

# Evaluation of CLCG Research, Summer 2004

Jan van Eijck, Susan Herring, Wolfgang Klein, Stephen Pulman  
with editorial support from Robbert Jan Bron

Summer 2004

# Contents

<b>1</b>	<b>General Information about the Assessment</b>	<b>3</b>
1.1	Context of the Assessment . . . . .	3
1.1.1	National System for Assessing the Quality of Research . . . . .	3
1.1.2	Outline of RUG Protocol . . . . .	4
1.2	Scope of the Assessment, Assignment to the Committee . . . . .	5
1.3	The Review Committee . . . . .	6
1.3.1	Composition . . . . .	6
1.3.2	Distribution of Tasks . . . . .	7
1.4	Input for the Research Assessment . . . . .	7
1.5	Working Procedure of the Committee . . . . .	8
1.6	The Site Visit . . . . .	8
<b>2</b>	<b>CLCG as a Whole</b>	<b>10</b>
2.1	Mission and Goals . . . . .	10
2.2	Self-Evaluation . . . . .	10
2.3	Leadership . . . . .	11
2.4	Overall Shape of the Centre . . . . .	11
2.5	Position of CLCG within the Department . . . . .	12
2.6	Balance of Strengths and Weaknesses . . . . .	12
2.7	Facilities . . . . .	14
2.8	Graduate Studies and Selection of PhD Projects . . . . .	14
2.9	Overall Assessment of the Institute . . . . .	15
<b>3</b>	<b>The Research Programmes</b>	<b>16</b>
3.1	Phonetics and Phonology . . . . .	16

3.2	Syntax and Semantics . . . . .	18
3.3	Discourse and Communication . . . . .	20
3.4	Descriptive and Historical Linguistics . . . . .	22
3.5	Computational Linguistics . . . . .	24
3.6	Educational Linguistics . . . . .	25
3.7	Neurolinguistics and Experimental Linguistics . . . . .	26
<b>A</b>	<b>Curricula Vitae of Members of Evaluation Committee</b>	<b>28</b>
<b>B</b>	<b>Declaration of Independence</b>	<b>30</b>
<b>C</b>	<b>Extended Description of the 5 point scale</b>	<b>34</b>
<b>D</b>	<b>Checklists for the Research Evaluation</b>	<b>35</b>
<b>E</b>	<b>Schedule of the Site Visit</b>	<b>38</b>

# Chapter 1

## General Information about the Assessment

### 1.1 Context of the Assessment

#### 1.1.1 National System for Assessing the Quality of Research

In 2003, the Dutch system for assessing the quality of research underwent a major change. The system of national, external assessments of individual disciplines, co-ordinated by the office of the Association of Dutch Universities (VSNU), was discontinued. In its place, the Executive Boards of the universities now determine the design and organisation of the research quality evaluations. They are bound by the “Standard Evaluation Protocol 2003-2009” (SEP), which is endorsed not only by VSNU but also by the Netherlands Organisation for Scientific Research (NWO) and the Royal Netherlands Academy of Arts and Sciences (KNAW).

The three main aims of the Standard Evaluation Protocol are improving the quality of research, improving research management and direction and improving accountability, both internal (by the unit to be assessed to its immediate superiors within the RUG) and external (by the RUG to government and society). The SEP is concerned with two main procedures:

- an external assessment once every six years (by a peer review committee conducting a site visit)
- a self-evaluation once every three years (one in preparation for the external assessment and one intermediate evaluation three years later, the ‘mid-term review’).

The most important conclusions of the external assessment committee, the reaction to these by the assessed unit and the final conclusions with regard to the future applied to

them by the Executive Board will all be published. An independent meta-committee, set up by KNAW, NWO and VSNU, will check the design and implementation of the new system by the various institutions and publish its findings annually.

### 1.1.2 Outline of RUG Protocol

The SEP provides a framework to guarantee - as far as possible - comparable procedures and criteria. Within this, it provides room for specific input by the own institution, which the RUG has set out in the “Protocol for Quality Assurance at the University of Groningen”, known as the RUG Protocol.

The following principles inform the RUG protocol:

1. close connection with the RUG quality policy
2. clear division of tasks and responsibilities
3. external assessment is transparent, authoritative and can be applied to both internal policy and external accounting
4. the aim is professionalising and minimum disruption for researchers.

**Re 1 RUG Quality policy with regard to research** The heart of this policy is that the RUG regards quality improvement as the dominant principle in its research policy as well as in that of the government. A crucial part is played by the peer reviews, external assessments by independent, objective researchers with expertise in the disciplines of the unit to be assessed. The peer reviewers should preferably be recognised international authorities and base their assessment not only on the self-evaluation of the unit but also on actual knowledge of the most important output, where possible supplemented by quantitative and qualitative indicators. Further, external research assessments should concentrate on:

- providing direct, swift feedback from the peer reviewers about the position of the research, measured against national and international standards for quality, productivity, relevance and vitality
- assessing both past performance and future expectations, the ambitions and the scientific and social impact of the research
- evaluating the management and the academic leadership of the unit in relation to the mission and ambitions
- the context of the research unit, for example how the unit is embedded in the department, the university as a whole, the national and international context, as well as disciplinary and interdisciplinary contacts with regard to content.

Before formal acceptance of the findings of the peer review committee as laid down in the assessment report, the Executive Board of the university will apply the principle of hearing both sides of the case.

**Re 3 Usability** The results of an assessment must be sufficiently informative to serve as the basis for policy decisions. This is why the possibility of adding a lower aggregation level than that of the programme as referred to in the SEP is deliberately left open. In practice, the aggregation levels of research programmes may vary strongly. If the Executive Board of the university believes that a research programme is too large for an evaluation at that level to be sufficiently usable for internal policy decisions, a supplementary evaluation at a lower aggregation level will be requested from the Faculty board. The external assessment at this lowest level can, if desired, remain confidential. The SEP provides for this eventuality in the management letter: ‘Matters of personnel policy and sensitive decisions are generally treated in the confidential management letter to the board and do not form part of the public report.’

