



Centre for Language and Cognition Groningen

ASSESSMENT OF RESEARCH QUALITY

1998-2003



RuG

Rijksuniversiteit Groningen

VER
BYM
DNI
LV
CER
NA

Assessment of Research Quality

Centre for Language and
Cognition Groningen
1998-2003

RuG

University of Groningen

P.O. Box 72
9700 AB Groningen
The Netherlands

Department of Academic Affairs and International Relations

Phone +31 50 363 5370
Email g.b.de.vries@bureau.rug.nl
Website www.rug.nl/kwaliteitszorgonderzoek/rapporten

Centre for Language and Cognition Groningen

Postbus 716
9700 AS Groningen
Phone +31 50 363 5858
Email clcg@let.rug.nl
Website www.rug.nl/let/onderzoek/onderzoekinstututen/clcg

Ontwerp Dorèl en anderen, Groningen

Foto's Elmer Spaargaren, Groningen

Druk PlantijnCasparie, Groningen

Assessment of Research Quality

Centre for Language and Cognition
Groningen
1998-2003

Contents

1	General information	7
1 1	Introduction	7
1 1 1	National system for assessing the quality of research	7
1 1 2	Outline of the RUG Protocol	8
2	Review of the Centre for Language and Cognition Groningen (CLCG)	11
2 1	The review committee	11
2 2	Scope of the assessment, assignment to the committee	11
2 3	Input for the research assessment	12
2 4	Working procedure of the Committee	12
2 5	The site visit	13
2 6	Assessment criteria and rating	14
3	Assessment of the Centre for Language and Cognition Groningen (CLCG)	19
3 1	Mission and Goals	23
3 2	Self-Evaluation	23
3 3	Leadership	23
3 4	Overall Shape of the Centre	23
3 5	Position of CLCG within the Department	23
3 6	Balance of Strengths and Weaknesses	23
3 7	Facilities	23
3 8	Graduate Studies and Selection of PhD Projects	23
3 9	Overall Assessment of the Institute	23
4	Research Programs	23
4 1	Phonetics and Phonology	23
4 2	Syntax and Semantics	23
4 3	Discourse and Communication	23
4 4	Descriptive and Historical Linguistics	23
4 5	Computational Linguistics	23
4 6	Educational Linguistics	23
4 7	Neurolinguistics and Experimental Linguistics	23
	Reaction Management CLCG	35
	Appendix A: Quality assurance at the University of Groningen	39
	Appendix B: Selection Criteria and Guarantee of Independence for Peer Review Committees	47
	Appendix C: Curricula vitae of the peer review committee	51
	Appendix D: Research input and output of CLCG	53

Master's students in Linguistics.



General information

1.1 Introduction

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),¹ which is approved not only by VSNU but also by the Netherlands Organisation for Scientific Research (NWO), the Royal Netherlands Academy of Arts and Sciences (KNAW) and the Dutch ministry of education and science.

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 its immediate superiors within the university) and external (by the university to government and society). The SEP is based on two fundamentals:

- 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 SEP requires the committee members to evaluate the research institute or school as a whole and the relevant parts of the institute (research programmes) individually, 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 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 the KNAW, NWO and VSNU, will check the design and implementation of the new system by the various institutions and publishes its findings annually.

¹ This can be downloaded from: <http://www.qanu.nl/?contentid=144>.

112 Outline of the 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 universities. Subsequently, the RUG developed the so-called “Protocol for Quality Assurance at the University of Groningen”.

The following principles underlie the RUG protocol:

- a *There is a close connection with the RUG quality policy*
- b *There is a clear division of tasks and responsibilities*
- c *The external assessment is transparent, authoritative and is relevant for both internal policy and external accounting*
- d *The aim is professionalisation and minimal workload for researchers.*

Re a) RUG quality policy with regard to research

The RUG regards quality improvement as the dominant principle in its research policy. 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 feedback from the peer reviewers on 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 faculty and/or the university as a whole, its national and international context, as well as its disciplinary and interdisciplinary contacts.

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 c) Usefulness

The results of an assessment must be sufficiently informative to serve as the basis for policy decisions. Therefore, the possibility of adding a lower aggregation level – compared to the aggregation level referred to in the SEP – exists. In practice, the aggregation levels of research programs may vary strongly. However, if a research program is believed to be too large to receive an adequate judgement on all research covered by this programme, the Executive Board of the university may request for a supplementary evaluation at a lower aggregation level. 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 d) Minimal assessment

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 Peer Review Committee (PRC) to assess several Groningen Institutes. This option is offered to faculties aiming to cluster their multidisciplinary research institutes



Scientific Director John Nerbonne at a meeting of graduate students.



