NAACL 2010 Program Chairs Report

The 2010 NAACL HLT program includes innovative high-quality work spanning computational linguistics, information retrieval, and speech technology. The NAACL HLT 2010 consists of oral and poster presentations of full and short papers, pre-conference tutorials, application demonstrations, a student research workshop, and post-conference workshops. This year, we chose the special Noisy Genre theme to acknowledge the significant work taking place in that area across several disciplines. We would like to thank L. Venkata Subramaniam of IBM Research-India for this suggestion. A one-minute madness session has been introduced this year to highlight the poster and demo contributions. We are excited to have two very interesting and diverse keynote speakers: David Temperley, University of Rochester, whose talk is entitled*, "Music, Language, and Computational Modeling: Lessons from the Key-Finding Problem"*, and Steve Renals, University of Edinburgh, *"Recognition and Understanding of Meetings."* In addition, there is a panel session reflecting the Noisy Genre conference theme: *"Recent and Future HLT Challenges in Industry,"* chaired by Kristina Toutanova.

Conference Papers. This year 291 full papers were submitted and reviewed, of which 90 papers were accepted (a 30.9% acceptance rate); and 159 short papers were submitted and reviewed, of which 56 were accepted (a 35.2% acceptance rate). One of the accepted short papers was withdrawn to appear elsewhere, leaving a total of 55 short papers in the proceedings. Full papers describe substantial, original, completed and unpublished work, and whenever appropriate, included concrete evaluation and analysis. They were judged on appropriateness, clarity, originality/innovativeness, correctness/soundness, meaningful comparison, thoroughness, significance, contributions to research resources, and replicability. Short paper submissions were required to describe original and unpublished work with one of the following characteristics:

- A small, focused contribution
- Work in progress
- A negative result
- An opinion piece
- An interesting application nugget

Eighty-six full papers and 25 short papers will be presented as oral presentations. The remaining 34 accepted full and short papers will be presented as posters. Whether a paper was chosen for oral or poster presentation was determined based on input from the reviewers and senior program committee members (i.e., the area chairs), taking into account the constraints of putting together a sensible and coherent program. Although we asked that the oral versus poster decisions be based on the nature rather than on the quality of the work, we found that the notion that a poster is for a lower quality paper is entrenched in the minds not only of the reviewers, but also of many of the area chairs. Some people have expressed the view that full papers should not be presented as posters; however, rather than create odd mixtures of papers in one or more session, we opted to place some full papers in the poster session. This year we have introduced a one-minute madness plenary session to highlight posters and demos to be presented in the Poster and Demo Session. At this plenary session, Poster and Demo presenters will have the opportunity to summarize, in 60 seconds or less, their Poster or Demo using a single slide. We hope that this will help make poster presentations more acceptable to the community.

A strict conflict-of-interest policy was in place during the review procedures (both paper review and best paper selection). Specifically, authors who were on the Program Committee in any capacity were removed from any responsibility related to their papers. Every conflict-of-interest paper was handled by another committee member. As well, the author was prohibited from participating in any discussion or decision making related to their paper. A similar policy also applied if a Program Committee member had an institutional or personal conflict with an author. All decisions on papers with these types of conflicts were made by members without conflicts. The acceptance rate for short papers was lower this year than last year. The reviewers' scores for short papers were statistically significantly lower than for full papers (two tailed t-test, $p < 0.000290593$). This

is possibly due to the fact that they were reviewed at the same time as the full papers. See Table 1 below. It may make sense to look at the statistics from 2009 to determine whether the difference in scores is an artifact of the 2010 review policy.

Paper Type	AVG	STDEV
Long All	3.03	0.74
Long Accept	3.87	0.28
Long Reject	2.65	0.55
Short All	2.77	0.71
Short Accept	3.48	0.43
Short Reject	2.39	0.50

Table 1. Review Score Statistics

A full breakdown of submission and acceptance statistics by area and paper length is presented in Table 2 in Appendix A at the end of this report. We also provide the distribution of the main contact's country for paper submissions in Table 3 in Appendix B.