**Re 4 Minimum assessment disruption** All institutes at the RUG are organised on a disciplinary and local level. Within the previous national system, an institute was assessed simultaneously with comparable research groups at other Dutch universities. In the current system national, disciplinary visitation is no longer compulsory but still an option, provided that the relevant Executive Boards approve. The RUG is determined to keep the option for national co-operation open, particularly because of the increased comparability of the assessments and the more efficient use of peer reviewers. An alternative for national co-operation would be to allow a single PRC to assess several Groningen institutes. This option is offered to departments aiming to cluster their multidisciplinary research institutes.

## 1.2 Scope of the Assessment, Assignment to the Committee

For the assessment of the research organised in the Centre for Language and Cognition Groningen (CLCG) an international peer review committee (PRC) was appointed by the the Executive Board of the University of Groningen. The PRC used the CLCG self-evaluation report 2003-2009 as a starting-point for their assessment.

Prior to the site visit the committee received the following documentation:

- the SEP and a summary thereof
- the RUG Protocol

- the self-evaluation report of CLCG (cf. SEP format, including a SWOT analysis, tables with input and output at institute and programme levels, publication lists and full text copies of key publications)

Within the Centre for Language and Cognition assessment, the following seven research programs were presented to the PRC for evaluation:

1. Phonetics and Phonology
2. Syntax and Semantics
3. Discourse and Communication
4. Descriptive and Historical Linguistics
5. Computational Linguistics
6. Educational Linguistics
7. Neurolinguistics & Experimental Linguistics

## 1.3 The Review Committee

### 1.3.1 Composition

The peer review committee members were appointed after a thorough selection procedure (see: Appendix B) assuring an authoritative, critical and independent assessment of the quality of the research institute. The University Board appointed Professor Henk Verkuil as the chairman of the committee. In consultation with the chairman the other members of the committee were selected. Due to illness Professor Verkuil was unable to fulfil his task. The Board of the university then asked Professor Jan van Eijck to take over the chair. Professor van Eijck accepted this invitation on very short notice.

The affiliations of the committee members:

- Prof. Jan van Eijck, Centre for Mathematics and Computer Science (CWI), Amsterdam, Universities of Utrecht and Amsterdam, Netherlands (chair)  
<http://www.cwi.nl/~jve>
- Prof. Susan C. Herring, Indiana University, Bloomington, USA  
<http://www.slis.indiana.edu/faculty/herring/>
- Prof. Wolfgang Klein, Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands  
<http://www.mpi.nl/world/persons/profession/klein.html>

- Prof. Stephen Pulman, University of Oxford, UK.  
<http://www.ling-phil.ox.ac.uk/people/staff/pulman/>

Robbert Jan Bron, MA, Department of Academic Affairs, University of Groningen, was appointed as secretary to the committee.

All members of the Committee signed a declaration and disclosure form to safeguard that:

- the panel members judge without bias, personal preference or personal interest,
- the judgement is made without undue influence from the institute, the programme or other stakeholders.

More information about the committee members and their curricula vitae can be found in appendix A.

### 1.3.2 Distribution of Tasks

Before the start of the site visit, the following distribution of tasks was agreed upon:

- Jan van Eijck: Assessment of *Syntax and Semantics* group, focus on research links to the outside world (collaboration with other departments, collaboration outside Groningen, international links).
- Susan Herring: Assessment of *Discourse and Communication* and *Descriptive and Historical Linguistics* groups, focus on graduate studies.
- Wolfgang Klein: Assessment of *Neurolinguistics* and *Language Acquisition* groups
- Stephen Pulman: Assessment of *Computational Linguistics* and *Phonetics and Phonology* groups.

## 1.4 Input for the Research Assessment

The input for the research assessment consisted of a self-evaluation describing the research conducted by the Centre for Language and Cognition Groningen over the period 1998-2003. It reports the institute’s mission, describes the research programmes, gives an overview of input in terms of research time and money, an overview of scientific output in different forms, an analysis of the institute’s impact both in and outside scientific circles and, finally, its plans for the future. The report was written in accordance with the directives of the “Standard Evaluation Protocol 2003-2009 For Public Organisations” (SEP) and contains two main parts. The first part comprises information on research at the

aggregation level of the institute; in the second part the individual research programmes are described in detail.

Together with the site visit (including interviews with CLCG staff), the self-evaluation report forms the basis for the assessment of research quality by the peer review committee.

## 1.5 Working Procedure of the Committee

The SEP requires the committee members to assess each of the research programs on four main aspects, namely:

- Quality (international recognition and innovative potential)
- Productivity (scientific output)
- Relevance (scientific and socio-economic impact)
- Vitality and feasibility (flexibility, management, and leadership)

The scores used within the assessment are: excellent (5), very good (4), good (3), satisfactory (2), and unsatisfactory (1). It should be noted that these ratings, which are specified in the SEP, differ from the ratings specified in the previously applied VSNU-Protocol (valid until 2003). The ‘old’ VSNU-protocol ratings were: excellent (5), good (4), satisfactory (3), unsatisfactory (2) and poor (1). From 2003 on, a new rating ‘very good’ was added between ‘excellent’ and ‘good’, while the rating ‘poor’ was omitted. When comparing scores this must be taken into account very carefully!

## 1.6 The Site Visit

Due to circumstances beyond his control, Professor Pulman was, at the last possible moment, prevented from leaving the UK to take part in the site visit. Fortunately, his written assessment was available to the other committee members. They used this material to corroborate the committee’s findings.

The site visit was conducted by the three remaining committee members and the committee secretary. A full schedule of the meetings is given in Appendix E. The first day was devoted to discussions with the dean of the faculty, Ger de Haan, with the Centre Director John Nerbonne, and with the various program leaders. Material for discussion was provided by previously distributed material, plus the short presentations that preceded each discussion.