10

2.1 The review committee

The peer review committee members for the research assessment of the CLCG were appointed by the Executive Board of the University of Groningen, after a thorough selection procedure assuring an authoritative, critical and independent assessment of the research institute (see: Appendix B). The University Board appointed Professor Henk Verkuijl as the chairman of the committee. In consultation with the chairman the other members of the committee were selected. Due to illness Professor Verkuijl 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 committee consisted of:

Prof. Jan van Eijck, *Centre for Mathematics and Computer Science (CWI), Amsterdam, Universities of Utrecht and Amsterdam, Netherlands (chair)*

Prof. Susan C. Herring, *Indiana University, Bloomington, USA*

Prof. Wolfgang Klein, *Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands*

Prof. Stephen Pulman, *University of Oxford, UK.*

Drs. R.J. Bron, *Department of Academic Affairs, University of Groningen*, was appointed secretary of the review 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, and
- › the judgement is made without undue influence from the institute, the program or other stakeholders.

Additional information on the committee members and their curriculum vitae can be found in appendix C.

2.2 Scope of the assessment, assignment to the committee

The review committee was asked to evaluate the CLCG as a whole as well as its seven research programs, with the emphasis on strategy and organisational aspects for the institute assessment and the focus on the results and quality of the scientific research and on the future for the programme assessments. In their task the PRC had to take into account the general rules laid down in the Standard Evaluation Protocol.

The committee based its assessments on the self-evaluation report of CLCG and on discussions with the program leaders and other members of the CLCG. Within the CLCG

assessment, the following seven research programs were presented to the PRC for evaluation:

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

2.3 Input for the research assessment

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

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

The input for the research assessment consisted of a self-evaluation describing the research conducted by the CLCG of the University of Groningen between 1998 and 2003. It reports the institute's mission, describes the research programs, 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 self-evaluation report was written in accordance with the directives of the 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 programs are described in detail.

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

2.4 Working procedure of the Committee

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. In order to guarantee an independent assessment the PRC members individually pre-assessed the programs prior to the site-visit. The chair divided the seven CLCG research programs among the committee members according to their expertise; 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.

Prior to the site visit, during the first meeting of the Peer Review Committee, individual pre-assessments were compared and discussed. Furthermore, the committee decided on specific questions to be asked during the coming site visit.

2.5 The site visit

A site visit of the PRC was scheduled on 1 and 2 July 2004 and was – due to the above mentioned absence of professor Pulman – conducted by the three remaining committee members (Jan van Eijck, Susan Herring and Wolfgang Klein) and the committee secretary. In the evening of the 1st of July, the peer review committee was welcomed by the dean of the Faculty of Humanities (Prof. dr. G. de Haan), the scientific director of the CLCG (Prof. dr. J. Nerbonne) and the CLCG program leaders.

The first day was devoted to discussions with the dean of the faculty, Ger de Haan, with the scientific 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.

2 6 Assessment criteria and Ratings

Assessment criteria

The judgements of the evaluators referred to the CLCG as a whole, and to the seven research programmes of the CLCG. The main criteria were always reviewed in relation to the mission of the CLCG. The committee acted upon the description of the SEP (see below), concerning the interpretation of the four main assessment criteria.

Description of the four criteria according to the SEP and interpretation by the PRC

Quality is to be seen as a measure of excellence and excitement. It refers to the eminence of a group's research activities, its abilities to perform at the highest level and its achievements in the international scientific community. It rests on the proficiency and rigour of research concepts and conduct; it shows in the success of the group at the forefront of scientific development. The members of the peer review committee judged quality both relying on their own knowledge and expertise and on discussions with the group leaders and other members, and on various kinds of systematic information. When an institute provides high quality state of the art facilities to the research community this can be considered as a measure of excellence.

Productivity refers to the total output of the group; that is, the variegated ways in which results of research and knowledge development are publicised. Usually, quantitative indicators measure this. In most cases this will be bibliometrics, which are indicators concerned with publications and citations of publications. In some cases technometrics (largely concerned with patents and citations of patents); or sociometrics (concerned with socio-economic performance or embedment of research) can be applied. The output needs to be reviewed in relation to the input in terms of human resources.

Relevance is a criterion that covers both the scientific and the technical and socio-economic impact of the work. Here in particular research choices are assessed in relation to developments in the international scientific community or, in the case of technical and socio-economic impact, in relation to important developments or questions in society at large.

Vitality and feasibility This dual criterion refers to the internal and external dynamics of the group in relation to the choices made and the success rate of projects. On the one hand, this criterion measures the flexibility of a group, which appears in its ability to close research lines that have no future and to initiate new venture projects. On the other hand, it measures the capacity of the management to run projects in a professional way. Assessment of policy decisions is at stake, as well as assessment of project management, including cost-benefit analysis. The questions to be answered with these assessments concern both the research institute and the research programmes. These questions are:

For past performance:

- › What are the quality and relevance of the institute?
- › What is the quality of the leadership, management, strategy and research programmes of the institute, its (human) resources, organisation and infrastructure and how can they be improved?
- › To what extent has the institute/research programme achieved its mission and goals formulated for the period under review?