START System. The START conference management system was used to manage paper submissions and the review process—Rich Gerber and the START team provided invaluable help with the system. There were some hiccoughs in START this year, as a result of which several people had trouble submitting. The cause of these remains unknown, although it is clear that it was more than just network interruptions. There were also some less troubling inadvertent double submissions that required us to clean the submissions up manually (tedious and potentially errorful).

Reviewing Process. The review process was organized into a two-tier system, with eighteen senior program committee (SPC) members (or Area Chairs), and 382 reviewers. The SPC members managed the review process for both the full and short paper submissions: each paper received at least three reviews. Areas with more than 40 submissions were Machine Translation, Information Retrieval and Extraction, Machine Learning (Text), Parsing, and Semantics. Although we anticipated a large number of submissions in Machine Translation (and so had two area chairs), the number of submissions in the Information Retrieval and Extraction and Semantics areas was unexpected. The next program committee may want to consider recruiting reviewers more heavily in these areas. Also, more than one area chair could help balance the workload to ensure a high quality review process.

It is clear that the review process for NAACL could be improved, but this was a challenge due to the tight review schedule we needed to adopt to accommodate resubmission to ACL. Here is the schedule we worked with:
- December 18 (Friday) Paper reviewer assignments completed
- Jan 11 (Monday) Reviews due back to area chairs
- Jan 12 (Tuesday) – Jan 23 (Fri) Full review discussion period
  - o Jan 12 (Tuesday) Send reviews to authors for the author response period
  - o Jan 15 (Thursday) Author responses due back to area chairs
  - o Jan 12 (Tuesday) – Jan 15 (Friday) Discussion of reviews with your reviewers
  - o Jan 18 (Monday) – Jan 21 (Thursday) Area chairs teleconference with PC co-chairs to make final decisions
  - o Jan 21 (Thursday) Area chair decisions due to program chairs
  - o Jan 23 (Saturday) final PC decisions made
- Jan 25 (Monday) Acceptance notifications sent

Area Chairs and Related Communication[1]. Area chairs (ACs) are on the front lines and are central to ensuring that good decisions are made. ACs must:

- Select reviewers for the pool of papers that are knowledgeable (after the fact, a couple of ACs actually made statements such as, —"this reviewer did not know the area." This meant that the AC had to go back and re-review that paper which is additional work).
- Get discussions going on papers that have mixed reviews, as well as papers with author responses, as well as encourage reviews to be updated to reflect the discussion
- Look very closely at papers with mixed reviews because they need to make a solid recommendation; sometimes these papers are fabulous and new.
- Be well-prepared to make decisions about papers to accept and reject, based on more than just score averages. The PC needs to look across all score areas, and also take into account reviewer comments if they are somehow in conflict with scores, which does seem to happen. As well, we also found that there were some papers that were submitted that got high scores, but were very similar if not identical to papers appearing in other conferences in the recent past.
- Provide suggestions for best papers together with a rationale for their consideration. The AC's should look closely at the high ranking papers in their area, as well as all papers nominated by one or more reviewer.
- Provide suggestions for Keynote Speakers

This year we attempted to add another tool for the ACs to work with: an author response period, which START now supports. The START author response system allowed authors to enter a 500-word response to the reviews between January 12 - 15, 2010. Authors were asked to limit their response to items such as correcting factual errors in the reviews, answering questions raised in the reviewer comments, and so on. The response period overlapped in time with the reviewer discussions due to the tight schedule, which caused some problems that gave us, as well as some authors, some concern. Active ACs encouraged discussion of marginal papers and notified their reviewers when responses were available for those papers, but not all ACs did so. Hence, some reviewers considered the author responses, while others did not, and some marginal papers were thoroughly discussed, but not others. In addition, the START support for author responses unintentionally allowed both reviewers and authors to modify their comments on the system during the response period in such a way as to allow for nearly live conversations about the submissions under review. Neither party was required to comment further, which resulted in a few cases of authors or reviewers being frustrated when further responses never arrived. Because the reviews themselves were the sole means of communication by reviewers in these back-channel discussions, the final state of the reviews sometimes contained remarks that would only have made sense to someone who had followed the entire back-channel discussion, and thus was less useful to the ACs and PC chairs as a summary of the final opinions of the reviewers.

