultations with Priscilla Rasmussen and Kathy McCoy, the following are the final numbers for NAACL-2001:

Income: \$202,927.25

Expenses: \$159,945.34

Surplus: \$42,981.91 of which \$33,104.27 belong to NAACL. The rest goes to the different SIGs that had workshops at the 2001 meeting.

#### 2. ANLP/NAACL-00.

The books have been successfully closed. The NAACL portion of the surplus amounts to \$3,275 which, added to the NAACL-01 surplus, gives us a total of \$36,379.27.

#### 3. NAACL shadow account.

The NAACL exec, in consultation with the ACL exec, decided to create a shadow account for NAACL which will include all surplus from ANLP/NAACL-00 and NAACL-01 as well as future NAACL events. The account will be used for the following purposes: JHU summer school, advances for NAACL-HLT-02, as well as for current expenses.

#### 4. JHU summer school.

The NAACL exec decided and the treasurer approved \$10K from the NAACL budget to be used to fund students to attend the newly revamped JHU summer school in Baltimore and to also register for ACL-02. Ten students out of 51 applicants were funded, with amounts ranging from \$475 to \$1,225.

Owen Rambow's report will describe the summer school in more detail.

5. 2002 NAACL conference.

There is no separate NAACL conference in 2002 (ACL-02 subsumes the NAACL meeting) so there are no numbers to report. However, after the SIGs get paid off, the surplus of the conference will be shared between ACL and NAACL.

6. NAACL-HLT-03.

A decision was made to create an account with \$5,000 in it for Ed Hovy, Marti Hearst, and Mari Ostendorf for use (after NAACL exec approval) for expenses related to the 2003 conference. As of this time, the account has not been used.

7. Getting funds from NAACL.

If you are organizing an NAACL-related event or if you think that you may incur NAACL-refundable expenses, please send mail to the treasurer (radev@umich.edu) first. It is highly unlikely that you will get reimbursed if you don't get an email confirmation from me before you incur the expenses.

## **1.8 Special Liason with Asia Report for 2002**

J.Tsujii (University of Tokyo, Japan)

(1) The AFNLP has started to move, while the movement is not so quick as I expected. We had a meeting when NLPRS (NLP Pacific Rim Symposium) took place in November 2001, in Japan. At the meeting, we agreed that we should take more systematic approaches to the federation i.e. the qualification of memberships, committee structures, their written charters, procedures of appointment of chairman, etc.

(2) We, therefore, established four tasks groups to prepare a draft of constitution of the federation (membership qualification, process of decision making, etc.) and drafts of the charters of the three major committees of the Federation, that is, the conference coordination committee, the resource committee and the publicity committee. Each committee will have prepared the drafts by the mid August, and their drafts will be discussed at the meeting we have at the Coling, Taipei. In particular, the conference coordination committee report will recommend to have a flagship conference of the federation and to re-organize several conferences in the region.

(3) In general, the atmosphere of cooperation among Asian colleagues emerges and become stronger.

The Thai group organizes a workshop on Asian MT systems and Multilingual NLP in May, 2002, supported by the UN. Participants were from countries like Vietnam, Myanmar, India, Indonesia and Malaysia as well as Japan, and Korea. They also took initiatives in organizing a workshop at Coling. The members of the resource committee also organize another workshop at Coling. A conference in Asia, IRAL (Information Retrieval with Asian Languages) is planning to organize their conference at ACL2003 at Sapporo as a workshop.

(4) The first issue on Transaction of Asian Language Information Processing (TALIP) by ACM finally appeared. This journal is to be one of the journals the Federation supports.

(5) The website established by the publicity committee become more informative than before. The resource committee is gathering catalogue information of linguistic resources in Asia and the catalogues will be made public through the website. A member association of AAMT also starts their multilingual website (English, Japanese, Korean, Chinese, and Thai) and will be linked with our Website.

(6) The situation in China is still unclear, but a professor of Beijing University may play the role of representative of China together with a representative from Hong Kong. The Micro Soft research lab in Beijing is also actively involved in research in this field and some of them have already been involved in committee activities. However, China is so large that we still cannot have a whole image of what is going on there. Bi-lateral meeting on NLP between Japan and China started last year and we will have another one October this year at Shanghai. We will have more information on activities in the Shanghai area.

## **1.9 Nominating Committee Report for 2002**

The 2002 Nominating Committee of the Association for Computational Linguistics, consisting of Phil Cohen, Wolfgang Wahlster, and Ed Hovy, is delighted to present the following slate of candidates for consideration by the membership of ACL:

VP-elect: Robert Dale

Martha Palmer

Treasurer: Kathy McCoy

ACL Exec Board: Norbert Reithinger

Annie Zaenen

## 2.1 General Conference Chair's Report for 2002

Pierre Isabelle

I was officially appointed as General Chair of ACL-02 on 2 October 2001, that is, about 9 months ahead. The Handbook timetable suggests GC's should be appointed about 16 months in advance; as we will see, my experience supports the idea of an earlier appointment.

### 1) Tasks

#### A- Appointing Chairs

- Local Chair: Martha Palmer had already been appointed. She set up a local committee comprised of Trisha Yannuzzi, Laurel Sweeney, Gilbert Kowie, Craig Martell, Yuan Ding, Paul Howard.
- Program Co-Chairs: Dekang Lin and Eugene Charniak had already been appointed. Worked on setting up a committee made up of 9 area chairs; discussions with exec took more than a month. Committee members: Daniel Marcu, Stephan Busemann, Keh-Yih Su, Bonnie Dorr, Steve Abney, Jan Hajic, Mark Steedman, John Carroll, and Ellen Riloff.
- Workshops Chair: I appointed Walter Daelemans, Oct. 23. Walter did not feel the need to set up a reviewing committee for workshop proposals.
- Tutorials Chair: Appointed Hwee Tou Ng, Oct. 22. Hwee Tou then enrolled Pamela Jordan Jakub Zavrel as proposal reviewers.
- Student Workshop: Appointed Susan McRoy as Faculty Advisor, Oct. 25. In consultation with her, I appointed Gideon Mann and Alexander Koller as Co-Chairs (Dec. 3). Susan and I advised them on how to set up a "balanced" PC. The PC was formed Jan. 30
- Webmaster: Appointed Aur\_.A\_Nilien Max, Nov. 21
- Demo Chair: Appointed Ronnie Smith, Nov. 29
- Sponsorships/Exhibits Chair: Appointed Debbie Dahl, Dec. 5.
- Publications Chair: Appointed Dragomir Radev, Feb. 20.

#### B- Help other Chairs work out basic calendars, prepare and send out their calls:

- main conference, preliminary CFP sent out Oct 13
- call for workshop proposals: Nov. 6

- call for tutorial proposals: Nov. 6
- call for student papers: Dec. 13
- call for demos: Jan. 24
- call for sponsorships: Feb. 1
- call for exhibits: Feb. 1

#### C- Electronic Submissions

Early on, I insisted that we needed to accept electronic submissions. We cannot assume that Program Chairs will be prepared to handle that chore, so I spent some time searching for someone to do it. But I was lucky: soon enough Dekang Lin decided to volunteer for the job; he set up and managed a very convenient Web-based interface that was used not only for submissions but also for final versions (CD-ROM production).

Dekang took up my suggestion of hiring a student to help him at submission time.

#### D- Publications

Given the extremely tight schedule, publishing all conference papers (main conference, student workshop, conference workshops) in both hardcopy and CD-ROM is a formidable challenge. We were lucky to have such a resourceful and dedicated Chair as Drago organize and coordinate the work.

Drago took up my suggestion of hiring a student to help him at submission time, and the quantity of work to be done by the student, especially for producing CD-ROM, exceeded our estimates.

A large number of issues arose during the critical weeks between reception of final submissions and shipping to Omnipress, generating several hundreds of e-mails.

#### E- Budget Issues

The ACL Secretary-Treasurer did most of the work, but I was involved in many discussions with her and other actors, especially in the period where we had to set registration fees. Given Kathy's experience, I found that aspect of the GC's role very easy.

#### F- Consult ACL/NAACL execs

For many important decisions (selecting Chairs, selecting PC's, selecting tutorials and workshop, I managed the process of consulting the execs and to interpret and implement their advice.

## II- Problems and recommendations

A- Schedule. Starting as late as we did made our life unnecessarily difficult. The GC/LC/PC really need to be appointed at least 12 months in advance. See if the exec can implement processes that will help guarantee that this is done.

B- Electronic submissions. In my opinion the ACL should have a policy of accepting electronic submissions for conferences (and maybe for the journal as well). And policy or not, a section on how to handle that should be added to the Handbook, along with pointers to useful pieces of software. Since electronic submissions processes now tend to be based on Web forms, it would make sense to place Web submissions under the responsibility of the conference Webmaster.

C- Electronic payment. We had some difficulty identifying a good way to handle secure electronic payments. Fortunately, Kathy McCoy managed to get the help of Greg Silber at Univ. of Delaware, who was able to arrange for secure payments. The ACL office should have a permanent facility for doing this.

D- Publications Chair. The task of the Publications Chair is extremely demanding, especially in view of the fact that most of the work is of necessity done under a very tight schedule. Some precious time was wasted in pondering who should be doing what. The duties of the Publications Chair should be described in detail in the Handbook, including recommendations on exactly how to share the work with the Program Chairs, the Student Chairs, the Workshops Chair and the Workshops organizer.

## III- Conclusions and acknowledgements

Serving as GC really gives one the feeling of being involved in everything. It can be a lot of fun, but if one does not have the right team for the job, I am sure it can quickly turn into a nightmare.

In my case, it was a great experience! I was truly privileged to be working with such an incredibly competent, resourceful and dedicated project team! Thanks to the ACL exec for giving me their trust,

as well as their constant support and encouragement. Thanks to all those who contributed to ACL-02.

## **2.2 Program Co-Chairs Report**

by Eugene Charniak

Initial actions of the program committee were routine with one exception, we decided to accept only electronic submissions, a first for ACL. This worked well due in large part to the abilities of one of the program chairs, Dekang Lin, who handled this responsibility. As in previous years we required electronic notification. In retrospect the usefulness of electronic notification is not completely clear if one is also doing electronic submission. One seemingly obvious benefit, allowing the program committee to start divvying up papers before they were submitted did not work well because of the large difference between the number of notifications and submissions. We received approximately 340 notifications of submission but only 258 actual submissions.

Of the 258 submissions one was withdrawn during review, and one was rejected without review (obviously not relevant to the conference). The submissions came from 22 countries and by percentages came out as follows: Americas 130, Asia-Pacific 53, and Europe-Africa 75.

Following the recommendation of last years program chairs, while the reviewing was blind, the area chairs were given the names of the authors, to better allow them to detect conflict of interest.

The program committee meeting was held March 30 and 31 in Toronto at the University, where Graeme Hirst made the local arrangements. (We did not select the locations of the program chairs because of flight availability.) The meeting was attended by the two program chairs and 8 of the 9 area chairs. To concentrate discussion on those papers that were most in doubt we organized the discussion as follows: The papers were sorted by the reviewer's average "overall" scores. We then went down the list starting from the top. Any paper with a score of 4.0 or higher (out of a possible 5) was accepted unless the paper's area chair had reservations. There was one such case, and if memory serves, it ended up rejected. Any paper with a score lower than 3.0 was rejected, again unless the area chair had reservations. (Ultimately all such papers were rejected.) All of the papers in between were discussed in detail.