For future plans:

- › Is the mission of the institute well chosen and phrased in view of the actual developments in the relevant research field(s)?
- › How do you assess the institute's research plans and is there sufficient coherence in the research portfolio of the institute?
- › What is the quality of the leadership, management and strategy of the institute, its (human) resources, organisation and infrastructure and how can they be improved?
- › Which of these aspects has room for improvement and how could that be accomplished?

Ratings

The PRC presents its assessment on quality, productivity, relevance and vitality/feasibility according to a five-point scale for each aspect. It should be noted that it will not always be feasible or satisfactory to measure on such a scale; in which case the verbal commentaries can supply more information.

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.

VSNU 1992-2002		SEP 2003-2009	
5	Excellent	5	Excellent
		4	Very good
4	Good	3	Good
3	Satisfactory	2	Satisfactory
2	Unsatisfactory	1	Unsatisfactory
1	Poor		

The extended description of the five point scale is as follows:

Excellent 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 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 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 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 Work that is neither solid nor exciting, flawed in the scientific and or technical approach, repetitions of other work, etc. Work not worthy of pursuing.

ERP research on the early identification of dyslexia. Fiona Koster is adjusting the cap.



18

3 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.

3 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.

3 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.

3 4 Overall Shape of the Centre

All in all CLCG seems to be in good shape. On average, the research staff members 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.

3 5 Position of CLCG within the Department

CLCG is one of three research centres (institutes) within the humanities department, the other two being Archaeology (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 fulfill 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.

3 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 actors), 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 as well. '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.

37 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.

3 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.

3 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.

Gertjan van Noord and Gosse Bouma leading a discussion among computational linguists.



24

The Research Programmes

4 1 Phonetics and Phonology

PROGRAMME LEADER ² :	Dr. Dicky Gilbers, Dr. Tjeerd de Graaf until his retirement in January 2003
TENURED STAFF:	1.20 fte (in 2003)
TOTAL RESEARCH STAFF:	4,4 fte (in 2003)

4 11 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 put 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.

² Programme leadership is on a provisional basis.

4 2 Syntax and Semantics

PROGRAMME LEADERS:	Prof. Jan Koster and Prof. Jack Hoeksema. Former programme leaders: Prof. Co Vet and Prof. Alice ter Meulen
TENURED STAFF:	4,55 fte (in 2003)
TOTAL RESEARCH STAFF:	8,95 fte (in 2003)

4 2 1 Assessment

This is – or has been – a highly productive group. The quality of the work is high, in fact it 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 put 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.

4.3 Discourse and Communication

PROGRAMME LEADER:	Prof. Gisela Redeker
TENURED STAFF:	2.20 fte (in 2003)
TOTAL RESEARCH STAFF:	3,5 fte (in 2003)

4.3.1 Assessment

This is a relatively small programme in terms of its staff, with one full professor and eight assistant professors, and no (current) non-tenured 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 members, 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:

the strong societal relevance of its subject areas (especially new media such as the Internet and the World Wide Web, and organisational applications), and 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.

4 4 Descriptive and Historical Linguistics

PROGRAMME LEADER:	Dr. Peter Houtzagers
TENURED STAFF:	3,90 fte (in 2003)
TOTAL RESEARCH STAFF:	6,25 fte (in 2003)

4 4 1 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	3
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.

4 5 Computational Linguistics

PROGRAMME LEADERS:	Prof. John Nerbonne, Dr. Gertjan van Noord
TENURED STAFF:	1.45 fte (in 2003)
TOTAL RESEARCH STAFF:	8,65 fte (in 2003)

4 5 1 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

4 6 Educational Linguistics

PROGRAMME LEADERS:	Prof. Kees de Glopper and Prof. Kees de Bot
TENURED STAFF:	2.15 fte (in 2003)
TOTAL RESEARCH STAFF:	4.25 fte (in 2003)

4 6 1 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.

4 7 Neurolinguistics and Experimental Linguistics

PROGRAMME LEADER:	Prof. Roelien Bastiaanse
TENURED STAFF:	3.30 fte (in 2003)
TOTAL RESEARCH STAFF:	13.50 fte (in 2003)

4 7 1 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. Overall the output of this project, in terms of journal papers, is a bit disappointing.

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.

foto 5: Tuba Yarbay, Eleanora Rossi and Erik-Jan Smit in a graduate seminar on methodology.

34



Reaction Management CLCG

Introduction

The Centre for Language and Cognition, Groningen (CLCG) was evaluated in the summer of 2004 resulting in the findings on the previous pages. The CLCG internal advisory board met on November 2, 2004 to discuss their reaction to the evaluation and in particular to the recommendations of the PRC. The present section summarizes CLCG's reactions to the PRC's report and sketches how the institute plans to deal with the PRC's recommendations, focusing on those recommendations which would represent deviations from earlier practice. We also sent this reaction to the PRC, who found it "fair" and "reasonable." The recommendations to continue present practice are not commented on separately, nor, for the most part, are criticisms which did not result in recommendations to modify practices. The evaluation scores are likewise ignored in what follows.