During the second day of the site visit, the committee had the opportunity to talk with graduate students, with postdocs, and with the director of graduate studies, Jack Hoeksema. Also, on the second day of the site visit, Van Eijck, Herring and Klein conducted

an extended telephone conference with Pulman for further consultation and discussion. During this conversation it became clear that the conclusions reached by the other three committee members were in complete agreement with Pulman's assessment.

The two-day site visit was concluded with an oral presentation of the preliminary findings of the committee to the scientific director and the dean of the faculty of humanities, in the presence of the programme leaders.

# Chapter 2

## CLCG as a Whole

### 2.1 Mission and Goals

The mission of the Centre for Language and Cognition Groningen (CLCG) is to promote linguistic research in the department of humanities. The position of the Centre brings special challenges, opportunities and responsibilities in the following four areas:

- CLCG recognizes that linguistic research in Groningen must keep in step with instructional needs of the department in the modern European languages;
- CLCG has taken on a special responsibility for Dutch language and dialects, plus the minority languages spoken in the North of the Netherlands, that is to say Frisian and Lower Saxon;
- CLCG supports both pure and applied research. CLCG recognises that opportunities for applied research are important for educational linguistics, aphasiology (neurolinguistics), discourse and communication and computational linguistics.
- CLCG is in a unique position to advance linguistic research in cooperation with cognitive neuroscience in Groningen. This opportunity presents itself mainly in the areas of neurolinguistics and computational linguistics.

### 2.2 Self-Evaluation

The committee was impressed by the clarity and honesty of the self-evaluation report provided by the Centre. The report provides a frank discussion of strengths and weaknesses of the institute, and creates a clear impression that those responsible for the future of the Centre are eager to benefit from any new insights the assessment might provide.

## 2.3 Leadership

The committee has praise for the strong intellectual leadership provided by the current director, Professor Nerbonne. It is clear that the director is an important factor in stimulating the open and creative atmosphere that has led to a very fair self-evaluation. It is also clear that he attempts to support and stimulate all programmes of the Centre, that he has a keen eye for current strengths and weaknesses, and that he is willing to provide every help in seeing to it that these weaknesses are overcome. He takes an active interest in where research is going in each of the programmes, and he is a stimulating factor in the acquisition of external funding in the various areas.

## 2.4 Overall Shape of the Centre

All in all CLCG seems to be in good shape. On average, the research staff are productive, and their output is of high standard, although individual variations in quality and productivity are considerable. There are good future prospects. The research is innovative, in particular in computational linguistics. Graduate students and post docs seem happy, facilities seem adequate, and there is considerable success in acquiring external funding.

The main challenges faced at the institute level are by no means unique to Groningen. Here is a list:

1. how to maintain research impetus under the heavy teaching loads that are typical for humanities departments;
2. how to get funding for enough graduate students and postdoctoral researchers to maintain ‘critical mass’ in each research area, in an increasingly unpredictable funding climate;
3. how to attract high quality PhDs in a humanities setting where research is not a main focus of the curriculum;
4. how to devise and stick to a long term hiring policy that is geared towards research, in a climate where teaching needs tend to be the decisive factor in hiring decisions.

Recommendations:

1. CLCG should continue its matching policy for externally funded projects.
2. CLCG should consider establishing an external advisory board, as a means for keeping more firmly in touch with developments outside Groningen University.

## 2.5 Position of CLCG within the Department

CLCG is one of three research centres (institutes) within the humanities department, the other two being Archeology (much smaller than CLCG) and Cultural Studies (much larger than CLCG). Traditional language studies do poorly these days, in terms of the numbers of students they attract. The dean of humanities expressed some concern about the balance between the three institutes, and would like to see CLCG attract more students. Ideally, all programmes within CLCG should attract enough students to justify their staff allocations (at present, the CLCG programmes vary enormously in the numbers of students they attract). There is no concern about overall student numbers in humanities. The dean: ‘We have enough students.’

According to the dean, research priorities in CLCG are in neurolinguistics, computational linguistics and semantics, but it is unclear whether the department has the means or the willingness for extra investment in these areas.

In the setting of the humanities faculty, officially allotted research time for professors and associate professors (UHDs) is 40 per cent, for assistant professors (UDs) it is 30 per cent. Research expectations are rather low: one published paper of any kind per year is enough to nominally fulfil one’s research obligations. In specific cases, extra research time can be allotted, or research time taken away, on the basis of past research performance.

If one compares CLCG with one of its sister institutes in the Netherlands, ILLC (Institute for Logic, Language and Computation) at the University of Amsterdam, then it is striking that CLCG is part and parcel of a single department while membership of ILLC is much broader. ILLC brings together linguists and philosophers from the humanities department, logicians and computer scientists from the science department, and artificial intelligence researchers from the social science department. It is unfortunate that institutional links with computer science, philosophy and artificial intelligence are so much weaker in Groningen, and it would certainly benefit CLCG if these connections could be strengthened.

## 2.6 Balance of Strengths and Weaknesses

The main strength of the Centre seems to be in computational linguistics. This is a particularly strong area, with great productivity and vitality. Neurolinguistics and experimental linguistics has been particularly successful in securing grant funding. Syntax and semantics has a particularly glorious past, and the future of educational linguistics looks particularly promising. This much for strengths.

According to the committee, the main apparent weaknesses (and the challenges they present) are in the following areas.

1. The Phonetics and Phonology group does not have a sufficiently large overlap of

interests among its members to follow a coherent research programme of its own. Recommendation: the work carried out in the Phonetics and Phonology programme should be incorporated in other programmes (possibly in computational linguistics, but other options are open).