Of the 258 papers we accepted 66.

We had many cases of dual submissions. In all cases except one the authors chose to present at ACL. (The one exception will appear at this years Cognitive Science conference.) Thus there are 65 papers to

be presented at the main session of the conference. The acceptance rate overall was 25%. By geography this broke down: Americas 40 (30%), Asia-Pacific 9 (17%) and Europe-Africa 17 (22). By area this breaks down as follows:

Area	Accepted	Rejected
Discourse and Dialogue	5	21
Generation and Multi-Modality	7	17
Machine Translation and Multilinguality	6	22
Lexicon and Semantics	7	19
Speech, Language Modeling and Statistical Methods	9	22
Word Segmentation, Shallow Parsing, Chunking and Tagging	9	23
Syntax, Grammars, Morphology and Phonology	5	16
Parsing	10	20
NLP Applications	8	30

As is to be expected many papers could fit into several areas, and there was some re-shuffling to even up area loads. Even so, it is clear that NLP Applications received the most papers and "Syntax..."the fewest.

Following last year we decided to have another "best paper" award. The program committee selected 9 papers to be considered. (These were the only papers to receive two nominations for the award from the reviewers.) As two of the papers were by area chairs, the remaining seven area chairs and two program chairs agreed to read 9 papers within two weeks and at which point we would reduce the group to a more discussable four. In practice it was quite easy to reduce the group to four. These four were all very strong, but after some discussion we finally decided that there was one paper that stood out, and it was decided to give this paper the award and not to share it (as was done last year).

Best paper selection was the last content-based action of the committee. What followed was all administrative, including production of the proceedings, finding sessions chairs, and producing this report.

### **2.3 Local Arrangements Report for 2002**

Preliminary Local Arrangements Report: (Budget at Exex meeting)

Registration figures as of June 25, 2002.

MAIN CONFERENCE: 615

#### **WORKSHOPS:**

EMNLP: 202

Effective Tools and Methodologies: 38

Discourse and Dialogue SIGDIAL: 86

Speech to Speech Translation: 41

Computational Approaches to Semitic Languages: 18

Natural Language Processing in the Biomedical Domain: 58

Unsupervised Lexical Acquisition: 84

Word Sense Disambiguation: 56

Morphological and Phonological Learning: 44

Natural Language Processing and Visualization: 13

Text Summarization: 70

#### **TUTORIALS:**

Machine Learning for Text Classification: 51

Toward Spontaneous Speech Recognition: 13

Kernel Methods and Support Vector Machines: 45

NLP Approaches to Reference: 38

BANQUET: 306

STUDENT LUNCH: 142

Student Volunteers: 61

Special Events for ACL-02

Sat, July 6

NAACL Dinner, Inn at Penn, Chancellor Bd. Rm., 3:45-11 p.m., 10 (+3)

Contact: Diane Litman

EMNLP Reception, Hall of Flags, 6:00 8:00, 200 guests

Contact: Jan Hajic

Sunday, July 7

ACL Executive Board Meeting, Golkin (223) in Houston, 8:30 - 6:00 15/20

Contact: Sandee Carberry

RECEPTION, AT Penn Museum in Chinese Rotunda, 6:30-9 p.m., 450 guests

3 food stations & 2 drinks, Entertainment: Early Music Group

Contact: Martha Palmer

Monday, July 8

Comp. Ling. Ed. Bd. Lunch, Houston Hall, 12:30-2:00

Contact: S. Carberry

EXEC Dinner, Inn at Penn, 6:30-10:00, 50

Contact: S. Carberry

Tuesday, July 9

Banquet at 30th St. Station, North Waiting Hall, 6:30-11:00, 400 guests

Southern Deluxe Buffet, 2 drinks, Band: City Rhythm,

Contact: Martha Palmer

Wednesday, July 10

ACL Conf. Comm.Mtg. Breakfast,, Inn at Penn, 8-10 a.m., #26

Sontact: S. Carberry

Student Luncheon, Hall of Flags, 12:30, pizza, 150

Contact: Gideon [gsm@cs.jhu.edu](mailto:gsm@cs.jhu.edu)

Thursday, July 11,

SIGDIAL Reception, Logan Terrace & Lobby, 6:30-8:30, 90-100

## **2.4 Tutorial Chairs Report for 2002**

ACL-02 Tutorials

Hwee Tou Ng

I prepared the call for tutorial proposals and the first call was publicly broadcast on November 7, 2001. The content of the call for proposals follows closely that of the previous years. The remuneration rate for each tutorial presenter is the same as last year: US\$500 per session plus \$25 per registrant in the range 21-50 plus \$15 per registrant in excess of 50. The various deadlines for submission and evaluation also mirror closely those of the previous year.

In mid December 2001, I enlisted the help of two tutorial reviewers, Dr. Pamela Jordan of the University of Pittsburgh, USA and Dr. Jakub Zavrel of Textkernel, the Netherlands.

By the submission deadline on January 11, 2002, I have received an enthusiastic response of 11 proposals. Many of them are of high quality, and this makes the selection process somewhat challenging since we only have slots for 4 half-day tutorials to be held in one day (i.e., acceptance rate of only 36%). The two reviewers and I deliberated on the selection individually, and each came up with an independent assessment and ranking of the proposals. We discussed among ourselves and resolved our differences to arrive at 4 selected tutorials. We then sought the views of the ACL-02 committee, the ACL executive committee, and the NAACL executive committee. After some discussions via email, the final accepted tutorials for ACL-02 are (in alphabetical order of the speakers' last name):

- (1) Sadaoki Furui, Toward Spontaneous Speech Recognition and Understanding
- (2) David D. Lewis, Machine Learning for Text Classification Applications
- (3) Jean-Michel Renders, Kernel Methods and Support Vector Machines - Application to NLP Problems
- (4) Michael Strube, NLP-Approaches to Reference Resolution

The final choices were made taking into account the appropriateness and timeliness of the topics, the speakers' reputation and knowledge of the topics, with a view to ensuring a balanced representation. Tutorial (1) and (2) will be held in the morning, with the other two held in the afternoon. We took care

to avoid scheduling (2) and (3) at the same time, since they may tend to attract the same corpus-based/statistical NLP audience.

I sent out the acceptance and rejection notices on January 25, 2002. By the end of March, tutorial descriptions were put up on the conference web site. The tutorial notes from all 4 speakers were also received largely on time, by mid May 2002, and they take up a total of 522 pages! This gives an idea of the amount of materials that will be covered in the tutorials.

The tutorial registration fee is set at US\$115 and US\$145 per tutorial session for early and late registration respectively. Each tutorial attendee will receive one hardcopy of the tutorial notes, which include the tutorial slides and a bibliography. The numbers of attendees who pre-registered are: Furui 10, Lewis 47, Renders 35, and Strube 35. The speech tutorial seems to attract fewer attendees, but the other numbers are encouraging, and these numbers are comparable to the attendance figures of the last ACL-01. Overall, I look forward to a successful tutorial program on Sunday, July 7, 2002.

## **2.5 PUBLICATIONS CHAIR Report for 2002**

REPORT by Dragomir Radev

It is a pleasure to be publications chair when everyone else involved in the process is friendly, timely, and reasonable. This list includes the editors of the individual proceedings, the conference chair, the printers at OmniPress, the local organizers, the CD-ROM editor, as well as Priscilla Rasmussen, David Yarowsky, and Jason Eisner. For the record, the most timely person to answer my emails was the general chair, Pierre Isabelle, which is amazing given the tons of other things that he had to do for this conference! Early on, I created a mail alias at aclweb.org that includes all people involved in the publications cycle. This mailing list proved to be quite valuable.

From a logistical point of view, a problem that haunted us until the very end and which could have been avoided was due to a violation of an important principle of data replication. To create the CD-ROM we needed a copy of the metadata for each paper (in Yarowsky "db" format). It turned out that these "db" files were maintained by each individual proceeding chair but also automatically generated by a program at Alberta using data typed by the authors themselves. As a result, the two versions were sometimes quite different, e.g., with authors in different orders, incorrect titles, varying numbers of pages, etc. For the future, I would like to recommend that all metadata for the papers be

created in one central location (e.g., on the conference Web site) and then used by the individual chairs to prepare their proceedings.

The style files for the conference and proceedings were adapted from previous years' versions by Dekang Lin. In general, authors were quite good about following the guidelines for the papers.

The set of tools developed by David Yarowsky for the CD-ROM and the individual proceedings was used with very few modifications. This is the third year that this software has shown to be invaluable!

The CD-ROM was made possible only after a number of sleepless nights spent by Jahna Otterbacher in order to get every single page, every singly link and figure in order. Fantastic job, Jahna! For the record: Jahna ended up spending 93 hours on the CD-ROM. In the future, I recommend that the position of CD-ROM chair should be created as the amount of work involved is way too much for the publications chair alone. I ended up spending a similar amount of time on the rest of the duties of publications chair.

One of the innovations in this year's CD-ROM is adding reference information for each paper in the margins of the PDF files. Thanks, Jason Eisner, for the idea and for helping make it possible.

Some additional suggestions for the future:

- there should be a uniform page order for all front matter in a workshop proceedings
- there should be a standard taxonomy of ACL topics that can be used to classify each submitted papers
- a standard for archival metadata should be agreed upon and followed. This standard should also address the inclusion of fonts in the soft copies of the papers
- the position of publications chair should be for two years. I learned a lot of useful things this year while doing this job which I should have been aware of earlier.

## **2.6 Student Session Chairs Report for 2002**

ACL-02 Student Research Workshop

Alexander Koller, Gideon Mann

### 1. Program Committee

The co-chairs of the ACL-02 Student Research Workshop, Alexander Koller (Saarland University) and Gideon Mann (Johns Hopkins University) were nominated by this year's general chair, Pierre Isabelle,

and approved by the ACL Executive Committee. Susan McRoy was appointed by the ACL Executive Committee as Faculty Advisor. The program committee was formed by the co-chairs and approved by the ACL Executive Committee. The final program committee consisted of 25 student members and 22 non-student members. Of the 47 reviewers, 19 were from North America, 19 from Europe, 8 from Asia, and 1 from South America.

## 2. Submission and Acceptance

We received 29 submissions to the Student Research Workshop. The submission length was set to six (6) pages using the ACL style files. Because of the unusually low number of submissions, and a number of reviewers which we had deliberately selected rather too high, we were able to get four reviews for most papers. 20 papers were assigned four reviewers each, 3 papers were assigned five reviewers, and 6 papers were assigned three reviewers. Each submission was reviewed by at least one student and one non-student reviewer. Reviews were generally done by e-mail.

We accepted 12 of the papers. One of the accepted papers was withdrawn because it had been accepted at another conference, and we accepted an alternative paper instead. Some statistics for the submissions and acceptance are shown below (Tables 1-3).