Scores were far more often lowered as a result of discussions/author input, for example, but this was in some cases inadequately captured in the reviews. A policy probably needs to be put in place, perhaps by the NAACL board, about score change as a result of the author response period, especially with regard to lowering of scores. Authors were obviously unhappy when an honest attempt to clarify issues for reviewers actually lowered their score. This could deter authors from engaging during the author response period. We also received emails from authors asking for the discussion threads, but these could not be forwarded because they would have disclosed the identities of the reviewers. Reviewers should have been encouraged by ACs to update their reviews when scores changed due to discussion and author responses. The START system should also be modified to freeze the state of the reviews prior to the discussion period, and then provide an extra section for reviewers to comment further in light of the responses. This feedback has already been communicated to the START support staff.

---

[1] See Appendix C for a sample invitation letter to Area Chairs.

Similar to recent years, we did not have a face-to-face meeting of the area chairs; instead we held a series of teleconferences between individual ACs and the PC chairs. We used doodle to collect the times that ACs were available in order to schedule the meetings in the short time we had between the completion of the review process and final decisions. The ACs were asked to prepare a spreadsheet that provided information about each paper (e.g., submission id, title, length, reviewer scores) and their recommendation for acceptance or rejection of the paper, as well as their suggestion for presentation mode. The spreadsheet was quite useful for focusing discussions on the borderline papers requiring discussion.

Best Papers. Two papers were selected for best paper awards this year:

- **Best Full Paper:** *Coreference Resolution in a Modular, Entity-Centered Model*, by Aria Haghighi and Dan Klein

- **Best Short Paper:** ``*cba to check the spelling"*: *Investigating Parser Performance on Discussion Forum Posts*, by Jennifer Foster

This selection was made by a committee chaired by Aravind Joshi, with Eugene Charniak, Michael Collins, Diane Litman, Daniel Marcu, and Drago Radev as its other members. We would like to congratulate the authors, Jennifer Foster, Aria Haghighi and Dan Klein, and especially thank the chair of the committee for managing the process so effectively and the rest of the committee for reading, discussing, and selecting the final best papers. The Best Paper process we used this year is as follows:

1. All reviewers were able to recommend a paper they reviewed in START for best paper.

2. At the program chairs' request, all of the area chairs were asked to look at these recommendations and other high quality papers in their area in order to nominate one or more paper for best full and best short paper.

3. Given this list, the program chairs assembled names of possible members for the Best Paper committee, and the program chair in charge of overseeing this process contacted these individuals in order to build a committee with 6 members, including a chair.

4. Best paper nominees were contacted to let them know that their papers were under consideration prior to the camera ready deadline. The hope was that the papers would be updated carefully for the best paper committee's perusal.

5. Once camera ready papers were submitted, all of the papers were sent to the committee for reading and assessment. At this time, 2 members identified a conflict of interest with 1-2 papers. Due to the late discovery of the conflicts, we decided that a conflict of interest policy should be put in place that required conflicted members recuse themselves from discussions of the papers for which they had conflicts.

6. The Best Paper committee chair then created the following process:

   a. For each paper each member was asked to notify the chair whether he/she had a conflict or not.

   b. Consider short papers first. Each member ranked the papers (abstaining on papers with a conflict), and then based on the outcome, further discussions were held without the conflicted members.

   c. Consider the full papers. Each member ranked the papers (or abstaining), and then based on the outcome, further discussions were held without the conflicted members.

Although we identified nominations for best papers and a committee well in advance of the final submission of camera ready papers, we did not share the details of those papers (e.g., authors) with the best paper committee until after the camera ready papers came in. When it was discovered that some of the members had a conflict-of-interest with one or two papers, we elected to work with the chair of the Best Paper Committee Chair to develop a conflict of interest policy rather than to replace the conflicted committee members due to the limited

time remaining for reading the papers and deliberation. Members with conflicts did not vote or express an opinion on any paper they had a conflict with.

The overall process could be improved in the future with some common-sense modifications. Best papers must be sent out in a timely manner to give reviewers plenty of time. It is critical to provide more information about the papers to be considered as the committee is being recruited and assembled (especially with regard to area expertise and conflict-of-interest), or by recruiting a larger pool of members with some held in reserve in the event of a conflict-of- interest (Eugene Charniak's recommendation, seconded by Aravind Joshi). To increase the pool of Best Paper candidates, Eugene Charniak suggested asking reviewers to indicate which papers are significantly better than the other papers they reviewed for the conference instead of nominating best paper candidates, something which reviewers (and in some cases area chairs) seemed reluctant to do.