2. The language acquisition work seems out of place in the programme Neurolinguistics and Experimental Linguistics. Recommendation: Part of this work (Behrens) should be incorporated in the Educational Linguistics programme, another part (Van Hout and Hollebrandse) might fit into Syntax and Semantics.
3. The Descriptive and Historical Linguistics programme is in need of more cohesion. The work here suffers from the somewhat individualistic and isolationist attitude of the people involved. If these researchers look into possibilities for joint work, this will no doubt improve their profile and will help them to compete for funding.
4. The Discourse and Communication programme has suffered from a lack of productivity, partly due to huge teaching loads. Also, there have been serious management problems partly caused by prolonged illnesses of some staff members. It is clear that this group has gone through hard times. Recommendation: since there are two tenured positions to fulfil, the committee advises the Centre to appoint an associate professor with great organisational and management skills, with the specific task of assisting the programme leader in matters of management. The committee approves of the plan to let hiring policy reflect the wish to branch into computer mediated communication.
5. The Syntax and Semantics programme has been very productive in the past. The future looks less bright, however. The syntax work runs the risk of becoming isolated, because the Generative Grammar paradigm is no longer as dominant as it once perhaps was, and because the connections with the computationally oriented work in the Computational Linguistics programme are not strong enough. No attempts at large scale grammar development are being pursued, and the range of syntactic phenomena researched is rather narrow and has hardly changed over the years. It is not clear to the committee whether the programme leadership is fully aware of the dangers of isolation. The work in semantics, traditionally one of the main strengths of Groningen, is becoming increasingly thin: the links to logic and philosophy have disappeared (largely due to external factors), and many of the researchers in this area have dropped out, for a variety of reasons.
6. The Computational Linguistics programme is very strong, but the tenured staff is small. The programme might get in serious trouble if one of its tenured staff decides to accept an offer from elsewhere.

According to the committee, all of these weaknesses can be repaired, and it is our hope that as a result of this review they will be addressed.

Below are three recommendations for name changes:

1. ‘Educational Linguistics’ reflects teaching involvement rather than research orientation. It is recommended to change this into ‘Language development over the life span’, a title that seems to reflect the focus of the whole group quite well.
2. ‘Experimental Linguistics’ is a misleading title for the research carried out by Stowe and her co-workers. ‘Language in the Brain’ or ‘Language and the Brain’ would seem to be a more adequate title.
3. If the researchers in Descriptive and Historical Linguistics take the committee’s recommendations to heart, then a name change might be in order here too. ‘Language Variation and Change’ might be a good common denominator for this group after revitalisation.

Finally a general recommendation: there seems to be a clear correlation between group vitality and institutionalised group interaction in the form of regular meetings. Not all groups have regular meetings, but all excellent groups do. Group meetings are essential for creating and maintaining an intellectually stimulating academic atmosphere, and for nurturing common research efforts. Regular group meetings should be part of the routine in all groups.

## 2.7 Facilities

Facilities for CLCG seem to be adequate. This impression is confirmed by the stories of graduate students and post docs. Two points of concern remain:

1. Making available technical support in the field of statistics would benefit almost all of the research groups.
2. More adequate computational support seems needed for the computational linguistics group. In the context of a humanities department, adequate support in this area is always a potential weakness.

## 2.8 Graduate Studies and Selection of PhD Projects

The institute has developed a Research Master programme for talented students. While the focus on theoretical linguistics and neurolinguistics may have been necessary for getting official approval for this, now that this Master programme is in place it should be used as a preparation for *all* research programmes within the Centre. For this, enrolment will have to pick up from one research master student last year, and three lined up for the coming academic year, but also the programme will have to be broadened.

The Research Master programme is meant to prepare students for their postgraduate training, and with the present set-up there is a danger that future PhD students will only get proper training for two of the seven research groups.

We recommend redesigning the Master programme in such a way that it can accommodate excellent students from a wider variety of linguistic backgrounds. This can easily be done by making a wider selection of regular Master courses available for the Research Master programme. The committee has no problem with making some courses in theoretical linguistics and neurolinguistics obligatory for all students.

At present, PhD projects do not get divided evenly over the different groups. This need not be a problem, but the current system seems to have biases that some groups profit more from than others. The PhD project allocation system should be evaluated and — if necessary — revised.

## **2.9 Overall Assessment of the Institute**

### **Quality**

4

### **Productivity**

4

### **Relevance**

4

### **Vitality**

4

These grades were arrived at by taking the averages of the individual research programmes, and rounding off upwards. The justification for upwards rounding off is provided by the excellent way the Centre is directed.

# Chapter 3

## The Research Programmes

### 3.1 Phonetics and Phonology

Programme Leader: Dr. Dicky Gilbers. Programme leadership is on a provisional basis. Former programme leader Dr. Tjeerd de Graaf retired in January 2003.

**Assessment** On the basis of the review material and the presentation by Gilbers and Gooskens, it became clear to the committee that there is little group cohesion within this programme. Clearly, the retirement of De Graaf has been a severe loss to the group: it has left the group without a leader. Also, the two most active group members are pursuing research interests of their own, with no great overlap of interests. As the self-assessment has it, meetings are ‘infrequent,’ and there seems to be little agreement on a common research agenda and research strategy. This creates the impression of different individuals each following their own interest, with little communication. One group member even appears to have switched to purely literary interests.

Some of the topics that one would think of as the domain of this group are in fact also being pursued outside it (e.g., in neurolinguistics). The level of published output is fair, but the rate of completion of PhD theses seems very poor. The level of external funding is not commensurate with the size of the group. The committee feels that this should not remain a separate group. They should merge with computational linguistics (if a specialist in the field of speech processing were to be appointed a more coherent group would emerge), with the discourse and communication group, or wherever they would feel most at home intellectually.

#### Quality

3

**Productivity**

3

**Relevance**

2

**Vitality**

2

These grades reflect our verdict that this group has become too small and incoherent for survival as a separate programme.