## 3. Presentation Format

The papers will be presented in half-hour slots (organized into three parallel sessions) allocated from the main conference. Continuing last year's format, for each accepted paper we have invited two senior researchers to act as panelists to give detailed feedback to the student author(s). Each slot will thus start with author presentation (18mins), followed by panelists feedback/discussion (7mins) and general Q&A (5mins).

## 4. Panelists

Currently, we are in the process of recruiting 2 panelists for each presentation.

## 5. Funding

As in previous years, the faculty advisor obtained funding for student travel from NSF. An application was submitted in early December (the deadline for grants to be funded in time for the ACL in July), using the current call for papers and historical information about the workshop. \$18,000 was requested and awarded.

The student co-chairs accepted funding applications from all students attending ACL who wished to

request support. To avoid conflicts of interest, McRoy reviewed the applications and made recommendations for funding. Of the 12 students attending the workshop, 8 requested funding and all will receive funding which will pay for most of their expenses. Funding levels were determined based on distance travelled to the conference. Because there was money remaining, with the permission of the NSF program director, McRoy also made recommendations regarding support for students not involved with the student research workshop. These were forwarded to the Conference chair and the ACL Treasurer.

## 6. Problems

The co-chairs found the experience from organizing the workshop invaluable and rewarding, and we note that some problems of earlier years have been solved for us -- most notably, scheduling and funding issues. We would like to thank ACL and the general chair, Pierre Isabelle, for offering us this opportunity.

### 6.1. Timing

The co-chairs were selected on December 3, which we found rather late. The late date forced us to prepare the call for papers and set up the WWW infrastructure very quickly, and effectively pushed reviewer selection into the new year. Even so, our call for papers was distributed two weeks after that of the ESSLLI student session, a major competitor for student papers, which may have contributed to the small number of submissions.

### 6.2. Reviewer Selection

It took us all of January to compile a satisfactory list of reviewers. By the time we could send invitations, some potential reviewers had already committed to reviewing for other conferences, and either declined or asked to be assigned only few reviews.

We spent one week of this period building up an initial proposal, and three weeks revising it to achieve a better geographical balance of reviewers. In order to improve efficiency in the future, we advise to make the requirements with respect to balances in the ACL Conference Handbook more precise.

To improve the efficiency further, we suggest that reviewers for the Student Research Workshop should be confirmed by the General Chair, not the ACL Exec. This would make it possible to get early feedback from the final decision maker, which is much harder if the final decision is made by a committee. In addition, this would allow for a more flexible allocation of reviewers. We chose a very high number of reviewers, which turned out to be too many for the submissions we had, because we

found it impractical to enter a committee discussion about additions to the reviewers list.

### 6.3. Student Status

All submissions were accompanied by either a CV or a letter from the student's advisor. This ensured effectively that all submitted papers were really written only by students. However, what precisely constitutes valid proof of studentship, and whether undergraduate students (who don't necessarily have an advisor) are eligible to submit, was not well defined.

For the future we advise to require that each author should send an email stating informally that they are a student and aren't planning to submit their PhD thesis within a year after submission of the paper. This should have the same effect while removing bureaucracy.

### 6.4. Distribution of Accepted Papers

Of the 13 accepted papers (including the reserve paper accepted after another's paper withdrawal), 8 were submitted by students from only four universities: Two each from the Carnegie Mellon University, University of Edinburgh, University of Cambridge, and Saarland University. We made the acceptance decision purely on the grounds of the reviewers' recommendations, but we are not happy with this clustering of successful papers to only a few universities: Those are exactly the places where a student has good access to feedback anyway.

Future acceptance decisions should be monitored for whether there is a trend towards a focus on a small number of universities, and possibly discuss measures for reversing it.

Table 1: Papers by country

Country	Submitted	Accepted
USA	11 (37.9%)	5
UK	6 (20.7%)	4
Germany	3 (10.3%)	2
Japan	3 (10.3%)	2
France	2 (6.9%)	0
Australia	1 (3.4%)	0
Canada	1 (3.4%)	0
Turkey	1 (3.4%)	0
Vietnam	1 (3.4%)	0

"Accepted" counts both the paper that was withdrawn and the paper that was accepted as an alternative.

Table 2: Papers by geographical area

Area	Submitted	Accepted	Acceptance rate
North America	12	5	41.7%
Europe	11	6	54.5%
Asia	5	2	40.0%
Oceania	1	0	0.0%

Table 3: Papers by topics

Topic	Submitted	Accepted
Machine Translation	5 (17.2%)	1
Generation	3 (10.3%)	2
Information Extraction	3 (10.3%)	1
Lexicon	3 (10.3%)	2
Models of Language	3 (10.3%)	0
Corpus-based Language Modeling	2 (6.9%)	0
Morphology	2 (6.9%)	1
Syntax	2 (6.9%)	2
Dialogue	1 (3.4%)	1
Information Retrieval	1 (3.4%)	1
Semantics	1 (3.4%)	1
Speech recognition and synthesis	1 (3.4%)	1
Summarization	1 (3.4%)	0
Message and narrative understanding	1 (3.4%)	0
Pragmatics	0	0
Phonetics and Phonology	0	0
Question Answering	0	0
Multi-lingual processing	0	0
Language in multi-modal systems	0	0

## 2.7 ACL02 Workshop Chair Reports

As a response to the Call for Workshop Proposals (distributed from November 7, 2001), 13 proposals were received, for a total of 16 workshop days. All proposals were of excellent scientific and organisational quality. Ten proposals were accepted for 12 workshop days. Two proposals were unified into one workshop. The workshops chair made the selection, with valuable feedback and advice from the ACL organizers and ACL and NAACL execs. Later on, the EMNLP was added to the list of ACL workshops. Of the organization committees of the accepted workshops only 4 are uni-continental (1 Europe, 3 US), the other 7 are mixed (6 Europe-US, 1 Europe-Asia). Four accepted workshops are organized by ACL SIGs (SIGLEX, SIGDIAL, SIGPHON, SIGDAT). At the request of the proposers, one workshop was scheduled in parallel with the Tutorials Session (on tools for teaching CL).

All CFPs of the workshops were announced around mid-January 2002, after a check by the workshops chair for consistency and uniformity, and distributed on the ACL workshops website (via the webmaster) and by email via Priscilla Rasmussen.

As the registration and programme information was distributed only electronically this year, only electronic workshop descriptions and programs are being made available.

A few workshops received external funding, one workshop was allocated an amount of money for ACL registration waiver of the invited speaker, workshop fees were waived for all invited speakers.

Relatively late appointment of a publications chair and some unfortunate and accidental miscommunication between him and the workshops chair caused some minor confusion about when, how and where to deliver the camera-ready proceedings material (CD-ROM version, paper version, and copyright release statements). The exact procedure and timetable for this should preferably be worked out between Omnipress, the publications chair, the workshops chair, and the conference organizers at a much earlier stage than was the case this year, and communicated to the workshop organizers as soon as possible. The tools developed by David Yarowsky are an excellent way of achieving uniformity and efficiency in this process, and should preferably be distributed to the workshop organizers at an early stage as well.

At the time of writing, no complete information is available yet about number of workshop papers and number of registrations for the different workshops.

Some issues which could be discussed within the ACL-exec are the following:

- Policy on workshop proposals submitted by ACL SIGs. Currently these proposals are treated as any other proposal. However, as these SIGs by definition represent a sizable part of the membership and are concerned with an area judged by ACL to be of importance, it seems strange NOT to accept them. In case of many SIGs submitting a workshop proposal, this would then leave only limited room for

other proposals.

- Conference-like workshops. Some workshops, like EMNLP, DUC and SIGDIAL, have grown to conference-sized events on their own. They could get a special status (e.g. being scheduled not necessarily parallel with the other workshops, and having autonomy over organisational issues like registrations fees, proceedings production, etc.).
- Policy on 2-day workshops. Usually, ACL conferences have limited capacity for workshops. Two-day workshops take up much of the available capacity and should probably be limited in number.
- Restrictions on size of workshop papers. If it is necessary to have such restrictions, this should be included in the policies. Preferably, such a restriction would be a maximum number of pages per workshop day rather than a maximum per paper.

Finally, ACL needs workshops! I was surprised by some of the policies of ACL regarding workshops, e.g. the fact that they are not allowed to set their own registration fees and to use this money to fund invited speakers, the fact that the organizers can't have their names on the cover of the proceedings, restrictions on size of workshop papers, and the fact that some members of the exec prefer a limited number of workshops regardless of capacity. In my opinion, more is better, and if possible, all scientifically and organisationally well-prepared workshops should be accommodated, and maximum flexibility should be allowed to the organizers to make organizing one as attractive as possible. The reason is the following: (i) the ACL main conference programme has room only for a limited number of papers; (ii) most researchers can only travel with accepted papers; (iii) all workshop attendees register for the main conference (and thereby start or continue ACL membership). Hypothesis: without workshops, ACL attendance would be half what it is now, and with more workshops, ACL attendance will increase. Furthermore, workshops, even if limited in scope, popularity or size, serve as an excellent tool to introduce new topics into the mainstream.

Walter Daelemans

Antwerp, June 4, 2002

## **2.8 Exhibits/Sponsorship Chair Report for 2002**

Deborah Dahl

Sponsorships:

We had several levels of sponsorships-- Gold, Silver, Bronze, and Sponsor, as well as special event

sponsorships. We had 4 Silver Sponsors, 5 Bronze sponsors, and 4 special event sponsorships. The lowest, Sponsor, category, for \$500, was new this year, but no one signed up for that category. The total value of sponsorships was \$31,000. Several companies who had been sponsors in the past were unable to sponsor this year's meeting, or had to decrease their level of sponsorship, due to bad economic conditions. The most successful type of contact was an individual email to a personal contact at a potentially sponsoring organization. Several general emails were sent out with very limited results. First, general emails were sent to 80 potential sponsors in February and March. In addition, a general mailing was also sent by Priscilla to the entire ACL membership in April. Only a few sponsorships appeared to be the results of either of these mailings. Most sponsorships came from big companies, such as IBM, Microsoft, and Sun. Only two startups (Teragram and Basis) signed up as sponsors. There were also three sponsorships from non-profits (Cornell Intelligent Information Systems, the Linguistic Data Consortium, and the University of Pennsylvania Computer Science Department).

Points for discussion:

1. We should think about how far we want to go in commercializing the sponsorships. One sponsor suggested that Gold sponsors should have the ability to have their logos put on the attendees badges, and another potential sponsor wanted access to the ACL's mailing list to use for marketing promotions. On the first request, the sponsor decided to become a Silver sponsor and dropped the request to put their logos on the badges. The second request was turned down after discussion with the ACL Exec. We need to discuss how much sponsorship we need and how far we're willing to go in commercializing the conference in order to obtain sponsorships.
2. One Silver sponsor was reluctant to contribute because their last year's sponsorship check had not been cashed, and they finally had to cancel it. We have to make sure that this doesn't happen again.
3. We may want to maintain a database of companies and contacts in natural language related businesses for reference when we're seeking sponsorships for future conferences.
4. Most sponsorships come through personal contacts. If we want to increase the level of sponsorship, it definitely would help to have a team to work on the task of contacting potential sponsors.