The Best Paper committee also had a lively discussion about the overall Best Paper approach. Two contrasting ideas were discussed, both of which would require a policy change at the NAACL Board level. The chair of the best paper committee (Aravind Joshi) recommends that we consider having attendees vote online after the presentations are made, with some adjustment for conflicts of interest with one's institution. The rationale is: "After all, we go to the conferences to "hear" people talk about their work. Our current system of judging the papers only on the basis of the written versions is a poor version of evaluating the papers for a journal, as there is not enough time and also the authors have no chance of rebutting some basic misunderstandings. Such misunderstandings can possibly be corrected after an oral presentation during the Q/A period." In contrast, another committee member (Daniel Marcu) recommends dropping the best paper award altogether and instead give an "a most influential paper award" that looks back over papers from the past 5 or 10 years. In light of this discussion, another committee member (Drago Radev) pointed out the following resource which could be helpful in making decisions about future Best Paper awards: "I have raw data (collected with Bonnie Webber and Ali Hakim) on the most cited papers in the ACL Anthology in any given year. It is possible to intersect this list with the best paper awards in these years. I looked at a few years already and the ACL best paper in these years was always one of the most highly ranked ones by citations." Drago indicates that this is a small sample, but it is nonetheless, a reasonable way to get started.

One important rationale for having best papers in our conference is as an example of premier work for our students. We believe that the Best Paper committee worked well this year; the discussion was solid and they selected two solid and interesting papers. This type of award should not prevent us from having awards for past papers with impact. This would be an award administered by ACL and could involve both conference and journal papers. IEEE has this type of award, as well as "Fellow of the IEEE" based on the work done in one's career. These are important awards that make our promotion and tenure letters more effective. We have grown to the point where perhaps we should be doing more to celebrate our good work, not less.

One final concern was raised by Aravind concerning the high numbers of papers being submitted to the ACL conferences, and the impact not only on best paper decisions but also on the overall quality of reviewing. "There are other problems with our meetings that make the "Best X" kind of selection even harder. The total number of submissions at ACL is close to 1000, I am told. I don't know what the number is for NAACL/HLT. For COLING the number is over 860. The degree of arbitrariness in the selection of papers (long and short) is increasing rapidly. Many rejected papers at ACL, I was told, have only two reviews, and so on. The powers that be need to look at this situation and think of alternate, perhaps more innovative ways for dealing with this situation. Arguments such as our major conferences are just as good (or even better than) the journal papers is losing force with our colleagues from other disciplines. This, of course, concerns those who are in the academic world more than others."

Committee Communication[2]. For future conferences, we believe it is vital to identify and communicate early in the process, the major tasks of the program chairs, as well as for people in the other roles, so that areas of

---

[2] See Appendix D for a list of suggested process improvements

overlapping responsibility are recognized and the needed collaboration occurs. Below is a list of activities that we found fundamental to our program chair role. To help program committees in subsequent years get a head start on assembling a top notch program, we suggest developing a hand-off process from one program committee to the next. This could be in the form of a conference call or actual meeting when the new process begins. The next set of program chairs should also be encouraged to contact the past chairs for advice. Perhaps a kick-off meeting should be held where the PC co-chairs and General Chair meet each other, and work together to develop a full list of tasks and dates so that there is a common understanding of roles and responsibilities. This is also a helpful link: http://www.cs.jhu.edu/~jason/advice/how-to-chair-a-conference.html. Below is a list of the most important tasks from this year:

1. Develop "Important Dates Schedule"
2. Define areas
3. Identify START Manager (either a Program Chair or perhaps, an additional person who can manage this exclusively to lighten/support the PC workload)
   - Coordinate with Rich Gerber of Softconf
4. Recruit of Area Chairs
5. Write and Post Preliminary and Final Call for Papers (CFPs)
   - See the 2010 CFP at http://naaclhlt2010.isi.edu/cfp.html
   - Send CFP and other materials to local arrangements chair to post on the website
   - Send CFP to Priscilla to post on the ACL listserv; send also to other appropriate lists
   - Add information concerning the fact that the ACL anthology will be updated to contain the conference proceedings two weeks in advance of the first day of the conference.
6. Learn about START
7. Schedule and run meetings to make decisions with area chairs about acceptances
8. Support Special Lunch Panels (e.g., Industry Panels)
   - Recruit Organizer to pick theme and organize
   - Offer general support to organizer
9. Communicate acceptance notifications
10. Field concerns of the recipients of the notifications
11. Recruit and Support Keynote Speakers (e.g., this year, a speaker asked for a piano)
12. Conference Best Papers:
    - Organization and Support of Best Paper Committee
    - Communicate Best Paper Award Notifications (winners and other nominees)
    - **IBM Best Student Paper Award**: IBM has been giving a yearly award for the Best Student Paper at NAACL (and ACL). The student receives a monetary award.   The student can publish with other non-student authors, but has to be first author, and has to have been a student the entire time that s/he worked on the paper, and at the time of submission. There is a provision that the student can be eligible up to one year after finishing his or her degree. The Best Paper Committee should select this paper in tandem with the Best Full and Short Papers. In 2009 and 2010, this award was given to one of the two Best Papers. However, when Best Paper candidate papers are selected, if there are no student papers in the mix, a Best Student paper should be selected. The logistics around this should be to be considered by the PC. That said, in the START system, the PC might want to have first authors enter if they are students or non-students. This could also support the process of selecting a Best Student paper, in the event that no Best Paper candidates are students. If the student paper is not the Best Full or Short Paper, time should be allocated so the paper can be presented at the Best Paper Session.

13. Develop conference program (with feedback from Local Arrangement Chairs and Priscilla)
14. Arrange for Session Chairs
15. Work with Local Arrangements and Publication chairs to make sure the technical program is correctly represented in the conference book and the proceedings
16. Organize Poster and Demo Session (in collaboration with Demo Chair, Local Arrangement Chairs, and Priscilla)
17. Write "Program Chair Preface" for the proceedings, and the Final Conference Report for the NAACL and HLT boards.
18. Prepare Business Meeting Slides
19. Prepare Presentation for One-Minute Madness (if this session is to be held)
20. Attend conference, and support conference on-going activities (e.g., introduction of Keynotes, Business Meeting, etc.)

We have included as much information as we can think of to help next year's Program Committee, and perhaps, also, the NAACL Board in any new policy setting. Please feel free to get in touch with us if you have any questions.

Jill Burstein,   Educational Testing Service (jburstein@ets.org)
Mary Harper,   University of Maryland & Johns Hopkins HLT COE (mpharper@umd.edu)
Gerald Penn,   University of Toronto (gpenn@cs.toronto.edu)

Appendix A: Full and Short Paper Submission Statistics by Area

Area	# Long Submissions	# Long Accepts	Long Acceptance Rate	Accepted But Withdrew	Withdrew After Review Reject	Withdrew Before Review
Acoustic Models Speech	3	1	0.333	0	0	0
Dialog	6	3	0.500	0	0	0
Discourse	13	4	0.308	0	0	0
Generation	9	3	0.333	0	1	0
Grammar Engineering	6	2	0.333	0	0	0
Information Retrieval and Extraction	38	9	0.243	0	1	1
Language Models	6	2	0.333	0	1	0
Machine Learning (Speech)	7	4	0.571	0	0	0
Machine Learning (Text)	33	7	0.219	0	0	1
Machine Translation	48	18	0.383	0	4	1
Mathematical Linguistics	1	0	0.000	0	0	0
Morphology Phonology	11	5	0.455	0	0	0
Parsing	27	9	0.333	0	0	0
Semantics	35	12	0.343	0	0	0
Sentiment	21	4	0.190	0	0	0
Summarization	19	5	0.263	0	1	0
Word Sense Disambiguation	11	2	0.182	0	0	0
TOTALS	**294**	**90**	**0.309**	**0**	**8**	**3**