General Reactions

The PRC consisted exclusively of accomplished, experienced scientists who all brought a broad view of linguistics to their work, one which resonates with the CLCG's own view. Moreover, the evaluation, while jibing with the institute's own views in most major points, also indicated unexpected areas where improvements might be made. The institute is deeply grateful to the very professional conduct of the PRC, the dispatch with which they conducted their work, and the clarity with which they formulated recommendations.

General Recommendations

Since the internal advisory board was not enthusiastic about the recommendation to establish an external advisory board, fearing further distraction from primary research, the director and deputy director will explore opportunities to obtain outside advice with little bureaucratic overhead, and no obligation of involvement on the part of all groups.

- CLCG agrees reluctantly with the recommendation to dissolve the Phonetics and Phonology group, agreeing in particular that the group is too small to function, and that there is too little promise of fruitful collaboration. The institute feels that phonetics and phonology is an absolutely essential part of linguistics and hopes the small number of specialists in this area can be strengthened in the near future. We explicitly include this expertise among those which need to be strengthened.
- We also agree that the language acquisition work will be more at home within the Language development group, and our language acquisition experts have accepted this suggestion.
- The Discourse and Communication program forewent strengthening in managerial skills, but has already hired two new members from whom we hope to see serious bridge-building efforts

in the direction of computational linguistics on the one hand and experimental linguistics (part of neurolinguistics) on the other. In addition, both the new members have expertise in semantics.

- › The syntax and semantics groups have worked at maintaining ties and collaborations with other research in Groningen (which the PRC in correspondence emphasized as the main thrust of its criticism), but also agrees that renewed efforts would be useful. The institute sees important results of contact between syntax and other research groups, particularly neurolinguistics. The group agrees that semantics needs shoring up (and as we note above, we have in fact moved to do just this).
- › The groups which have had fewer meetings to-date have agreed to start modest efforts at regular, perhaps initially infrequent meetings.
- › All of the relevant name changes suggested will be adopted.
 - ‘Educational Linguistics’ will henceforth be called ‘Language and Literacy Development over the Life Span’.
 - ‘Experimental Linguistics’ is not a group, but a subgroup, and we leave it to the group how it is to be called. We agree to call the larger group ‘Neurolinguistics’ simpliciter.
 - Everyone is enthusiastic about renaming, and hopefully improving the focus of the Descriptive and Historical Linguistics, which hence bears the name ‘Language Variation and Change’.
- › CLCG has identified statistics as an area in which strengthening is highly desirable (with semantics and phonetics/phonology).
- › The computational linguistics group has approached the computer department about computational support. All parties have asked that this be part of a master plan for centralizing support, which the central computing center is also backing.

Graduate Studies and Selection of PhD Projects

Everyone concerned supports the concept of a Master’s program open to all CLCG specializations.

The selection of PhD projects is an ongoing topic of discussion at CLCG advisory board meetings. The university policy calls for favoring stronger groups, so some imbalance is seen as desirable. There has also been critique that the current system gives incoming PhD students a great deal of freedom in determining the direction of a topic. This has the advantage that students are enthusiastic and committed about their projects, but this advantage leaves the research programs of the groups little opportunity to provide direction. We would like to find a happier balance between these desiderata.

Specific Research Groups

Phonetics and Phonology

This group has reluctantly decided to accept the recommendation that it dissolves, obviating the need for further comment.

Syntax and Semantics

This group feels that it is less insular than the assessment suggests, but emphatically agrees that cooperation with computational linguists and neurolinguists should be part of its research agenda.

Discourse and Communication

This group particularly appreciated the positive assessment of the relevance and quality of its work. As noted above, the group has initiated a search for a specialist in computer-mediated communication, and the group will take to heart the suggestion that student projects be harnessed to contribute to the group's research.

Descriptive and Historical Linguistics

This group finds many remarks in the assessment worthy of notice. First, as noted above, the group agrees that the name 'Language Variation and Change' covers the work of its members well, and suggests foci for cooperation within the group. Second, it has begun to explore the possibilities for common initiatives, such as joint (PhD) projects and international conferences. A conference with the preliminary title 'Language Contact in Times of Globalization' is being planned for the autumn of 2006. Third, the group will meet more regularly to discuss (future) joint projects and ongoing research. PhD students and postdocs, in particular, will be invited to discuss their work. One group member now co-organizes the weekly CLCG Linguistics Colloquium, which may help to promote the relevance of historical linguistics to other CLCG members. Finally, the group members have the impression that the relevance of publishing for specialist audiences, including those best addressed in languages other than English, is underestimated by the Committee. At the same time, they agree that addressing broader issues and a broader academic readership should be a point of special interest. One recent new member has in fact published a good deal on the grammaticalization phenomena and their relevance to general theories of language change.