## 3.2 Syntax and Semantics

Programme Leaders: Prof. Jan Koster and Prof. Jack Hoeksema. Former programme leaders are Prof. Co Vet and Prof. Alice ter Meulen

**Assessment** This is — or has been — a highly productive group. The quality of the work is high, in fact it is has been at the international forefront for several decades. Given the size of the tenured staff, the number of PhDs is perhaps a bit disappointing, but on the other hand the track record for on-time-completion is excellent.

The two subgroups ‘Syntax’ and ‘Semantics’ were merged in 2002. There is some interaction, but the impression persists that the merger was inspired by the fact that the two subgroups were becoming too small. As the self-assessment has it, “The subgroups still function rather independently.” This may be an understatement of the fact that there is less common ground than one would hope. There is some research interaction in the study of logical form and anaphora, but a common agenda of open issues in these areas, stated in a framework independent manner, seems to be missing. Conspicuously absent is collaboration with philosophers on semantics and philosophy of language, with logicians on inference, and with computer scientists (or computational linguists inside CLCG, for that matter) on formal grammar analysis and grammar complexity.

The committee has some concerns regarding the future. The research within the syntax group certainly has international visibility, but it is perhaps becoming somewhat esoteric, in the sense of drawing strongly on paradigms that used to be dominant, but that seem to be losing ground around the world. There is some danger that this may lead to isolation in the future. The fact that syntactic work in the computational linguistics group within CLCG is based on radically different assumptions is quite telling. The committee recommends that the group should put much more emphasis on the emerging cooperation with computational linguists and neurolinguists. We don’t recommend a framework change, but we would like to see some awareness of the fact that times are changing for Generative Grammar and Minimalism. What used to be a forest now is reduced to a collection of two or three oak trees, admittedly still impressive, but also maybe the last that are still standing.

Semantics has been one of the highlights of Groningen linguistics for several decades. In the glorious 1980s there were strong links with logic and philosophy of language. For various reasons these links have all but disappeared (Van Benthem left Groningen without leaving any logicians with an interest in natural language behind, Sanchez died, Ter Meulen opted for a management position on the board of NWO, Zwarts was appointed Vice Chancellor — Rector Magnificus — of the University, Hoeksema shifted his research focus). This means that, for all kinds of readily understandable reasons, the group has thinned considerably over the past few years. The research agenda has remained virtually unchanged for a long time, with polarity phenomena, quantification, and anaphora still figuring as key elements. If no new semanticists are appointed in the near future, if the

links with logic are not reestablished, and if the research agenda is not updated (e.g., by branching into common areas with computational linguistics), the committee sees a problem concerning the long term vitality of this research area.

**Quality**

5

**Productivity**

5

**Relevance**

3

**Vitality**

3

The somewhat lower marks for relevance and vitality are motivated by the fact that the group takes its relevance perhaps too much for granted. There is scope for enhanced relevance and vitality if the collaboration with computational linguistics and neurolinguistics is vigorously pursued. The overall average scores for the program are still high, and we do not wish to suggest that the Syntax and Semantics program should be excluded from consideration for additional funding.

### 3.3 Discourse and Communication

Programme Leader: Prof. Gisela Redeker

**Assessment** This is a relatively small programme in terms of its staff, with one full professor and eight assistant professors, and no (current) nontenured staff. Two funded PhD projects were completed in an average of 7.5 years, the smallest number of PhDs and the longest time to completion of any of the programmes in the CLCG. The programme also has the lowest research output, as measured in publications, although output is variable across the staff. Several of the staff (e.g., Redeker and Bax) are internationally visible.

This programme faces a number of challenges. Its subject matter is popular with students, thus teaching loads are heavy. It is somewhat surprising that the programme has not been given additional staff hires before now, if its courses are generating revenue for the department as reported. It seems particularly understaffed. It has also merged two different areas, discourse/pragmatics and media studies, and has yet to achieve a coherent synthesis. The programme's future projection to hire a specialist in computer-mediated communication would help to bridge the gap. This group has also suffered from few (and weak) PhD applicants, some under-productive faculty, and some that have had health and stress-related problems, although the latter two situations appear now to be resolving themselves. It is unclear to what extent these issues are interrelated, or related to the relatively low amount of external funding that the programme has attracted during the evaluation period.

The strengths of the programme, in addition to generating revenue through teaching upon which the institute depends, include

1. the strong societal relevance of its subject areas (especially new media such as the Internet and the World Wide Web, and organisational applications), and
2. its leadership, which appears to have aptly identified the programme's current weaknesses and articulated a vision which could capitalise nicely on the present relevance and interest of media studies, by creating a focus on multimedia and non-verbal modalities of discourse. This is a cutting-edge topic, and there is yet to be much competition from programmes in other universities in Europe or abroad, so CLCG could establish a ground-floor advantage. Such a focus could also enable collaboration with media and computer interface designers, which in turn could enhance the programme's broader impact and external fund raising capacity.

In sum, although this programme's past research performance has not been stellar, it has future potential. Its future viability and vitality will hinge on new hires who are productive researchers, a strong vision, and institutional support to protect research time from being consumed by teaching.

It is recommended that student projects should be used to bridge the gap between teaching and research. The nature of research in this area, plus the fact that there is a large reservoir of students to recruit from, makes this quite feasible.

**Quality**

4

**Productivity**

2

**Relevance**

5

**Vitality**

3

Highest marks for relevance, as reflected in the attraction to students. The high mark for quality is motivated by the fact that several of the staff are internationally visible. The low mark for productivity is a reflection of the past. Vitality of the research is quite uneven, and the mark is an average.

## 3.4 Descriptive and Historical Linguistics

Programme Leader: Dr. Peter Houtzagers

**Assessment** This programme has the second-largest number of tenured staff (after Syntax and Semantics), comprising four full professors, three associate professors, and six assistant professors. Although its ranks are declining, it continues to attract and support several PhD projects, and its PhD students publish their work, which is commendable. The staff are also active in publishing, some highly so, including a number of monographs. However, only Werner Abraham (who retired in 2002) publishes in international journals; most of the other publications are in conference proceedings, in-house publications, or specialist journals in other languages, which inevitably limits their accessibility and impact. Although a number of these researchers appear to be respected specialists in their fields, the fields are defined narrowly; moreover, a number of the researchers have been working on the same topics for years. This gives the impression of over-specialisation.