Exhibits:

There were 6 exhibitors, most of them book publishers. Emails were sent to 80 potential exhibitors in February and March but most exhibitors found the conference on their own.

Points for discussion:

1. Do we want to increase the number of exhibitors? If we do, what is our objective? That is, do we want to make available information about products of interest to the attendees, do we want to increase revenue for the conference, or do we want to expose the attendees to commercial technologies? Except for books and linguistic data, it's not clear how ACL attendees could be customers for most commercial natural language businesses. Consequently, it's hard to find a motivation for why any company other than a book publisher would be interested in exhibiting at the ACL.

## **2.9 Demos "STATUS REPORT ON SOFTWARE DEMONSTRATIONS PROGRAM" Chair Report for 2002**

June 20, 2002

Chair: Ronnie W. Smith

Completed:

- \* Demo Schedule, 4 sessions, 3 on Tuesday the 9th and 1 on Wednesday the 10th, 19 demos total
- \* One iteration of confirmation with both demonstrators and Local Arrangements in terms of required equipment/logistical support

To be done:

- \* Second iteration of confirmation on logistics support
- \* Securing services of student volunteers to assist during demo sessions
- \* Finalizing procedures for setup and removal of demo materials, especially during the tight transition time between demo sessions 1 and 2 on Tuesday morning (10 minute transition window)
- \* Communication of final instructions to demonstrators
- \* Walk through of demo room on Monday the 8th with Local Arrangements and student volunteers.

## **3.1 Journal "Computational Linguistics" Report for 2002**

57 papers were submitted to Computational Linguistics in 2001. The mean time to first decision was 110 days, down from 190 days for 2000. Of the total submissions, there were 8 in the short paper category with a mean time to first decision of 77 days. The acceptance rate for 2001 was 31.6%.

Disposition of manuscripts by type of first decision:

	2001	2000	1999	1998	1997	1996
Submitted	57	64	47	48	69	57
Accepted	18	15	13	9	21	9
Rejected	12	11	9	7	15	15
Resubmit	22	27	12	4	27	26
No decision yet	0	7	8	28	5	6
Withdrawn	3	3	2	0	1	1
Resub as squib	2	1	3	0	0	2

In 2001 we published one special issue, on Computational Anaphora Resolution, edited by Ruslan Mitkov, Bran Boguraev, and Shalom Lappin. As of June 1, 2002, there are 39 submissions for 2002, 16 of which are for a special issue on the Web as Corpus, edited by Adam Kilgarriff and Greg Grefenstette.

=====

For 2000 papers, the completed disposition of mss was, with the acceptance rate 26.6%.

The chart for disposition of 2000 mss

Submitted	64
Accepted	17
Rejected	13
Resubmission	30
Withdrawn	3
Resub as Squib	1

=====

Distribution of submissions by location of first author for 2001:

<b>ASIA/PACIFIC</b>	
New Zealand	1
P. R. China	1
Japan	1
Hong Kong	1
Total	4
<b>EUROPE</b>	
Germany	3
Norway	1
Ukraine	1

Spain	2
Switzerland	1
Turkey	1
UK	12
Portugal	1
France	2
France	1
Total	25
<b>NORTH AMERICA</b>	
USA	25
Canada	3
Total	28
<b>GRAND TOTAL</b>	57

The new 2002-4 editorial board is:

Elisabeth Andre, Universita"t Augsburg

Eugene Charniak, Brown University

Hitoshi Isahara, Communications Research Laboratory

Maria Lapata, University of Edinburgh

Inderjeet Mani, MITRE and Georgetown University

Kemal Oflazer, Sabanci University

Roni Rosenfeld, Carnegie Mellon University

Marc Swerts, University of Antwerp and Technische Universiteit Eindhoven

CL's new editorial assistant is Kathy Vesonder.

Julia Hirschberg, Editor

### **3.2 Book Reviews Published**

Computational Linguistics

Book Review Editor's report

Graeme Hirst

June 2002

In 2001, we continued to publish about five book reviews plus a couple of brief notices in each issue of Computational Linguistics. This number will certainly drop in 2002-03, as the number of books published in computational linguistics and closely related areas has dropped noticeably over the last year.

Most reviews are published in a timely manner -- that is, within 12 months of receipt of the book. This allows six months for the reviewer (some take less) and five months for journal production.

#### MATERIAL REVIEWED

I am continuing to be fairly strict in deciding if a book is to be reviewed, but try to include all books that are in "core" computational linguistics, as well as a variety of books from adjacent and overlapping disciplines that are likely to be useful in CL. We do not review technical reports, doctoral theses, conference proceedings, or workshop proceedings, except if revised for publication as a book by a recognized publisher.

#### PRODUCTION MATTERS

I am indebted to Nadia Talent for long hours of reading out loud with me to check the galleys.

### **3.3 Squibs and Discussions report for 2001**

Pierre Isabelle, Squibs editor

At the beginning of year 2001, there was only one submission in the pipeline. In the course of year 2001, 12 additional papers were (re-) submitted. At the end of the year, there was only one submission left in the pipeline. Thus, 11 decisions were made over year 2000. The results were as follows:

- \* 4 papers accepted
- \* 4 papers rejected
- \* 3 invitations to revise and resubmit

The mean time taken for these 11 decisions was 83 days.

The geographical distribution of the 12 submissions by location of first author was:

- \* Asia: 1
- \* Europe: 7

\* Israel: 1

\* North America: 3

Thus far, 4 papers only have been (re-) submitted in 2002.

### **3.4 CUP/ACL Book Series "Studies in Natural Language Processing"**

Bran Boguraev and Steven Bird

The goal of Studies in Natural Language Processing is to identify and publish state-of-the-art work on topics of interest to the CL/NLP community. Feedback from sales suggests that interest in the community is maintained at level similar to prior years; levels of sales are in line with the types of scholarly books that have been published recently. The editors welcome proposals for new book projects.

The series is currently migrating between Acquisitions Editors at CUP, from Andrew Beck to Andrew Winnard, and moving the series headquarters from the CUP office in New York to the CUP headquarters in Cambridge. The series will shortly be up for renewal.

#### STATUS OF CURRENT PROJECTS

1. Books published during the last 12 months:

0521631963 HB KIRAZ Computational Nonlinear Morphology 12/17/01

0521780489 HB BUSA/BOUIL The Language of Word Meaning 01/15/01

2. Book projects approved, with manuscripts under review:

\* Wilks, Y., "Machine Translation; Its Scope and Limits"

\* Karttunen, L. & "Finite State Morphology"  
Beesley, K.

3. Forthcoming titles, without projected publication dates:

\* Asher, N. & "From Discourse to the Lexicon:  
Lascarides, A., Interpretation in Context"

\* Basili, R. et al., "Lexical Acquisition for Practical NLP Systems"

- \* Carroll,J., "Practical Natural Language Parsing"
- \* Wilks,Y. "Language, Cohesion and Form: Selected Papers of Margaret Masterman"
- \* Carl Pollard "Generalized Phrase Structure Grammars"

### 3.5 ACL Anthology Report for 2002

Steven Bird

ACL Anthology is a digital archive of research papers in computational linguistics. Once it is completed, it will contain approximately 50,000 pages of documents spanning a 40 year period. All of this material is being made freely available to ACL members and non-members alike, in commemoration of our 40th anniversary. The Anthology will be officially launched on the opening day of our 40th annual meeting, in Philadelphia on July 8, 2002.

The impact of the ACL Anthology will be far reaching. It opens our literature to newcomers and makes our field more accessible. The full content index makes it easy for us to discover past papers that are relevant to our research. As authors we benefit from the fact that online articles are more highly cited, and the fact that our papers can be discovered and indexed by citation services (e.g. citeseer). As teachers we can assign readings to our students as easily as sending them a URL. With easy access to early materials we can study the history of ideas in our field, or analyze the development of new trends. Finally, we can turn our own NLP technologies loose on our corpus of publications, and develop innovative new services based on the content.

#### CURRENT STATUS

ACL Proceedings	1979-1999	introduction by John Nerbonne
CL Journal	1980-2001	introduction by Julia Hirschberg
EACL Proceedings	1983-1999	introduction by Donia Scott
ANLP Proceedings	1983-2000	introduction by Sergei Nirenburg
NAACL Proceedings	2000	introduction by Diane Litman
TINLAP Proceedings	1978,1987	introduction by Bonnie Webber

Total Size: 22,000 pages, 120 volumes, 3,100 papers

equivalent to a stack 6'6" / 2m tall

## ORGANIZATION

### Anthology:

- \* PDFs in the format "PDF Image with Hidden Text", with HTML index pages
- \* Full content indexed by web search engines (e.g. Google)

### Archive:

- \* 600dpi greyscale tiffs
- \* RTF OCR output

## THE PROCESS OF BUILDING THE ANTHOLOGY

### 1. Inventory

- (a) Proceedings held in the ACL office - Priscilla Rasmussen  
(includes page counts, number of copies, location, date)
- (b) Maintaining inventory spreadsheet, track donors - Steven Bird

### 2. Collecting out-of-print materials

- (a) Old ACL proceedings - Mike Rosner  
(located several proceedings for late 1970s)  
(still looking for any papers or abstracts prior to 1975)
- (b) ACL workshop proceedings - Steven Bird  
(located donors for about 20 volumes)
- (c) COLING proceedings - Nicoletta Calzolari  
(located donors for all COLINGs, 1965-)

### 3. Obtain copyright

- (a) Identify copyright holder for each volume - Priscilla Rasmussen
- (b) Obtain letter from U Montreal for COLING/ACL-98 - Elliot Macklovitch
- (c) Announce intention to republish materials - Steven Bird
- (d) Notify Chicago and MIT Press re MT&CL - Steven Bird

### 4. Scanning

- (a) Specification - Steven Bird & Pinehurst Technologies
- (b) Collecting, packing and shipping proceedings - Priscilla Rasmussen
- (c) Scanning, deskewing, PDF compilation, OCR - Pinehurst Technologies

(d) Quality control - Steven Bird

5. Sponsorship

(a) North American sponsors - Eduard Hovy

(b) European sponsors - John Nerbonne

(c) Publicity, processing payments - Priscilla Rasmussen

6. Website

(a) Upload, indexes, search interface, logos from sponsors - Steven Bird

(b) Write introductory essays - John Nerbonne, Julia Hirschberg, Donia Scott, Sergei Nirenburg, Diane Litman, Bonnie Webber

(c) Write historical essays - Aravind Joshi, Karen Sparck-Jones, Mitch Marcus

ONGOING AND FUTURE WORK

Current scanning work (June 2002): 60 ACL workshops, TINLAP 1975

Future scanning work:

\* COLING proceedings (1965-), COLING workshops (July-September 2002)

\* Any ACL proceedings prior to 1975

\* microfiche journal issues (1974-78)

\* Finite String Newsletter (1964-1995)

\* Mechanical Translation and Computational Linguistics (1961-68)

Other tasks:

\* Integrate proceedings already published on CDROM - David Yarowsky

\* BibTeX data - Doug Arnold

\* Establish process for integrating future proceedings - Steven Bird

\* Other formats: DjVu, XML OCR data - Yann LeCun, Steven Bird

\* Clean up metadata - TBD

\* Extract abstracts - TBD

\* Extract references and crosslink - TBD

\* DVD publication - Steven Bird

\* Maintain long-term website - TBD

\* Specialized metadata for subject classification - TBD

\* Specialized metadata for language identification - TBD

\* Identify mirror sites - Steven Bird

## COSTS

The original estimate was 60c/page, for 300dpi b+w scans. This is the current industry standard for written documents. However, at the ACL executive committee meeting in Toulouse, it was agreed that 600dpi would be preferable if it wasn't too much more expensive. Considering the b+w, 16-level and 256-level grayscale options, I have selected 16-level, balancing quality and cost issues.