Computational Linguistics

This group has taken the advice to heart to pursue opportunities for collaboration with Computer Science and Artificial Intelligence more vigorously. In addition the group is investigating possibilities to add Discourse and Communication (within CLCG) to this list of potential fruitful collaborators. The group would also welcome the opportunity to hire a phonetician, as would CLCG (noted above).

Educational Linguistics

This group will adopt the suggested name, and it will also remain together rather than split up. Finally, it welcomes the addition of the language acquisition specialists who until now were part of Neurolinguistics.

Neurolinguistics and Experimental Linguistics

The group, especially its senior members, agrees with the criticism that it needs to increase its scholarly productivity (exception duly noted by the PRC). This group will also simplify its name to 'Neurolinguistics' in conformance with the suggestions of the committee. As noted above, the two language acquisition specialists will leave this group to join the Language Development group. The group also agrees with the committee's criticism of the available administrative and secretarial infrastructure and will again bring this to the attention of the faculty administration. The group emphasizes that it intends general contributions to the understanding of language and the brain, even if its inspiration has mainly been drawn from Government and Binding theory, now outdated. The group will work to clarify this in presentation. The dyslexia project has not published enough, which project members feel is due to the many other demands on the project, especially the large number of children who need to be tested. It will be difficult to channel more energy into publication. The Neurolinguistics leader will take this up with project leader, the dean, and the rector (who is the original project leader), but as of this writing there is no clear strategy for improvement.



Quality assurance at the University of Groningen

National system for assessing the quality of research

In 2003, the Dutch system for assessing the quality of research was changed radically. 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 organization of the research quality evaluations. They are bound by the “Standard Evaluation Protocol 2003-2009” (SEP),³ which is endorsed not only by VSNU but also by the Netherlands Organisation for Scientific Research (NWO) and the Royal the Netherlands Academy of Arts and Sciences (KNAW).

The most important elements of the new system, set out in the SEP, are:

- › Three main aims:
 - improving the quality of research;
 - improving research management and direction;
 - improving accountability;
 - internally – by the unit to be assessed to its immediate superiors within the RUG, and externally – by the RUG to government and society.

- › Based on two fundaments:
 - an external assessment once every six years (by a peer review committee which conducts a site visit);
 - self-evaluation once every three years (one in preparation for the external assessment and one intermediate evaluation three years later, the ‘mid-term review’).

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

To this end, the Executive Board of the University of Groningen (CvB) will draw up a schedule for all the units to be assessed and ensure that all the research is evaluated.

.....
³ This can be downloaded from: <http://www.qanu.nl/?contentid=144>.

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.

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. With regard to the instructions for peer reviewers before they are appointed by the CvB, the text below explains the most important points in the SEP and the RUG Protocol.

SEP – outline of the main points

1 General

The evaluation applies to:

- > the quality of research according to the standards of the relevant academic disciplines
- > the way that the research results are reported to the academic world.

Depending on the mission of the unit to be assessed, the evaluation also examines:

- > socioeconomic aims
- > technological or infrastructural aims
- > cross-disciplinary aims.

The evaluation will be both retrospective and prospective. The results are intended to assist the research organization, the management of the research units and the individual researchers in decision-making about future research, research policy and research management.

The three central concepts in the SEP are:

Board: the Executive Boards (CvB's) of the universities and the boards of KNAW and NWO are responsible for the organization and procedural processing of the evaluation of the 'institutes' which fall under their responsibility.

Institute: the unit to be assessed is referred to in the SEP as '(research) institute' and defined as follows: 'An institute may be defined as "a group of researchers with an articulated shared mission operating under the same management". Each "institute" will have a director, board and/or research leader(s) with a final responsibility. Throughout this document they will be referred to as "the management".'

Research programme: this is the unit to be assessed, for which there is no specific definition in the SEP. Each programme should submit a title, programme leaders, research field and mission, as well as the research capacity of the academic staff, the share of the research resources within the unit to be assessed and the research output.

The peer review committee (PRC) reports to the Board (CvB). The CvB will make policy decisions concerning the Institute based on this report and the discussions about it with the Institute. The decisions of the CvB and the evaluation report together form the results of the evaluation. These results will be reported to the Minister of Education, Culture and Science (OC&W) via the normal channels (annual reports).

2 Assessment criteria

The evaluations will differ per institute and per programme:

- › per institute: the emphasis is on strategic and organizational aspects
- › per programme: the emphasis is on results, quality and the future of the research.

The main criteria are:

QUALITY:	international recognition and innovative power
PRODUCTIVITY:	scientific output
RELEVANCE:	scientific and socioeconomic impact
FEASIBILITY:	flexibility, management, leadership.

These criteria will always be related to the mission of the institute or group, which may, for example, limit itself to national scientific tasks.

The feasibility criterion includes the ability to close dead-end lines of research and start new projects. With regard to management, the criterion includes the ability to implement projects in a professional manner. This covers policy decisions and project management, including an analysis of the costs and benefits.