The retirement of Werner Abraham obviously has been a severe blow to this group.

The staff do not collaborate or meet to discuss their research, nor do they express an interest in doing so. They prefer to work individually on their particular topics of interest, which often involve different languages, making it additionally difficult to share their work. The group defines itself loosely as not theory-driven (presumably in contradistinction to the syntax and semantics group) and descriptive; this seems not the strongest conceptual basis for constituting a research group, which in fact it really is not, since its members do not work together. Its a-theoretical approach also calls the scientific relevance of the work into question. Much of the research carried out in this programme seems “old school” and not attuned to recent developments in the fields of historical and functional linguistics, such as corpus linguistics, grammaticalisation, and variationism. It may be for these reasons that the group has difficulty in securing external funding. In certain respects, it seems anachronistic.

As such, while it continues established traditions and might well ride on this momentum for a number of years more, this is not a forward-looking programme. Unless it becomes more flexible and dynamic, it may not have a long-term future. Recommendation: The members of this group should seek to increase the relevance of their work by addressing more often its broader (including interdisciplinary) implications, e.g., for understanding the nature of language variation and change. Indeed, a variationist would make a good addition to the staff. They should also seek to publish in more visible venues, and to address (occasionally) a more general readership.

### Quality

**Productivity**

5

**Relevance**

3

**Vitality**

3

The mark for quality reflects the lack of visibility in international journals, and the perhaps somewhat conventional focus of the research. The mark for relevance could easily be improved: many of the research questions pursued here have potential interest for a broader audience. To improve the mark for vitality, the tone of resignation should give way to a more enthusiastic response to the issues and exigencies of our modern times.

## 3.5 Computational Linguistics

Programme Leaders: Prof. John Nerbonne, Dr. Gertjan van Noord

**Assessment** Although small, this is one of the strongest Computational Linguistics groups anywhere. In terms of permanent contributions to the field the group has accomplished much more than many larger groups. Their strengths are in the theoretical combination of traditional symbolic linguistics along with probabilistic modelling, along with — unusually — large scale serious implementations. The committee has nothing but praise for them.

However, small groups full of talented individuals are always going to be vulnerable if one or more of the main players is attracted elsewhere. In an ideal world an enlightened department should be trying to reward their success in attracting external funding and their excellent international reputation.

During their presentation the group made a request that the CLCG policy of providing matching for externally acquired funding should be continued. We think this is altogether reasonable.

Connections with computer science and artificial intelligence should be vigorously pursued, not the least for the opportunity to tap these reservoirs of suitable students.

In sum: this group deserves more tenured staff. The committee also advises hiring a phonetician, if the speech processing path is being pursued.

### Quality

5

### Productivity

5

### Relevance

5

### Vitality

5

## 3.6 Educational Linguistics

Programme Leaders: Prof. Kees de Glopper and Prof. Kees de Bot

**Assessment** This group obviously has had some problems in the past. Until the appointment of the new programme leaders, this programme was not functioning well, with productivity only fair, and number of PhD students very low.

The department should be congratulated on the appointment of two new professors who seem to have brought more focus to the programme. Since their appointment productivity has gone up, quality has improved and the number of PhD students has increased.

The committee is impressed by the nice balance of theoretical and applied research, by the combination of qualitative and quantitative approaches, and by the recent outpouring of papers in international journals. Nice to see, for once, more journal papers than book chapters.

During the presentation, De Glopper and De Bot put forward the question whether they should continue to join forces or split up in a group studying language development with intervention and one focusing on development with little or no intervention. Our advice is for them to stay together: it does not seem wise to change or split up a winning team. ‘Educational Linguistics’ is a bit of a misnomer. We recommend changing the title of the programme into ‘Language development over the life span’.

### Quality

4

### Productivity

4

### Relevance

5

### Vitality

5

The high marks for relevance and vitality reflect the turn that this group has taken under the direction of the new programme leaders.

## 3.7 Neurolinguistics and Experimental Linguistics

Programme Leader: Prof. Roelien Bastiaanse

**Assessment** The entire programme includes four subgroups which differ considerably in orientation, size and performance. Since there are only six permanent staff members altogether, these subgroups are relatively small. Their common denominator is that they all suffer from insufficient infrastructure (i.e., administrative, secretarial and technical help).

Among the subgroups “Aphasiology” was by far the most successful over the last five years. Their findings are well published and they are well received. The subgroup was very successful in getting external grants. Their record with respect to dissertations is very good. One critical point concerns the linguistic background. Linguistic assumptions are mainly phrased in terms of the “classical” Government and Binding theory, as developed in the early 1980s by Chomsky. This is surely not the state of the art in linguistic theory, neither in Generative Grammar nor elsewhere.

Another subgroup — with the misleading name ‘Experimental Linguistics’ — focuses on the neurological basis of language processing. This work is done in a competent way. External funding came from NWO (Pionier). The most interesting output is a recent, very comprehensive paper by Stowe, Haverkort and Zwarts ‘Rethinking the neurological basis of language’ which presents a kind of synthesis of older and newer work (including the authors’ own) on the various areas involved in comprehension and — to a somewhat lesser extent — other types of language processing.

The third group focuses on dyslexia. This work was made possible by a huge grant to Zwarts (1998–2005), but there are few publications so far. A mitigating factor here is that the project is still in the process of carrying out a longitudinal survey. The work so far is very competent, but one would need more evidence to be able to declare this project a success.