Area	# Short Submissions	# Short Accepts	Short Acceptance Rate	Accepted But Withdrew	Withdrew After Review Reject	Withdrew Before Review
Acoustic Models Speech	4	0	0.000	0	0	0
Dialog	8	4	0.500	0	1	0
Discourse	6	2	0.333	0	0	0
Generation	5	1	0.200	0	0	0
Grammar Engineering	2	1	0.500	0	0	0
Information Retrieval and Extraction	15	6	0.429	0	2	1
Language Models	4	1	0.250	0	0	0
Machine Learning (Speech)	7	3	0.429	0	1	0
Machine Learning (Text)	8	3	0.429	0	0	1
Machine Translation	31	8	0.258	0	1	0
Mathematical Linguistics	0	0	0.000	0	0	0
Morphology Phonology	9	3	0.333	0	0	0
Parsing	20	9	0.450	0	1	0
Semantics	16	5	0.313	0	0	0
Sentiment	12	4	0.333	0	0	0
Summarization	8	4	0.500	1	0	0
Word Sense Disambiguation	6	2	0.333	0	0	0
TOTALS	**161**	**56**	**0.352**	**1**	**6**	**2**

Table 2. Long and short paper statistics

Appendix B: Submissions by First Author Country

Argentina	2
Australia	15
Austria	1
Belgium	2
Brazil	1
Canada	13
China	21
Cuba	1
Czech Republic	1
Denmark	1
Egypt	1
France	5
Germany	30
Greece	1
Hong Kong Special Administrative Region of China	4
India	24
Ireland	6
Israel	2
Italy	9
Japan	13
Netherlands	3
New Zealand	1
Norway	1
Qatar	1
Republic of Korea	2
Singapore	6
Spain	11
Sweden	1
Switzerland	3
Taiwan	2
Turkey	1
United Kingdom	30
United States	239
Viet Nam	1

Table 3. Country statistics for the main contact authors of submitted papers.

Appendix C: Area Chair Letter

Dear Area Chairs,

We would like to thank you for agreeing to be area chairs for the NAACL-HLT 2010 conference to be held in Los Angeles from June 1-June 6th, 2010. This year, as in previous years, there will be short papers submitted in addition to long papers, but with the same deadline. The time frame for this year is as follows:

Dec 1 (Tues)	Long and short papers due
December 9 (Tues)	**Receive papers in your area**
December 18 (Fri)	Paper reviewer assignments completed
Jan 11 (Mon)	Reviews due back to area chairs
Jan 12 (Tues) – Jan 23 (Fri)	Full review discussion period

Review discussion period details:

Jan 12 (Tues)	Send reviews to authors for the **author response period**
Jan 14 (Thurs)	Author responses due back to area chairs
Jan 12 (Tues) – Jan 15 (Fri)	**Discussion of reviews with your reviewers**
Jan 18 (Mon) – Jan 21 (Thurs)	Area chairs teleconference with PC co-chairs to make final decisions.
Jan 21 (Thurs)	Area chair decisions due to Jill, Mary and Gerald
Jan 23 (Sat)	Final PC decisions made

Jan 25 (Mon):	Acceptance notifications sent
Mar 31 (Wed)	Final, camera-ready papers due

Please note that there will be an **author response period** this year. Author responses will not be sent to reviewers, but will be used along with the reviews to you as assist area chairs in making acceptance decisions. Note that discussion of review will take place during author response period.

Your expertise and assistance is vital for helping to ensure the quality of the papers in this conference. Your oversight is vital at several different points in time:

1. **You will be called upon to recruit a pool of reviewers.** We attach a letter that you can use for your recruiting effort. It will be important to recruit sufficient high quality reviewers that you will not have to overload them. Note that each paper requires three reviews. If your area receives 30 papers and each has 3 reviews that means you have 90 review events. If you have only 10 reviewers, that will heavily load the reviewer, and could lead to poorer quality reviews. As a rule of thumb, each reviewer should be expected to review a total of no more than 7 submissions (a typical load might be 5 long papers and 2 short papers, or 4 long papers and 3 short). Please try to recruit all your reviewers **by November 15, 2009**. It is easier to get people to say "yes" well before the papers come in! We attach a pdf file with the number of papers assigned to the areas from last year. This year we have a slightly different area distribution, so you may have to use your best judgment

   Note that we would like you to send us the list of reviewers in the following comma separated format so we can check for overlaps: First Name,Last Name,Institution,email
   That way we can let you know of overlaps so that the multi-area reviewer is not overloaded (and you will know to recruit an additional reviewer if needed). **Please send us your list as soon as possible, so we can identify overlaps and let you know.**

2. **You will need to assign papers to reviewers based on your knowledge of your reviewer pool.** You will receive papers in your area by December 1[4th]. You should complete the paper assignment process no later than December 18th, 2009. Prior to this, we will send you a template for the review form by

November 1[6th], 2009. There will be a very short 1-2 day period of time for addressing mis-assignments. Please note that reviewing will be blind and anonymous and will be supported by the START system.