The questions to be answered include:

Retrospectively:

- 1 What is the quality and relevance of the Institute?
- 2 What is the quality of the leadership, the management, the strategy and the research programmes of the institute, the personnel and material resources, the organization and the infrastructure, and how may this be improved?
- 3 To what extent has the institute or programme realised the mission and goals of the period to be assessed?

Prospectively:

- 1 Has the mission of the institute been chosen and expressed well, given the current developments in the relevant field of research?
- 2 How can the research plans of the institute be assessed and is there sufficient coherence in the institute's research portfolio?
- 3 What is the quality of the leadership, the management, and the strategy of the institute, the personnel and material resources, the organization and the infrastructure, and how may this be improved?
- 4 Which of these aspects has room for improvement and how may this be realised?



The CvB may ask the PRC to investigate additional questions. These may refer to specific tasks of the institute that are not directly related to research, or to specific circumstances such as major changes to the organization or mission of the institute, or to specific demands from stakeholders who significantly contribute to the financing of the institute. If desired, confidential parts of the assessment can be included in a management letter to the CvB.

3 **Documentation for the PRC**

In preparation for the site visit, the peer reviewers will be sent a self-study report, the Specific Visitation Protocol and any supplementary questions from the CvB. Further, the chair of the PRC may request supplementary documentation.

Self-study report

Appendix 3 of the SEP states the format of the documentation to be supplied for a self-evaluation. This must serve as the basis for a strength-weakness (SWOT⁴) analysis, as set out in Chapter 4 of the SEP. Together they form the self-study report, which is in principle identical for both the mid-term review and the self-evaluation in preparation for a visitation.

The Specific Visitation Protocol

The SEP must be supplemented by the profile of the PRC, a list of supplementary questions and any supplementary information for the PRC. This enables the protocol to be adapted to the specific wishes of the CvB. Together with the SEP, this comprises the Specific Visitation Protocol for the external assessment in question.

Outline of the RUG Protocol⁵

1 **Starting points of RUG policy concerning assessing the quality of research**

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 must be transparent, authoritative and able to be applied to both internal policy and external accounting;
- 4 the aim is professionalisation and minimum disruption for researchers.

Re a) Quality policy with regard to research

The heart of this policy is that the RUG regards quality improvement as the dominant principle

.....
⁴ Analysis of 'Strengths, Weaknesses, Opportunities and Threats'.

⁵ The RUG-protocol can be downloaded from <http://www.rug.nl/Corporate/onderzoek/kwaliteitszorg/index>

in its research policy. Quality not only plays an important role in its own policy but also 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 recognized 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.

The principle of listening to both sides of the case will be applied before an external assessment will be accepted.

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 faculty, the university as a whole, the national and international context, as well as disciplinary and interdisciplinary contacts with regard to content.

Re c) Usability

The results of an assessment must be sufficiently informative to serve as the basis for policy decisions. This is why the possibility to add 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 vary strongly. There may be good reasons for working with larger programmes, something which thus cannot be ruled out in advance. However, if programmes are chosen that are so large that in the opinion of the Executive Board the assessment is not sufficiently usable for internal policy decisions, a supplementary evaluation at a lower aggregation level will be requested. The Faculty Board will itself submit a motivated proposal that must then be approved by the Executive 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 d) Minimum assessment disruption

All Institutes at the RUG are organized locally and according to discipline. In the old national system, an Institute was assessed at the same time as comparable research groups at other universities by one PRC. In the new system, although the national, disciplinary visitation is no longer compulsory it is certainly still an option. Voluntary co-operation is still possible, on condition that it is approved by the relevant Executive Boards. The RUG is determined to keep the option for national co-operation open, particularly because of the greater 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 faculties aiming to cluster their multidisciplinary research institutes.

2 **Composition of a Peer Review Committee (PRC)**

The responsibility for appointing PRCs is borne by the Executive Board. The RUG abides by the following guidelines:

- › The PRC must comprise nationally or internationally renowned scientists who are experts in the disciplines or subdisciplines of the unit to be assessed.
- › The expertise in the PRC as a whole must sufficiently cover all the subfields within the unit to be assessed.
- › The peer reviewers must be authoritative but may not be interested parties. In order to guarantee the independence of the peer reviewers, they are obliged to sign a standard declaration of independence before accepting membership of a PRC.
- › The Executive Board reserves the right to submit the list of prospective candidates to external experts before they are appointed. The aim is to create a national code of behaviour with regard to this. The KNAW, NWO and VSNU are expected to submit a proposal.

Secretary

The CvB will appoint the secretary of the committee and ensure, after consultation with the Faculty Board, that he/she is properly instructed. The secretary must on the one hand be independent of the research unit; on the other, he/she must be sufficiently familiar with the local situation. In principle, the secretary will be someone from the RUG Office. The secretary will be appointed at the same time as the committee.

3 **PRC procedure and reporting method**

Instructions for the committee

In consultation with the Faculty Board, the CvB will commission the committee and ensure that it is instructed. Within the framework of the Specific Visitation Protocol, the RUG protocol and the commission, the committee will determine its own procedure.