The fourth group is the Language Acquisition group. This group seems misplaced in this research line. The work done here is doubtlessly competent and well received, but this research would seem to fit better into Educational Linguistics (Behrens) or Syntax and Semantics (Van Hout and Hollebrandse).

Recommendation: change the programme title to either “Language and the Brain” or “Language in the Brain”.

### Quality

4

## **Productivity**

4

## **Relevance**

4

## **Vitality**

4

The mark for productivity is an average. Bastiaanse, as the leader of the group, has been exceptionally productive. The external funding received by other members of the group (NWO Pionier grant, Dyslexia grant) has led to high expectations; however these expectations have not yet been fully met by the research output of these projects.

# Appendix A

## Curricula Vitae of Members of Evaluation Committee

Jan van Eijck (1951) is professor of Computational Linguistics at the University of Utrecht and Senior Researcher at the Center for Mathematics and Computer Science (CWI) in Amsterdam. He obtained his B.A. in Philosophy at the University of Amsterdam, his B.A. in Dutch Language and Literature at the Free University of Amsterdam, his M.A. and Ph.D. in Philosophy at the University of Groningen. From 1997 until 2002 he was Scientific Director of the Dutch Graduate School in Logic (OzsL). His main interests are computational linguistics, natural language semantics, and applications of dynamic and epistemic logic in computer science and in natural language analysis.

Susan C. Herring (1955) is professor of Information Science at the School of Library and Information Science at Indiana University in Bloomington, USA. She obtained a B.A. in French at the State University of New York and a M.A. and a Ph.D. in Linguistics at the University of California in Berkeley. She worked as assistant, subsequently associate professor at California State University in San Bernardino, before accepting the post of associate professor in Linguistics at the University of Texas at Arlington. Since 2003 she has been full professor at Indiana University. Her main research interests are: Computer-mediated communication; discourse analysis; content analysis; gender and technology; new media and language and historical linguistics.

Wolfgang Klein (1946) is Director of the Max-Planck Institute for Psycholinguistics in Nijmegen, the Netherlands. He studied German and Romance Philology and Philosophy at University of Saarbrücken, Germany and holds a Ph.D. in sentence parsing from that University. From 1972 until 1976 he was associate professor at the University of Heidelberg, before obtaining the Chair of Linguistics at the University of Frankfurt. Since 1980 he has been Director of the Max-Planck Institute. Professor Klein is an expert in Psycholinguistics, language acquisition and sociolinguistics.

Stephen Pulman (1949) is professor of General Linguistics at Oxford University. He obtained a B.A. (hons) at the University of London and his M.A. and Ph.D. at the University

of Essex. From 1978 until 1984 he was a lecturer in Essex and at the University of East Anglia. From 1984 until 2000 he was University Lecturer, subsequently Reader in Natural Language Processing at the University of Cambridge. Since 2000 he has occupied the Chair of General Linguistics at Oxford University. His main areas of expertise are: the development, application and evaluation of all types of natural language processing technology; machine learning for language applications; text, dictionary and corpus processing; information extraction and text mining; spoken language dialogue systems.

# Appendix B

## Declaration of Independence

Quality assurance at the University of Groningen

Selection Criteria and Guarantee of Independence for Peer Review Committees

Peer review and quality assurance committees are expected to produce authoritative, critical and independent assessments of the quality of the research schools, institutes or programmes they have been asked to examine. This means that the members must meet high standards with regard to quality. The authority of the assessment in terms of quality, objectivity and influence stands or falls with the independence of the assessing peers. It is in everyone's interests that such peer review committees be carefully selected in order to guarantee their independence. This appendix lists selection criteria for members of peer review committees as well as instruments to guarantee the independence of these committees.

Contents

1. selection criteria for peer review committees
  2. reporting obligation for the research schools and institutes to be assessed if they foresee potential conflicts of interest, prejudice or influence by potential/proposed peer review committee members
  3. code of behaviour, including a declaration of independence for peer review committee members
- 
1. Guidelines for selecting a Peer Review Committee When choosing an external peer review committee (PRC) which conforms with the criteria of independence, expertise and academic quality, the following points must be taken into consideration when selecting potential candidates:
    - Authoritative scientific expertise in at least one discipline or sub discipline of the department to be assessed

- National or international authority in the field
- Independence with regard to the department to be assessed and to the researchers within the department
- Insight into, and if possible some expertise in, related disciplines and sub disciplines
- Insight into and an overview of national developments in the field \* Insight into and an overview of international developments in the field
- Insight into relevant interdisciplinary developments
- Some familiarity with how research is organised in the Netherlands.

In order to determine the independence of the potential chairperson and members of the visitation committee, the following issues at least must be considered:

- Excluded from a PRC are:
  - (former) employees or PhD students of the institute to be assessed,
  - (former) members of an advisory body for the institute to be assessed (or the associated Research School),
  - co-authors of scientific publications from employees or PhD students of the institute to be assessed.
- Has the potential candidate ever worked intensively with members of the department to be assessed, for example, long-term participation in alliances, regular participation in PhD assessments?
- Has the potential candidate close links with one or more members of the department to be assessed, for example as the PhD supervisor of a member, or as a member of the same research group, joint editorships?

If one or more of these questions must be answered with yes, then this must be clearly stated by the institute when proposing the candidate in question. It should also be made clear why the board is of the opinion that the independence of the proposed candidate can be sufficiently guaranteed.

When potential candidates are approached with the request to participate in a PRC, they will be asked to sign a standard declaration of independence, including a brief code of behaviour (see below), before accepting. During the final meeting, the members of the committee will be asked to confirm or expand the declaration they signed earlier, and to state that they have actually fulfilled their commitments.

## 2. Reporting obligation

The list with potential peer review committee members must be presented to the heads of the programmes, research schools and institutes to be assessed before it is sent to the Executive Board. The former are obliged to report any potential conflicts of interest, prejudice or influence on the part of the proposed peer review committee members and must be able to report and substantiate their objections in writing to the Faculty Board.