The price for 600dpi, 16-level grayscale scanning is 80c/page. This figure includes deskewing, creation of TIFF, RTF and PDF files according to the specification. The XML index files cost around 50c/paper. Delivery on DVD-ROM is \$175 each (supplied in duplicate for safe storage in two locations). When the per-paper and per-DVD costs are added in, the project comes to almost exactly \$1/page. If all the materials can be collected, the entire project will cost around \$50k.

The scanning company, Pinehurst Technologies Inc, was highly recommended by Michael Ley (Editor, ACM SIGMOD Anthology) and Bernard Rous (Electronic Publishing Program Director, ACM), and their work has been of the highest quality.

As of June 2002, over 90% of the needed funds have been either donated or pledged. This funding is being used solely for the scanning work. The rest of the anthology is "sweat-ware".

## COPYRIGHT

The ACL and ICCL hold copyright on all the volumes being scanned. However, in the case of proceedings volumes, this copyright only applies to the compilation; the copyright of the individual papers remains with the authors. Any authors who object to having their papers included in the anthology should contact Steven Bird to have their papers removed.

## ANTHOLOGY SUPPORTERS

As of June 2002, the supporters of the anthology project are as follows:

DONORS OF MULTIPLE VOLUMES: Betty Walker, Jane Robinson, Helen Gigley, Karen Sparck Jones, Jussi Karlgren, Eva Hajicova, Petr Sgall, Antonio Zampolli, Nicoletta Calzolari, Elliott Macklovitch, Kenneth Church, Eduard Hovy, Steven Bird.

DONORS OF SINGLE VOLUMES: Elizabeth Andre, Sabine Bergler, Jason Eisner, Andy Kehler, Guy Lapalme, Alberto Lavelli, Winfried Lenders, Bente Maegaard, Mark Maybury, Kathleen McCoy, David McDonald, Ruslan Mitkov, Johanna Moore, John Nerbonne, Yael Netzer, Miles Osborne, Hannes Pirker, James Pustejovsky, Mike Rosner, Donia Scott, Evelyne Viegas, Bonnie Webber, Yorick Wilks.

DONORS OF FINANCIAL SUPPORT: Margaret Fleck, Lillian Lee, Pierre Nugues.

INSTITUTIONAL DONORS: Gold level: Advanced Research and Development Activity; Silver level: German Research Center for Artificial Intelligence, Information Sciences Institute (USC), Linguistic Data Consortium (UPenn), MITRE Corporation, Netherlands Organization for Scientific Research, Xerox Research Centre Europe; Bronze level: Center for Language and Speech Processing (JHU), City University of Hong Kong, Columbia University, Information Technology Research Institute (Brighton), Macquarie University, School of Behavioral and Cognitive Neurosciences (Groningen).

AUTHORS OF ANTHOLOGY INTRODUCTIONS: John Nerbonne, Julia Hirschberg, Donia Scott, Diane Litman, Mark Johnson, Martin Kay, Sergei Nirenburg, Bonnie Webber, Ralph Weischedel, Mitch Marcus, Karen Sparck-Jones, and Aravind Joshi.

ASSOCIATE EDITORS: Eduard Hovy, John Nerbonne, Mike Rosner, Nicoletta Calzolari, David Yarowsky and Doug Arnold.

The ACL Executive Committee expresses its deep gratitude to these individuals and institutions for their wonderful support.

#### **4.1 Report on HLT-02**

Kathleen F. McCoy

This year HLT-02 was sponsored by ACL. This means that ACL provides the financial backing for the conference, and keeps the profits or losses that the conference generates. This was an unusual sponsorship, and was undertaken primarily as a gesture of good will towards HLT (which will be holding its 2003 conference jointly with NAACL).

Usually when ACL sponsors a conference we require ACL membership of all members attending. For

this conference we relaxed this rule and instead provided a 50% ACL membership coupon for all of those attending (thus the conference itself paid for the other 50% of the membership price for all of those who took advantage of the offer).

The conference took place at the end of March, and had around 235 participants. While the conference finances have not been finalized, it looks as if the conference will make a small profit. To date, it looks like only a small number of people have taken advantage of the 50% off membership coupon. On the other hand, Priscilla, who handled the conference through her company ARCS, has noticed that ACL-02 has gained some new participants because of HLT-02.

## **5.1 ACL-03 local arrangements Report for 2002**

Preparations for the ACL 2003 in Sapporo

June 23, 2002

Kenji Araki ([araki@media.eng.hokudai.ac.jp](mailto:araki@media.eng.hokudai.ac.jp))

This report describes the status of preparations for ACL 2003. The prices given do not include Japan's 5% consumption tax. The exchange rate is about 124 yen to the U.S. dollar.

### **1. ROOMS**

We are planning for ACL 2003 to be held at Sapporo Convention Center, which is scheduled to be completed in December 2002, officially opening on June 1, 2003. This gives us six months between completion and the conference. We have also reserved conference rooms at Hokkaido University as a fallback. Rooms reserved at Sapporo Convention Center are as follows:

- Plenary sessions (3 rooms): 3 large conference rooms (capacity: 600, 320, 191) x 3 days
- Workshops (11 rooms): 3 large conference rooms (capacity: 600, 320, 191) x 2 days;  
2 medium-sized conference rooms (capacity: 180 each) x 2 days; 6 small conference rooms (capacity: 80 each) x 2 days
- Tutorial (2 rooms): 2 large conference rooms (capacity: 320, 191) x 1 day
- Demonstration & exhibition: 2 medium-sized conference rooms x 6 days
- Anterooms for organizers: 1 small meeting room (capacity: 24) x 6 days
- Reception (welcoming cocktail party): 1 large conference room (capacity: 600) x 1 day
- Banquet: 1 large conference room (capacity: 600) x 1 day

Audiovisual equipment (Internet, projectors, OHP) is also available.

## 2. SCHEDULE

Planned schedule of ACL 2003:

July 7, 2003: 2 tutorial sessions

July 8 - 10, 2003: main conference (3 concurrent sessions)

July 11 - 12, 2003: 11 workshops

We are planning a digital poster session. In this session, presenters will explain their research through posters and demonstrations. We think the presentations will be very helpful in giving the attendees an understanding of the research. The schedule of the digital poster session is as follows:

July 9, 2003: 16:00 - 18:00

July 10, 2003: 16:00 - 18:00

Presentations number 15 per session, for a total of 30 presentations.

The details of excursions and social events such as banquets and receptions have not been decided.

## 3. SPONSORS

The City of Sapporo and the Hokkaido Government have agreed to provide subsidies. Those from the City of Sapporo total 1,500,000 yen and 200,000 yen from the Hokkaido Government. We are planning to secure 3,000,000 yen from local companies paying demonstration fees at the exhibition.

## 4. PUBLICITY

We have made a Website (<http://www.ec-inc.co.jp/ACL2003>), which includes information on accommodations and Sapporo. We plan to make a pamphlet and poster for distribution at ACL 2002 and COLING 2002, and for posting to major NLP researchers so that they may distribute them to their communities. Moreover, we will distribute a call for papers using electronic media.

## 5. ACCOMMODATIONS

We have reserved 400 rooms (capacity: 540 persons). These hotels are divided into 3 classes by room rates.

Group A: 13,000 yen for single, 22,000 yen for twin

Group B: 8,500 yen for single, 15,000 yen for twin

Group C: 7,000 yen for single, 13,000 yen for twin

For attending graduate students we have made reservations at Sapporo Guest House, managed by the

City of Sapporo.

The low room rates are as follows: 5,000 yen for single, 9,600 yen for twin, 18,000 yen for triple

This facility has only 9 rooms (capacity: 13 persons), so we will reserve another inexpensive accommodation: Clark Kaikan, managed by Hokkaido University. The room rates are as follows: 4,500 yen for single, 7,400 yen for twin. This facility has 10 rooms (capacity: 17 persons).

## 6. BUDGET ESTIMATES

(1) 500 attendees

### Income

Registration Fee	18,000,000 yen
Regular @40,000×400 =	16,000,000
Student @20,000×100 =	2,000,000
Subsidy from the City of Sapporo	1,500,000 yen
Subsidy from the Hokkaido Government	200,000 yen
Sponsorship income	3,000,000 yen
total	22,700,000 yen

### Expenditure

Costs of hiring conference venue (including fixed equipment)	2,000,000 yen
Costs of PCO contracts or costs required for pre-/on-site registration	500,000 yen
Printing	1,500,000 yen (Including Proceedings, Circulars, etc.)
Hire of audio visual equipment and technicians	3,000,000 yen (Including PC, projectors, telephone and fax rental costs)
Hire of poster panels	2,500,000 yen (Including signs, booths, lighting equipment, etc.)
Costs for invited speakers	1,500,000 yen @500,000×3
Estimated Local Committee expenses	2,800,000 yen
(Including costs for expendable supplies, conventions bags, bookbinding, communications, labors, transportation, etc.)	
Meeting of Program Committee	600,000 yen
Signs, banners, etc.	300,000 yen
Refreshments (coffee, tea, juice, cakes, etc.)	1,000,000 yen

Welcome Reception costs	2,000,000 yen
Banquet costs	5,000,000 yen @ 10,000×500
total	22,700,000 yen

(2) 400 attendees

Income

Registration Fee	14,400,000 yen
Regular @40,000×320=	12,800,000
Student @20,000× 80=	1,600,000
Subsidy from the City of Sapporo	1,500,000 yen
Subsidy from the Hokkaido Government	200,000 yen
Sponsorship income	3,000,000 yen
total	19,100,000 yen

Expenditure

Costs of hiring conference venue (including fixed equipment)	2,000,000 yen
Costs of PCO contracts or costs required for pre-/on-site registration	425,000 yen
Printing	1,200,000 yen
(Including Proceedings, Circulars, etc.)	
Hire of audio visual equipment and technicians	3,000,000 yen
(Including PC, projectors, telephone and fax rental costs)	
Hire of poster panels	2,000,000 yen
(Including signs, booths, lighting equipment, etc.)	
Costs for invited speakers	1,500,000 yen @500, 000×3
Estimated Local Committee expenses	1,720,000 yen
(Including costs for expendable supplies, conventions bags, bookbinding, communications, labors, transportation, etc.)	
Meeting of Program Committee	480,000 yen
Signs, banners, etc.	300,000 yen
Refreshments (coffee, tea, juice, cakes, etc.)	800,000 yen
Welcome Reception costs	1,600,000 yen
Banquet costs	4,000,000 yen @ 10,000×400
total	19,175,000 yen

## 7. OTHER TASKS

We would like to decide on all chairs, including program chairs, the sponsorship chair and the demonstration chair.