Programme site visit

Under the responsibility of the Faculty Board, and in consultation with the chair of the PRC, the management of the research unit will design the programme for the site visit, bearing in mind the provisions of the SEP. The Faculty Board will inform the CvB of the programme. The CvB will receive the PRC at the start of the site visit.

Format of the report

With regard to content, the committee will be guided by the SEP and by any supplementary questions posed by the university.

To achieve uniformity in the visitation reports, the secretary of the PRC will use a basic format for the preparation of the final assessment.

Verification of facts

The PRC will present the draft report to the management of the research unit for verification of the *facts* (correction phase).

Report to the CvB and the Faculty Board

After the facts have been verified, the PRC will report its findings simultaneously to the CvB and the Faculty Board in a draft report. If the faculty is the research unit being assessed, the committee will report only to the CvB.

Check of formal requirements

In consultation with the Faculty Board, the CvB will check whether the draft report is complete and consistent with the SEP and the supplementary questions posed by the RUG. If either the CvB or the Faculty Board wish expansion or explanation of the report, the CvB will request the chair of the committee, via the secretary of the committee, to provide these additions or explanations.

Acceptance or rejection of the visitation report

In consultation with the Faculty Board, the CvB will decide whether to accept the visitation report if, in its view, the visitation report conforms with all the requirements in the SEP as well as those set by the RUG.

On occasion, and after consultation with the Faculty Board, the CvB may decide to reject the visitation report. The CvB will make such a decision known to the chair and members of the PRC.

Adoption of the evaluation report

The report by the committee and the reaction of the research unit together form the final evaluation report that will be adopted by the CvB, in consultation with the Faculty Board. With this, the CvB concludes the external evaluation.

Right of response

The draft report will be presented by the Faculty Board to the management of the assessed research unit who will be asked to *comment regarding content* and to react to the findings of the external assessment.

Responsibility of the CvB

The CvB, on the basis of its own conclusions, is responsible for reporting to the Supervisory Board.

The CvB will determine *in general terms* the way in which evaluation reports will be presented and published externally. This includes reports to the Minister of OC&W, VSNU, KNAW, NWO and sister institutions as well as publication on the university's website. External reporting of the managerial implementation of the research assessments will be included in the annual report. The starting point is the conditions set by the SEP for the public evaluation report.

Management letter

In addition to the public evaluation report, the PRC will be asked, if necessary, to submit a confidential management letter to the Faculty Board, with a copy to the CvB. This management letter shall be based on meetings with the management of the research unit and include any sensitive information concerning personnel or company-sensitive information about the current or future position of the research unit. If necessary, the Faculty Board, after discussions with the CvB, will discuss the management letter with the chair of the PRC.

If the faculty is the research unit being assessed, the committee will address the management letter only to the CvB.

During the correction phase and in consultation with the Faculty Board, the management of the unit may ask the PRC to move parts of the report to a management letter in certain circumstances, such as contractual obligations to third parties, restrictions on making things public, etc.

4 Managerial implementation of the visitation report

Managerial assessment and measures

The CvB will ask the Faculty Board for a managerial assessment of the visitation report. In the event of shortcomings revealed by the report, the CvB will ask the Faculty Board, as the responsible and authorized body, which measures the Faculty Board has in mind to effect improvements. The Faculty Board may submit the evaluation report to the advisory board or supervisory board of the research unit and/or to the faculty academic committee(s) and ask for advice. The Faculty Board will inform the CvB of the measures to be taken. The CvB will discuss the measures with the Faculty Board, and testable agreements and how they are to be monitored will be formulated. In line with the Higher Education and Research Act (WHW), the Faculty Board is responsible for the design and the quality of the research.

Conclusions arising from the management letter

If a confidential management letter is submitted, it will be discussed by the CvB with the Dean of the Faculty. The Dean will be asked to give those in charge of the research unit an opportunity to respond. In mutual consultation, the CvB and the Dean of the Faculty will discuss the conclusions and any measures to be taken.

Appendix B

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 subdiscipline 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 subdisciplines
- › 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 organized 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 A member of the peer review committee bases his/her assessment primarily on:
 - > the Standard Evaluation Protocol 2003-2009 for Public Research Organisations
 - > the 'specific peer review protocol' ascertained by the Board of the University
 - > if applicable: additional instructions of the Board of the University

- 2 In giving a judgement on the quality of research, a member of the peer review committee grounds his/her assessment on the following information:
 - > the self evaluation report and accompanying documentation
 - > if applicable: additional information provided on request of the peer review committee
 - > interviews, lectures and talks carried out within the framework of the assessment

- 3 A member of the peer review committee meets the generally known quality demands within scientific research, including:
 - > competence and professionalism
 - > independence and objectivity
 - > care and consistency
 - > transparency and impartiality

- 4 A member of the peer review committee experiences no personal, scientific, financial or any other potential conflicts of interest in participating in the research assessment of [*name Institute*] and is therefore both qualified and competent to carry out his/her task as an independent assessor.