### 3. Code of behaviour + declaration of independence for peer review committees

The following will be sent together with the invitation to participate to the individual members of the peer review committee and must be signed and returned before the site visit takes place.

#### Competence and independence of peer review committee members

1. Members of the peer review committee must base their assessment primarily on:
  - the Standard Evaluation Protocol 2003-2009 for Public Research Organisations
  - the 'specific peer review protocol' adopted by the Executive Board of the University
2. When judging the quality of research, members of the peer review committee must base their assessment on the following information:
  - the self-evaluation report and accompanying documentation
  - possible additional information provided at the request of the peer review committee
  - interviews, lectures and talks conducted within the framework of the assessment
3. Members of the peer review committee must meet the generally accepted quality demands within scientific research, including:
  - competence and professionalism independence and objectivity care and
  - consistency transparency and impartiality
4. Members of the peer review committee may not have any personal, scientific, financial or any other potential conflicts of interest when participating in the research assessment of ..... [name of institute] and are therefore both qualified and competent to carry out their task as independent assessors.
5. Members of the peer review committee must report any potential conflicts of interest in the assessment procedure to the chairman of the review committee.

I declare that I have read the above and that I will comply with it to the best of my ability.

- Place and date:.....

- Signature:.....
- Name: .....

1 Based on the format of the national organisation "Quality Assurance Netherlands Universities" (QANU) d.d. Dec. 2003

2 This national protocol can be downloaded from: <http://www.qanu.nl/?contentid=144>

3 The RUG protocol and other relevant documents will soon be available at: <http://www.rug.nl/kwaliteitszorg-onderzoek> Appendix to the Groningen protocol for assessing research

# Appendix C

## Extended Description of the 5 point scale

**Excellent (5)** Work that is at the forefront internationally, and which most likely will have an important and substantial impact in the field. Institute is considered an international leader.

**Very good (4)** Work that is internationally competitive and is expected to make a significant contribution; nationally speaking at the forefront in the field. Institute is considered international player, national leader.

**Good (3)** Work that competitive at the national level and will probably make a valuable contribution in the international field. Institute is considered internationally visible and a national player.

**Satisfactory (2)** Work that is solid but not exciting, will add to our understanding and is in principle worthy of support. It is considered of less priority than work in the above categories. Institute is nationally visible.

**Unsatisfactory (1)** Work that is neither solid nor exciting, flawed in the scientific and or technical approach, repetitions of other work, etc. Work not worth pursuing.

# Appendix D

## Checklists for the Research Evaluation

The evaluation committee may use the following checklists for the assessment of an institute and its research programmes. Filled in checklists will not be published but are meant as a tool only.

- 5 = excellent,
- 4 = very good,
- 3 = good,
- 2 = satisfactory,
- 1 = unsatisfactory

**Institute** How do you evaluate the institute with respect to:

1. Leadership
2. Mission and goals
3. Strategy and policy
4. Adequacy of the resources
5. Funding policies
6. Facilities
7. Academic reputation of the institute

8. Societal relevance of the institute
9. Balance of the strengths and weaknesses of the institute

Overall assessment of the institute

Remarks and questions:

**Research Programme** How do you evaluate the programme with respect to:

1. Leadership
2. Mission and goals
3. Strategy and policy
4. Adequacy of the resources
5. Funding policies
6. Facilities
7. Academic reputation
8. Societal relevance
9. Balance of the strengths and weaknesses

Overall

**Quality** How do you evaluate quality with respect to:

1. originality of the approach and ideas
2. significance of the contribution to the field
3. coherence of the programme
4. publication strategy
5. prominence of the programme director
6. prominence of the other members of the research group
7. quality of scientific publications (scientific impact)
8. quality of other results

Overall assessment of quality

**Productivity** Considering the number of staff, how do you evaluate the productivity with respect to:

1. number of Ph.D. theses
2. number of scientific publications
3. number of professional publications
4. other results (if applicable)
5. distribution of published output within the group

Overall assessment of productivity

**Relevance** Considering the stated mission of this programme, how do you evaluate the relevance of the research with respect to:

1. the advancement of knowledge
2. the dissemination of knowledge
3. the implementation of knowledge

Overall assessment of relevance

**Vitality and feasibility** Considering the present status and future developments (if known) of staff and facilities, how do you evaluate the long-term viability of the programme:

1. in view of the past scientific performance
2. in view of future plans and ideas
3. in view of staff age and mobility

Overall assessment of vitality

# Appendix E

## Schedule of the Site Visit

**Thursday, July 1** dinner with review committee, dean, director, representative of faculty, representative of university board

**Friday, July 2 8:30** Coffee

**9:00** Welcome by dean Ger de Haan Faculty perspective on review

**9:45** John Nerbonne, Director of CLCG

**10:30** Break

**11:00** Phonetics/Phonology

**11:30** Syntax/Semantics

**12:10** Discourse/Communication

**12:50** Lunch

**13:50** Discussion of first groups

**14:20** Descriptive/Historical

**15:00** Coffee/Tea

**15:30** Computational Linguistics

**16:10** Educational Linguistics

**16:50** Neurolinguistics

**17:30** Discussion of afternoon groups

**18:00** Closing day one.

**19:30** Dinner

**Saturday, July 3 8:30** Coffee

**9:00** Graduate Students

**9:30** Director of Graduate. Studies, Jack Hoeksema

**9:50** Postdocs  
**10:10** Kees de Bot, Chair, CLCG Board  
**10:30** Coffee  
**11:00** Discussion of morning session  
**11:30** General Discussion. Preparation of Preliminary Report  
**12:30** Lunch  
**13:30** Further discussion at committee's discretion, with group leaders, Nerbonne, and de Haan available for consultation.  
**14:30** General Discussion of Report  
**15:30** Presentation of preliminary findings  
**16:30** Instructions to committee members by chair  
**17:00** Closing, day two.  
**19:00** Dinner