### **5.2 ACL-04 site selection Report for 2002**

ACL'04 Conference Organization

Mark Johnson

The call for bids for ACL'04 has closed, and we received several excellent bids from sites in Europe. At the ACL winter meeting we decided to ask two sites to prepare more detailed bids. At the same time an ACL'04 organizing committee was formed, consisting of members of the ACL and EACL executive committees. We received two excellent detailed proposals, and it is clear that either site could have hosted an excellent ACL meeting. After a detailed review and much discussion, we decided to ask the team from Barcelona to host ACL'04.

### **6.1 General Report for 2002**

Special Interest Groups

Mark Johnson

2001-2002 was a very active year for many of the ACL SIGs. Many of the SIGs have grown to the extent that they are as large and as active as the whole of ACL was not so long ago.

SIGDAT had not submitted a report by the time this summary was written.

SIGDIAL held elections in 2001. It sponsored or supported seven events this year.

SIGGEN held elections in 2001. It is sponsoring one event this year.

SIGHAN held elections in 2001. It sponsored two events this year.

SIGLEX has set up a specialized subgroup, SENSEVAL, to run the SENSEVAL workshops, and is sponsoring four events this year.

SIGMEDIA sponsored or supported three workshops or meetings this year.

SIGMOL had not submitted a report by the time that this summary was written.

SIGNLL held elections in 2001. It sponsored several events this year.

SIGPARSE had not submitted a report by the time this summary was written.

SIGPHON is running biannual workshops, including one co-sponsored by SIGNLL.

SIGSEM had not submitted a report by the time this summary was written.

I'd like to thank Hiroshi Nakagawa

<<http://www.r.dl.itc.u-tokyo.ac.jp/~nakagawa/English.html>

for his excellent web-based report system, which made the preparation of this summary very easy!

## **6.2 SIGDAT - 2002 Annual Report**

(Ken Church, David Yarowsky)

SIGDAT is ACL's special interest group for linguistic data and corpus-based approaches to NLP.

In 2002, SIGDAT organized the 2-day Conference on Empirical Methods in Natural Language Processing (EMNLP-2002). The meeting was scheduled immediately prior to ACL-02 at the University of Pennsylvania in Philadelphia on July 6 and 7.

Jan Hajic served as conference chair and Yuji Matusumoto served as co-chair. 35 papers were accepted for presentation out of 142 submissions (a 24.6% acceptance rate), yielding a program and proceedings equivalent in size to full ACL conferences before 1997.

Preregistration to EMNLP-2002 exceeds 215 persons. Generous sponsorship was received from Justsystem Corporation and CLAIRVOYANCE Corporation.

SIGDAT also co-sponsored the Sixth Conference on Natural Language Learning, with SIGNLL, to be held in Taipei on August 31-September 1. Dan Roth, Antal van den Bosch and Erik Tjong Kim Sang are co-chairs.

## **6.3 SIGDIAL Report 2002**

REPORT ON SIGDIAL ACTIVITIES

SIGdial is the ACL and ISCA Special Interest Group on Discourse and Dialogue which was formed in November 1997. This report covers the period from July 2001 to June 2002.

PEOPLE:

The second SIGdial elections were held in the autumn of 2001. As a result of the elections and

subsequent nominations, SIGdial is now being run by the following team:

President: Laila Dybkjær; Vice President: David Traum; Secretary: Jennifer Chu-Carroll; Science Advisory Committee: Jan Alexandersson, Kristiina Jokinen, Akira Kurematsu, Alexander Rudnicky, Ronnie Smith, and Jan van Kuppevelt; Information Officer: Claudia Soria assisted by Kristiina Jokinen, Eric Chang, and Pamela Jordan; ACL Student Liaison: Weiqun Xu and Holmer Hensen; CHI Liaison: Nils Dahlbäck; ICSLP/ISCA Liaison: Julia Hirschberg; Mailing list maintenance: Laurent Romary. SIGdial has 250 members registered on its membership web page.

#### WORKSHOP ENDORSEMENT

During the past year several workshops with strong relations to the SIGdial field of interest have been endorsed by SIGdial. The workshops endorsed by SIGdial which have taken place within the reported period include:

July 6-7 2001: ACL-EACL Workshop on Evaluation for Language and Dialogue Systems (Toulouse, France).

August 5 2001: 2nd IJCAI Workshop on Knowledge And Reasoning In Practical Dialogue Systems (Seattle, USA).

August 20-24 2001: ESSLLI Workshop: Information Structure, Discourse Structure and Discourse Semantics (Helsinki Finland).

October 22-24 2001: ISCA Tutorial and Research Workshop on: Prosody in Speech Recognition and Understanding (Red Bank, USA).

June 4 2002: ITS 2002 Workshop on Empirical Methods for Tutorial Dialogue Systems (San Sebastian, Spain).

June 17-19 2002: ISCA Tutorial and Research Workshop on Spoken Dialogue in Mobile Environments (Kloster Irsee, Germany).

Workshops which have been endorsed by SIGdial but haven't been held yet include:

September 4-6 2002: EDILOG 2002: Sixth Workshop on the Semantics and Pragmatics of Dialogue (Edinburgh, Scotland).

#### SIGDIAL WORKSHOP SERIES

Following the first and successful SIGdial workshop on Discourse and Dialogue held on 7-8 October 2000 in Hong Kong in conjunction with ACL'2000, the second SIGdial workshop on Discourse and Dialogue was held on 1-2 September 2001 in Aalborg, Denmark, in conjunction with Eurospeech'2001. This workshop was a very successful event with 90 participants - more than twice as many as

participated in the first SIGdial workshop. The workshop received 57 paper submissions. Of these 29 were selected for presentation. The online workshop proceedings are available at [http://www.sigdial.org/sigdialworkshop01/acceptance\\_list.php](http://www.sigdial.org/sigdialworkshop01/acceptance_list.php). Hardcopies can be bought via ACL. Based on some of the best workshop submissions a book is now being edited by the workshop chairs Jan van Kuppevelt and Ronnie Smith. The book will be published by Kluwer later this year. The third SIGdial workshop will take place on 11-12 July in Philadelphia, USA, in conjunction with ACL'2002. Information on SIGdial workshops can be found at <http://www.sigdial.org/workshops.php>.

## WEB

SIGdial has a web site at <http://www.sigdial.org>. The web site is hosted by NISLab. Some pages are maintained by a particular person, such as the information officer, while other pages are intended to be maintained by the membership collectively. For instance, the web site enables members to easily update their personal information in the membership database. Also a resource page has been established which members are encouraged to collectively maintain by adding references to resources they are aware of and which may be of interest to the SIGdial community. This page includes references to e.g. coding schemes, language resources, spoken dialogue systems, and tools.

## OTHER ACTIVITIES

Since autumn 2000 SIGdial has a collaboration with Elsnets, the European Language and Speech Network. Elsnets has a quarterly newsletter, Elsnets, and Elsnets has generously offered one page in each issue of Elsnets for SIGdial purposes. SIGdial contributions in Elsnets can be found in issue 9.3 and onwards. Elsnets can be freely downloaded in pdf format from <http://www.elsnets.org/elsnets.html>. Contributions from the reported period include

1. Kristiina Jokinen: The USIX Interact Project: Adaptivity in Dialogue Systems.
2. Hans Dybkjær: Stray Thoughts from a Crossroad.
3. Alexander Rudnicky: The DARPA Communicator Programme.
4. Yasuharu Den: Current Research Activities on Discourse and Dialogue in Japan.

A flyer on SIGdial was produced for distribution at Eurospeech'2001, the 2nd SIGdial workshop and the COCOSDA workshop. The flyer is available for download in pdf format from [www.sigdial.org](http://www.sigdial.org).

**SIGdial has a business meeting each year in connection with the SIGdial workshop. Agenda, slides and minutes are available at <http://www.sigdial.org/businessmeetings/>**

## **6.4 SIGGEN Report 2002**

Report on SIGGEN 01-02

CONFERENCES AND WORKSHOPS: The major NLG event of the year is INLG-02, which SIGGEN is sponsoring and supporting. It will be held in Columbia University's Arden House, near New York City, just before ACL. For more information see the conference web site at <http://inlg02.cs.columbia.edu/>.

SIGGEN did not this year receive any requests to offer "support in name" for NLG workshops. This is probably because there are no specific NLG events this year other than INLG. This would be worrying if it marks the start of a trend, but we hope it is just a one-off blip, and that next year we will again see NLG workshops either stand-alone or as part of other conferences.

MAILING LIST AND WEB SITE: One technical change, which is that the mailing list was moved to a new server. But in general both seem to be working well, and providing useful information to the community.

ELECTION: An election was held for new board members. Owen Rambow and Helmut Horacek replaced Daniel Marcu and Keith vander Linden as permanent board members, and Noemie Elhadad replaced Irene Langkilde as student member.

CONSTITUTIONAL ISSUES: Comments were solicited from the community on a formal constitution, and a draft constitution was posted on the SIGGEN web page. We expect to discuss the constitution at INLG-02, and then hold a formal acceptance vote.

## **6.5 SIGHAN Report 2002**

SIGHAN, the ACL Special Interest Group on Chinese Language Processing, was approved by the Exec, Sunday July 8, 2001.

The SIG has a website: [www.sighan.org](http://www.sighan.org). The constitution for SIGHAN can be found at [www.sighan.org/sighan-constitution.html](http://www.sighan.org/sighan-constitution.html)

The first slate of officers is as follows:

Chair: Martha PALMER, University of Pennsylvania

Vice-Chair/Treasurer: Richard SPROAT, AT&T Labs

Secretary/Web Master: Tom EMERSON, Basis Technology

Information Officers:

China: YU Shiwen, Peking University  
 ZHOU Ming, Microsoft Research

Hong Kong SAR: Benjamin T'SOU, City University of Hong Kong

Singapore: TAN Chew Lim, National University of Singapore

Taiwan: HUANG Chu-ren, Academia Sinica  
 WANG Hsiao-Chuan, National Tsing Hua University

United States: XIA Fei, IBM

SIGHAN is officially involved in the following activities in 2002:

1. The First SIGHAN Workshop on Chinese Language Processing, to be held September 1, 2002, in conjunction with COLING 2002, in Taipei.
2. Co-sponsorship of the tutorials on Computational Linguistics and Chinese Chinese Language Processing, August 24, 2002 at COLING 2002. SIGHAN will be sponsoring at least two student fellowships to attend these tutorials, funded by proceeds from the workshop.