- 5 A member of the peer review committee reports any potential conflicts of interest in advance to the chairman of the review committee.

I declare that I have read the above-mentioned and that I will follow these to the best of my ability.

Place and date:

Signature:

Name:

Curricula vitae of the peer review 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 D

Research input and output of CLCG

Table 1 **Research input (in fte) CLCG 1998-2003**

	1998	1999	2000	2001	2002	2003	TOTAL
Tenured staff	21.15	20.85	20.55	18.15	19.05	18.75	118.50
Non-tenured staff	7.20	8.85	11.50	11.15	11.10	6.90	56.70
PhD students	17.00	17.00	18.20	22.85	22.60	23.85	121.50
Total research staff	45.35	46.70	50.25	52.15	52.75	49.50	296.70
Supporting staff	-	-	-	-	-	-	-
Total staff	45.35	46.70	50.25	52.15	52.75	49.50	296.70

Table 2 Research staff (in fte) at programme level CLCG 1998-2003

	1998	1999	2000	2001	2002	2003	TOTAL
Phonetics and Phonology							
Tenured staff	1.25	1.55	1.55	1.65	1.65	1.20	8.85
Non-tenured staff	0.00	0.50	0.50	0.50	0.50	0.00	2.00
PhD students	0.80	0.80	2.40	2.40	2.80	3.20	12.40
Total research staff	2.05	2.85	4.45	4.55	4.95	4.40	23.25
Syntax and Semantics							
Tenured staff	5.65	5.65	5.65	5.20	5.15	4.55	31.85
Non-tenured staff	1.80	3.60	4.60	2.30	3.00	2.00	17.30
PhD students	2.00	0.80	0.80	0.80	1.60	2.40	8.40
Total research staff	9.45	10.05	11.05	8.30	9.75	8.95	57.55
Discourse and Communication							
Tenured staff	2.70	2.70	2.50	2.20	1.95	2.20	14.25
Non-tenured staff	1.00						1.00
PhD students	1.60	1.60	0.80	0.80	1.60	1.30	7.70
Total research staff	5.30	4.30	3.30	3.00	3.55	3.50	22.95
Descriptive and Historical Linguistics							
Tenured staff	6.00	5.70	5.30	3.90	4.10	3.90	28.90
Non-tenured staff							
PhD students	3.40	2.60	1.60	2.25	2.20	2.35	14.40
Total research staff	9.40	8.30	6.90	6.15	6.30	6.25	43.30
Computational Linguistics							
Tenured staff	1.10	1.20	1.20	1.20	1.45	1.45	7.60
Non-tenured staff	2.40	0.70	2.00	4.50	4.40	1.00	15.00
PhD students	3.40	4.20	4.60	7.80	6.60	6.20	32.80
Total research staff	6.90	6.10	7.80	13.50	12.45	8.65	55.40
Educational Linguistics							
Tenured staff	2.15	1.75	1.75	1.40	1.95	2.15	11.15
Non-tenured staff		0.25	0.50	0.25		0.70	1.70
PhD students		0.80	1.60	1.60	1.40	1.40	6.80
Total research staff	2.15	2.80	3.85	3.25	3.35	4.25	19.65
Neurolinguistics							
Tenured staff	2.30	2.30	2.60	2.60	2.80	3.30	15.90
Non-tenured staff	2.00	3.80	3.90	3.60	3.20	3.20	19.70
PhD students	5.80	6.20	6.40	7.20	6.40	7.00	39.00
Total research staff	10.10	12.30	12.90	13.40	12.40	13.50	74.60
Total	45.35	46.70	50.25	52.15	52.75	49.50	296.70

Table 3 Aggregated results CLCG 1998-2003

	1998	1999	2000	2001	2002	2003	TOTAL
Academic publications							
– in journals	53.0	60.0	49.0	63.0	43.0	47.0	315.0
– book chapters	75.0	83.0	87.0	84.0	65.0	89.0	483.0
Total Academic publications	128.0	143.0	136.0	147.0	108.0	136.0	798.0
Monographs	2.0	3.0	1.0	1.0	2.0	2.0	11.0
PhD theses	9.0	2.0	2.0	6.0	8.0	6.0	33.0
Professional publications and products							
– Prof. publ.	43.0	68.0	49.0	40.0	37.0	29.0	266.0
– Reviews	27.0	24.0	16.0	19.0	17.0	5.0	108.0
– Editorial act.	8.0	8.0	8.0	16.0	7.0	8.0	55.0

This publication is part of a series on the assessment of research quality
at the University of Groningen.

ISBN 90-367-2241-2

Also published in this series:

Northern Centre for Healthcare Research – University of Groningen
Assessment of Research Quality 1997-2002 (March 2005)



CLCG
Center for Language and Cognitive Sciences