3. **You will need to ensure that all of the reviews come in for papers in your area.** Using the START system, you will be monitor the state of the reviews, remind reviewers of upcoming deadlines, and "nag" those reviewers who are not delivering as promised. Note that there will be times when a reviewer will be unable to deliver for various reasons. It is better to over-recruit as a buffer, but inevitably you will need to have reviewers in the wings who can help out to ensure that three reviews are provided. It may be that you as area chairs will have to step in and review a paper or two. You might also consider recruiting 1-2 reserve reviewers who will pinch hit for you. **Reviews are due January 11, 2010, and reminders should be sent early in January and then as needed when the final due date comes closer.**

4. **You will facilitate discussion to ensure a fair review process.** Inevitably reviewers will have mixed evaluations for some papers. There are two tools available to you to help with this resolution. We will disseminate blind reviews to authors to elicit their responses, and there should be a discussion among reviewers to help resolve incompatible reviews. In the end, you will need to be the final arbiter for mixed-review papers. **January 12-15, 2010 is set aside as the time for review discussion and to elicit author responses.** Again the START system will play an important role in this effort.

5. **Conflict of interest**. We are currently looking into how conflict of interest issues (e.g., you or your student will plan to submit to your area, or another area) will be handled in START. Stay tuned.

6. **You will need to meet with us by teleconference during the week of January 18th (probably in 2-3 installments).** The program committee (pc co-chairs + area chairs) will make acceptance recommendations about papers, taking into consideration the author responses and the reviews. This will be a busy week, and it will be essential that all area chairs are available during this week.

Thanks again for agreeing to serve as area chairs, and we look forward to working with you on making NAACL HLT 2010 a success.

Best,
Jill, Mary, and Gerald

Appendix D: Suggested PC Process Improvements

PC requirements:

Since the PC co-chairs do not meet regularly in person, all communication is handled remotely, and co-chairs are selected without regard to individual working styles, it is critical that co-chairs agree a priori to communications guidelines to ensure efficiency and fair delegation of responsibilities, including (but not limited to) the following.

1) Regular meetings of the co-chairs with explicit action items
2) Clear partitioning of the jobs
3) A commitment to checking email (frequency to be determined by Program Committee, but 3 times a day over the course of a day would be the suggestion) if they expect to be included in all decision making, and especially during timely events (e.g., sending out acceptance notifications, responses to authors)
4) Understanding that there are critical tasks that must be addressed on particular days, e.g., monitoring the submission queue and handling email issues related to submissions as they come in. These tasks require significant personal attention.

Suggestions for future PC co-chairs:

1) Create an email account for email traffic that the PC chairs can monitor together, with auto forward to respective email addresses with a keyword that enables filtering.
2) Look into getting program creation software. It is really hard to put together a program that is sensible and make sure all papers are accounted for without that. This is especially difficult to coordinate via remote communication (phone and e-mail), and is very tedious and easily prone to error.
3) Identify up front all program constraints (e.g., events that outside the papers, such as demos, posters, student workshops, etc) so that there are no missteps in planning the program.
4) All of the publication chair work needs to be done in tandem with the program chair work.
   a) Paper guidelines must be updated and posted before the final call for papers
   b) Camera-ready guidelines and submission links need to be updated and posted soon after the acceptance letters go out so that when we inform the authors of poster or oral presentation, that information is already in place. It is also be important to put up guidelines for presentation and poster preparation on the web at that time, so that everything can go out in one letter.
   c) The hand off of the program is to **both** the publication chairs and the local arrangements chair in charge of the conference handbook. There is the need for a three-way communication to ensure consistency. We needed to discover this, but PC members should understand this up front.