As of May 23, 2002, the SIGHAN membership breaks down as follows:

name_en	count(*)
Australia	1
Canada	2
China	19
Finland	1
France	1
Germany	1
Hong Kong	5
Japan	3
Singapore	4
South Korea	1
Taiwan	3
United Kingdom	1

United States	24
United States Minor Outlying Islands	1

Richard Sproat

Vice-Chair/Treasurer, SIGHAN

## 6.6 SIGLEX Report 2002

At ACL01, SIGLEX held the SENSEVAL-2 workshop. SENSEVAL-2 was an evaluation event in which there were Word Sense Disambiguation exercises for twelve languages, with around ninety participating systems from thirty-five research groups. The workshop was attended by 86 people.

Following SENSEVAL-2, it became evident that, while the exercise was clearly a success, some people felt that the highly informal organization should be regularized. Thus we set up SENSEVAL as an organization with its own committee and constitution, answerable to ACL-SIGLEX. We sought candidates for Chair and committee, and in the event returned Phil Edmonds (unopposed) as Chair, with a committee of nine, all of whom had been organisers or participants in SENSEVAL-2. This group is now proceeding with an ACL-02 workshop:

Word Sense Disambiguation: Recent Successes and Future Directions

co-chairs: Phil Edmonds, Rada Mihalcea (to incorporate a special session on Prepositions, chaired by Patrick st-Dizier)

which will look forward towards SENSEVAL-3. A Special Issue of the Journal of Natural Language Engineering (editors: Edmonds and Kilgarriff, publication: early 2003) also came out of SENSEVAL-2. SENSEVAL also has a mailing list, see

<http://www.hum.gu.se/mailman/listinfo/senseval-discuss>

SIGLEX took a role in initiating the First Conference on Global WordNet, in Mysore, India, in January 2002, in which 81 people from 19 countries participated.

SIGLEX is also sponsoring another workshop at ACL-02: Unsupervised Lexical Acquisition (chair: Joseph Pentheroudakis) and is co-sponsoring, with SIGSEM, the 5th International Workshop on Computational Semantics, in Tilburg in January 2003.

Adam Kilgarriff

Chair

## 6.7 SIGMEDIA Reports 2002

(SPECIAL INTEREST GROUP ON MULTIMEDIA LANGUAGE PROCESSING)

June 11th 2002

CHAIR: Elisabeth André (University of Augsburg, Germany, [andre@informatik.uni-augsburg.de](mailto:andre@informatik.uni-augsburg.de))

MAILING ADDRESS: [sigmedia@dfki.de](mailto:sigmedia@dfki.de)

URL: <http://www.sigmedia.org> <<http://www.sigmedia.org/>>

### CONFERENCES AND WORKSHOPS:

In the last reporting period, SIGMedia has been serving as an official sponsor of three scientific events. SIGMedia has been serving as a sponsor of the second Conference on Computational Semiotics for Games and New Media which will take place in Augsburg (Germany) on September 2nd - 4th, 2002. It is a cross-disciplinary conference that explores the ways in which semiotics can be applied to creating and analysing computer-based media. The organizers of the Conference are: Elisabeth André (Universität Augsburg, Germany, Andy Clarke (Kinonet, UK), Clive Fencott (University of Teeside, UK), Craig Lindley (Interactive Institute, Sweden), Grethe Mitchell (University of East London and Kinonet, UK) and Frank Nack (CWI, Netherlands). The URL of the workshop is: <http://www.kinonet.com/conferences/cosign2002>

SIGMedia is also involved in a joint activity with the ACL Special Interest Group SIGDial. Both of them support the International Speech Communication Association with the preparation of the ISCA Tutorial and Research Workshop on Spoken Dialogue in Mobile Environments. The workshop will take place at Kloster Irsee, Germany from June 17-21, 2002. The organizers of the workshop are: Paul Heisterkamp and Wolfgang Minker, both DaimlerChrysler AG, Germany, Laila Dybkjaer, University of Southern Denmark and Elisabeth André, University of Augsburg, Germany.

The URL of the workshop is: <http://www.sigmedia.org/ids02/>

Together with MITRE, SIGMedia has been serving as a co-sponsor of the ACL Workshop on Temporal and Spatial Information Processing which took place in Toulouse on July 7th 2001. The organizers of the workshop were: Lisa Harper, MITRE, USA, Inderjeet Mani, MITRE and Georgetown University, USA and Beth Sundheim, SPAWAR Systems Center, USA.

### COOPERATIONS WITH OTHER NETWORKS AND INSTITUTIONS:

SIGMedia members have been involved in a number of Expressions of Interest (EoI) proposals for Networks of Excellence devoted to natural-language processing in multimedia applications that have

been invited by the European Commission. SIGMEDIA also intends to establish links to the FET Convivio Network which is currently under negotiation and will become the successor of i3net ( <<http://www.i3net.org/>> [www.i3net.org/](http://www.i3net.org/)) with which we had a number of common interests.

## **6.8 During the past year, the activities of the Mathematics of Language (MOL) SIG have included Report for 2002**

+ a joint Formal Grammar/Mathematics of Language Conference last August in Helsinki in conjunction with the European Summer School in Logic, Language, and Information (ESSLLI), including invited lectures by Jaakko Hintikka and Aravind Joshi, and a symposium on Model-Theoretic Syntax with contributions from Geoff Pullum (keynote), Patrick Blackburn, Jim Rogers, and Uwe M"onnich; for conference information, see:

<<http://www.let.uu.nl/~Richard.Oehrle/personal/fgmolscreen.pdf>>;

for the program,

<<http://www.let.uu.nl/~Richard.Oehrle/personal/fgmol01prog.pdf>>

The proceedings of the conference are being published by Electronic Notes in Theoretical Computer Science, volume 53, accessible via the ENTCS home page at

<<http://www.elsevier.nl/gej-ng/31/29/23/show/Products/notes/index.htm>>.

a selection of the papers will appear in a special issue of the Journal of Language and Computation, in preparation;

+ a one-week workshop during the first week of the ESSLLI summer school, with presentations by Christian Wartena, Aravind Joshi, Yiannis Moschovakis, Michael Moortgat & Dick Oehrle, and Ed Keenan;

+ a MOL-sponsored symposium at the 2002 AAAS Meeting in Boston, entitled 'Reconciling Expressivity and Tractability: Finite State Methods in Natural Language Analysis', with contributing talks by Aravind Joshi, Ron Kaplan, Lauri Karttunen, Steven Bird, and Richard Sproat.

+ Jim Rogers was elected VicePresident (2001-2003)/PresidentElect (2004-2005).

The MOL community would like to acknowledge their gratitude to Larry Moss for his contributions as President for the past two years.

Dick Oehrle  
MOL President  
20 June 2002

## **6.9 SIGNLL Report 2002**

### ACL SIGNLL - President's Report 2001

An election for SIGNLL officers was organized after the summer of 2001. New president is Walter Daelemans, secretary Dan Roth, and the advisory board consists of Thorsten Brants, James Cussens, Diane Litman, Raymond Mooney, John Nerbonne, Miles Osborne, David Powers, Adwait Ratnaparkhi, and Erik Tjong Kim Sang. Our SIGDAT Liaison Representative remains David Yarowsky, and our Information Officer Antal van den Bosch.

In 2001 SIGNLL has an estimated 220 members (judging from the on-line information, and controlled for inactive email addresses). The SIGNLL home page is located at URL <http://www.aclweb.org/signll/>.

The goals of the SIG (promoting of and informing about research on computational language learning) are served by the following two activities: (i) the maintenance of an informative and up-to-date website and associated mailing list, and (ii) the organization of at least one yearly event (CoNLL), and support of other related activities.

The web-pages, maintained by Antal van den Bosch, remain an important source of information, complemented by an email list for conference announcements. On the web-site, links can be found to relevant associations, networks, research cooperations, research departments, groups, institutes, individuals, mailing lists, archives, journals, bulletins, conference reports, online papers, online courses and slides, bibliographies, software, corpora, companies, meta-information sources etc. A major addition was made to the website in 2001 with the incorporation of a separate page for a CoNLL on-line proceedings archive, going back to the 1997 meeting in Madrid.

The main event in 2001 was the fifth CoNLL (SIGNLL Conference on Natural Language Learning), organized with the ACL in Toulouse. Again, this was a successful event, with 60 submissions of which 18 could be accepted, and invited talks by Eric Brill (sponsored by CLIF, <http://clif.uia.ac.be>) and John Nerbonne. A shared task on clause segmentation was organized as part of the CoNLL, to which 6 systems participated. More information can be obtained from the SIGNLL website. 75 persons registered officially for the workshop.

The sixth CoNLL will be organized in Taipei by Roth and van den Bosch, with a shared task

organized by Tjong Kim Sang on named entity recognition. CoNLL was merged this year with SIGDAT's VLC and the LLL workshop. 44 papers were submitted of which 23 were accepted for presentation and publication in the proceedings. 12 systems were submitted for the shared task, and will also be presented and included in the proceedings. We look forward to yet another successful CoNLL.

We think SIGNLL is still unique in its multiple focus (computational models of language learning both for language engineering and for testing psycholinguistic and linguistic theories; formal and empirical aspects of learning of both artificial and natural languages). We keep striving for complementarity with related SIGDAT events such as EMNLP and WVLC, and have contributed to interdisciplinary communication by our conference collocation policy.

Walter Daelemans

Antwerp, May 23, 2002

## **6.10 SIGPARSE Report for 2002**

2001-2002 report from SIGPARSE

-----  
Over the period of reporting, June 2001 - June 2002, the activities of SIGPARSE have, as usual, focussed on matters concerning the biennial 'International Workshop on Parsing Technologies' (IWPT) series. Ensuring the continuity of this series is the main aim of SIGPARSE.

The biennial rhythm of the IWPT workshop series (1989-91-93-95-97) was slightly disturbed in 1999, when it was decided to move the workshop that had originally been scheduled for late 1999, to 2000 for fear of logistic problems relating to the millenium change. The 6<sup>th</sup> International Workshop on Parsing Technologies thus became IWPT 2000 rather than IWPT'99; it took place in Trento, Italy.

It was subsequently decided to return to the original rhythm, and to have the 7th International Workshop on Parsing Technologies in 2001. It was for the first time held in Asia, at the University of Beijing, hosted by the Institute for Computational Linguistics, where it took place on 17-19 October. The preparations and the attendance were hurt by the September 11 events, which prevented one of the invited speakers (Michael Collins), several participants, and one of the local organisers (Fuliang Weng, who had just moved from Beijing to California) to travel to Beijing for the workshop. As a result, attendance was slightly less than at the previous workshop. The local organisation in Beijing was in the very capable hands of a committee consisting of Shiwen Yu, Houfeng Wang, Fuliang Weng, Youqi Cao and Qunxin Chen, with the support of Xiajing Bai. The local arrangements were organised in a truly

excellent fashion; the Institute for Computational Linguistics provided excellent facilities, the staff was very helpful, and the beautiful campus of the University of Beijing formed a wonderful setting. Giorgio Satta was program chair and Harry Bunt general chair; John Carroll advised on many issues arising in the organisation of the workshop. Aravind Joshi and William Cohen were invited speakers; Eric de la Clergerie and John Carroll filled in for Michael Collins and for another speaker who was unable to attend. The proceedings, containing texts of the 3 invited talks, the 15 accepted submitted papers, and abstracts of 14 poster presentations, were available at the workshop.

In the meantime, preparations have started for IWPT'03. The 8<sup>th</sup> International Workshop on Parsing Technologies will be held from 23 to 25 April in Nancy, France. Gertjan van Noord will be program chair; Guy Perrier will be local arrangements chair.

During the period reported here, Harry Bunt (Tilburg University) has continued as SIGPARSE officer. Alon Lavie at CMU has been most helpful in continuing to maintain the active SIGPARSE email list ('sigparse-project@cs.cmu.edu'). After 8 years of providing this service, Alon Lavie has per 17 June handed over the monitoring of the sigparse-list to Kenji Sagae at CMU.

A SIGPARSE website is maintained at the Department of Computer Science at the University of Twente in Enschede, the Netherlands, by Hendri Hondorp. This site can be found at <http://parlevink.cs.utwente.nl/sigparse/>.

Harry Bunt, June 2002.

## **6.11 SIGPHON Report for 2002**

2001-2002 Annual Report SIGPHON (Computational Phonology)

Jason Eisner

### **MEMBERSHIP**

SIGPHON is ACL's special interest group for computational phonology. Membership currently stands at 214 (up from 190 last year and 176 the year before), with 47 members (up from 39 and 36) declaring computational phonology as their "primary" interest.

### **WORKSHOP**

SIGPHON generally runs biannual workshops. The workshop at ACL-02, chaired by Mike Maxwell,

has the theme "Morphological and Phonological Learning" and is cosponsored by SIGNLL. This is SIGPHON's sixth workshop. The program includes an invited talk and 8 papers. It features 2 invited papers by academic linguists, in keeping with our desire (mentioned in last year's report) to increase crosstalk with the pure linguistics community.

#### ONLINE ACTIVITIES

SIGPHON continues to serve the community by maintaining a mailing list and online bibliographies at its web site, [www.cogsci.ed.ac.uk/sigphon/](http://www.cogsci.ed.ac.uk/sigphon/). Thanks to the efforts of the SIG's secretary, Steven Bird, the mailing list will soon be integrated with the LINGUIST list, which provides list archiving, search, and broader exposure.

#### FUTURE PLANS

The Executive Committee hopes to arrange for a special issue of a journal such as *Phonology*, focused on the contribution of computational phonology to phonology proper. To further highlight the relevance of computational phonology to practicing phonologists, we are considering holding our next workshop at a linguistics conference such as the LSA.

### **6.12 SIGSEM Report for 2002**

-----  
Annual Report on SIGSEM, June 2002

Patrick Blackburn and Harry Bunt  
-----

It has been a relatively quiet year for SIGSEM. Membership has stabilised at 559, and after a heady start we're now faced with the longterm task of trying to build a useful SIG.

One major event was held by SIGSEM over the last year. The Third Inference in Computational Semantics workshop (ICoS-3) was co-located with IJCAR in June 2001 in Siena, Italy. IJCAR, the International Joint Conference on Automated Reasoning, was a one-off event made up of the three main automated reasoning conferences (CADE, FTP, and TABLEAUX). By co-locating with IJCAR in this way, it was hoped to establish further links between the automated reasoning community and the computational semantics community. A report on ICOS-3 can be found on the SIGSEM homepage ([www.sigsem.org](http://www.sigsem.org)). The ICoS-3 proceedings are currently being edited: they are due to appear in a special issue of the Journal of Language and Computation.

The next major SIGSEM-endorsed activity will be IWCS-5 (The 5<sup>th</sup> International Workshop on Computational Semantics), scheduled for January 15-17, 2003 in Tilburg.

Another SIGSEM-endorsed initiative was proposed by Harry Bunt and Laurent Romary: A Working Group on Multimodal Meaning Representation. The idea is to create a working group devoted to this topic operating under the auspices of SIGSEM, using the SIGSEM web pages and newsletter for communication about the activities of the WG, and using SIGSEM-sponsored events like IWCS and ICoS as opportunities to meet and to present the activities to the computational semantics community. A preliminary meeting of the working group took place at LREC 2002 in May in Las Palmas, Spain; the first full-scale meeting will take place on January 14, 2003, in conjunction with IWCS-5.

## **7.1 ACL INTERNET SITE Report for 2002**

REPORT by Dragomir Radev

### INTRODUCTION

- o The ACL Web site is getting significantly more popular. From June 25, 2001 to June 25, 2002, the top-level page has been accessed 273,000 times. That is an average of 748 hits (accesses) per day (or an increase of 8% over the previous year).

### NEW ACTIVITIES

- o A vote for an ACL constitutional amendment was carried out electronically on the ACL Web site.
- o A new online election will be held soon after ACL-02.

### CONTINUING FEATURES

- o The ACL NLP/CL Universe, the catalog of Web-based NLP/CL resources, has grown at a steady rate. Since last year there has been a 10.1% increase in the number of pointers, bringing their total from 2279 to 2509. There are now 332 personal pages, 641 pointers to various resources (corpora, software, etc.), 366 academic or industrial labs, departments, and institutes, 65 professional organizations, etc. Babawole Akin-Aina has been in charge of adding new entries to the Universe. Mary Taffet helped a lot too.
- o The mailing list acl-news is used to announce new additions to the NLP/CL Universe. Currently, the list includes 587 subscribers, an increase of 60 since last year.
- o The unofficial Natural Language Processing FAQ (list of Frequently Asked Questions and Answers)

is still available through the ACL page.

Volunteers are sought to contribute to the list.

- o A major new initiative is the ACL Anthology, an online digital library consisting of all ACL, ANLP, NAACL, and CL journal papers ever published. More on the Anthology in Steven Bird's report.

#### ANNOUNCEMENTS

- o ACL members should make an attempt to announce all events, resources, or personal and lab pages to the [aclweb.org](http://aclweb.org) site using the on-line URL submission mechanism.

Please feel free to send me mail with suggestions or comments

([radev@umich.edu](mailto:radev@umich.edu))

## 7.2 ACL Archives Report 2002

Material continued to trickle into the archives during late 2001 and early 2002. I am happy to report that archival material is reaching me in a fairly timely manner now.

Updates during this period include the 2002 officer slate, the ACL 2003 bid, and some minor changes to the verbiage on the main index page.

Richard Sproat

ACL Archivist

## 7.3 ACL Course Survey

Following on from the surveys of computational linguistics courses carried out in 1986 by Robin Cohen and in 1994 by Bonnie Dorr, and the survey of graduate programs carried out by Martha Evens in 1992, we have put in place a new survey for 2002. This is the first time such a survey has taken advantage of the web for data capture and presentation. The survey is available at [www.clt.mq.edu.au/survey](http://www.clt.mq.edu.au/survey), and consists of a set of basic questions and set of more detailed questions; it takes between 10 and 20 minutes to complete, depending upon whether the visitor chooses to complete both parts. We encourage all those who teach courses that involve computational linguistics, natural language processing and language technology to complete the survey: we have deliberately drawn loose boundaries around what counts for inclusion in order to gather as much data as

possible.

Robert Dale and Mary Taffet

## **8.1 NLSR "Actual status of the ACL Natural Software Registry" Report for 2002**

---

During the last year the ACL Natural Language Registry (NLSR) team has been looking for extending the domain of application of the NLSR and for closer cooperation and interlinking with organizations and initiatives concerned with the listing and organization of NLP/CL resources.

In this context there have been among other discussions with ELRA, Inria (University Nancy) and the ISLE Metadata initiative (IMDI). Those discussions have resulted in the building of an European Consortium, which submitted a project proposal to the European Commission. A main topic of this (in the meantime accepted) project concerns the interoperability of language data and NLP tools repositories and the description of a web protocol supporting this interoperability. The protocol has been named "LREP" (Language Repository Exchange Protocol), and has been presented at the LREC 2002 conference. The ACL Registry will be modified in order to be able to access via this protocol the relevant language data repositories and to be accessed by queries coming from other repositories.

The NLSR has been invited by the Open Language Archives Community (OLAC) initiative to participate as a data provider. For this, the NLSR made its data conform to an OAI provider. Furthermore, an XML output been developed that is constantly available on the web, so that it can be harvested by the OLAC tools. There are still some few improvements to be implemented in order to make the ACL Registry fully compliant to the OLAC format.

The NLSR is participating as a data provider in the German Project "COLLATE" (Computational Linguistic and Language Technology), hosted in Saarbruecken at both DFKI and the University of the Saarland, and is now fully operational in the Virtual Information Center of this project ([www.lt-world.org](http://www.lt-world.org)). All the data of the ACL Registry are listed in the "Systems and Products" part of LT-World.

Thierry Declerck

Contact: [registry@dfki.de](mailto:registry@dfki.de)

## 8.2 OLAC Report for 2002

OLAC - Open Language Archives Community - [www.language-archives.org](http://www.language-archives.org)

Steven Bird and Gary Simons

OLAC, the Open Language Archives Community, is an international partnership of institutions and individuals who are creating a worldwide virtual library of language resources by: (i) developing consensus on best current practice for the digital archiving of language resources, and (ii) developing a network of interoperating repositories and services for housing and accessing such resources.

OLAC builds on the Open Archives Initiative and the Dublin Core Metadata Initiative, and is sponsored by the NSF/EC project "International Standards in Language Engineering" (ISLE).

The OLAC homepage, at [www.language-archives.org](http://www.language-archives.org), hosts many documents providing full technical details of OLAC, including the OLAC Metadata Set, the Metadata Schema, the OLAC Process document, information for implementers, and the OLAC-General Mailing list.

In the past 12 months, there have been three major developments with OLAC:

**LAUNCHES:** OLAC was formally launched in North America in a symposium held at the 76 Annual Meeting of the Linguistic Society of America in San Francisco (January 2002). OLAC was also launched in Europe in a symposium held at the 3rd International Conference on Language Resources and Evaluation, in Las Palmas, Spain (May 2002).

**PARTICIPATION:** Participation in OLAC has doubled. Today there are over 20 participating archives and services: ACL, ANLC, AISRI, APS, ATILF, CBOLD, DFKI, ELRA, Ethnologue, LACITO, LDC, LINGUIST, OLACA, OTA, Perseus Project, Rosetta, SCOIL, SIL, TalkBank, TRC and TRACTOR. Together these institutions have documented some 30,000 language resources (corpora, tools, publications, etc), and they can be searched simultaneously. Note that the ACL Anthology is indexed by OLAC.

**INFRASTRUCTURE:** Key software infrastructure is now in place, including Vida, a virtual data provider for proxied services; ORyX, and XML format for representing OLAC repositories in XML, ORE, the OLAC repository editor, a forms based metadata editor; and DP9, a web-crawler gateway permitting search engines to index all OLAC content.

There are three ways in which members of the community can become involved in OLAC:

1. Use OLAC to search for needed language resources. Users can search based on language, linguistic data type, linguistic software functionality, etc.
2. Document new resources using OLAC metadata. Use OLAC tools to create/publish descriptions of language resources.
3. Contribute to the development of controlled vocabularies for language resource description.

To find out more, visit [www.language-archives.org](http://www.language-archives.org) and join the OLAC-General mailing list.