The humanities are currently experiencing a ‘technological’ as well as a ‘cognitive’ turn. The Meertens Institute intends to include these important changes in its new research programme *Crossing Boundaries*. We believe that these two developments may considerably enrich our investigations into Dutch language and culture. It should be noted that we have already started integrating aspects of both turns into our research over the last few years.

As the technological means for searching, enriching, mining, analyzing and visualizing data come within reach of scholars in the humanities, the result will be a more objective and verifiable empirical perspective which will cause the differences between ‘hard’ and ‘soft’ sciences to gradually diminish. The availability of digital data, huge infrastructures and advanced applications for organizing the data and searching out complex patterns will increase the empirical validity of the research. I truly believe that we are on the threshold of a substantial change in humanities research. Naturally, future research questions will not strongly diverge from the questions we raise at present. Rather, we will observe a change in the objectivity and hence in the quality of the answers to existing questions, which will ultimately yield new questions and potentially new paradigms.

In the present research programme *Crossing Boundaries*, we see this development emerge in many new projects with a technological aspect, not only in the linguistic domain, which was traditionally quite advanced in this respect, but also in other domains of our research into the dynamics and diversity of Dutch language and culture. The technologically most advanced group at this moment investigates oral culture, i.e. Dutch popular songs and stories. As the description of the projects in chapters 3 and 5 shows, the Meertens Institute has invested substantially in technology during the past period and will continue to do so in the future. The Institute’s goal is to achieve a leading position in the national field and a strong international position.

Of course, the research methods best known in the humanities and social sciences communities (library and archival research, fieldwork and participant observation) will be continued. Many research questions raised in the humanities cannot or can only be partially answered by studying large corpora. A move towards technology does not therefore imply giving up the more traditional methods that have proved to be successful in studying language and culture.
The increase of the relevance of technology is not the only break with the past in this research programme. In our research plans two other developments stand out. First, we will attend to the cognitive aspects of language and culture more than before. Several projects in the domains of music, language, visual arts and poetry investigate the question as to what extent language and culture are shaped or predisposed by human nature. One of the major issues in this domain is the relation between the cognitive capacities under investigation and the way these capabilities interact.

Another feature of the new research plan is a more explicit distinction between internal and external approaches to language and culture. External approaches very clearly focus on questions of national or regional identity, ethnicity, and cultural heritage. They deal with how language and culture are constructed socially and how they resonate in the lives and experiences of the Dutch people.

Finally, a major organizational change introduced in this research plan, which will be in force as of January 2013, is the discontinuation of the two research groups Dutch Ethnology and Language Variation: there will now be a single research group comprising a number of selected theme groups. The actual research will be organized in temporary research projects, each of them related to one of the themes. The aim of this reorganization is to create a more flexible and dynamic organization in which the Institute’s researchers have well-defined research tasks within the projects.

In the coming years there will be closer collaboration with the humanities institutes in the Academy, in the framework of the KNAW Humanities Centre (see Contourennota KNAW, 2012). Especially in the fields of collections, technology and back-office activities, joining forces may have interesting consequences, possibly leading to innovation. Although nothing definite can be said yet, it is to be expected that more collaboration will have serious consequences for the organization of the Meertens Institute in the coming years.

It is the goal of this research plan to lay a solid base for a discussion on the Institute’s future.

Hans Bennis
director Meertens Institute

December 2012
| 1.1  | Introduction          |
| 1.2  | Towards temporary research projects |
| 1.3  | Objectives            |
| 1.4  | Research area: Empirical domains, perspectives and theoretical approaches |
| 1.5  | Organization          |
| 1.6  | Scientific quality    |
| 1.7  | Earning capacity      |
| 1.8  | Societal relevance    |
In a world increasingly determined by global connections and developments, there is a growing need for knowledge of one’s own society. The importance of the Meertens Institute, whose primary focus is the study and documentation of Dutch language and culture, increases correspondingly. On the strength of specific and unique empirical material, the Institute contributes to wide-ranging current scientific debates and developments in these fields of study by virtue of internationally appealing scientific research. Besides, the Meertens Institute also sets itself the task of sharing the generated insights with Dutch society at large, thereby meeting society’s legitimate need of reflection.

**Introduction**

There is a continuing need for disciplinary and interdisciplinary research into language and culture, well-established areas in themselves in the national and international research community. Our contributions to the fields of linguistics and ethnology will remain the major ingredients in our research efforts, specifically as far as the study of everyday culture and language variation and change is concerned. The empirical focus of this research is on Dutch language and culture, while the scientific focus is comparative and directed towards the international fields of linguistics and ethnology. The Meertens Institute will continue its successful research on these topics, while at the same time it intends to connect to innovative developments in our fields of study, such as computational approaches to the analysis of the diversity and dynamics of variation and research into the relation between cognition and patterns of variation. With its highly qualified research staff, its large digital databases and tools, and its scientific and technological expertise, the Meertens Institute should be able to maintain and even strengthen its position as a frontrunner in these developments.

Our research plan for the period 2013-2018 differs in a number of respects from the preceding ones. These changes are motivated by changing needs in the day-to-day practice of scientific research in the humanities and also to some extent by the fact that the context in which the Institute operates is changing. Society evinces a growing interest in our central research themes, i.e. in topics such as national, regional and local identities, intangible cultural heritage, multiculturality, language variety and change. The Institute is aware of questions in relation to these issues and we consider it our task to contribute to possible answers from a scientific point of view. Secondly, there is a rapidly advancing technological impetus in the humanities that will have a substantial impact on research activities. Another factor is the fact that the Dutch govern-
innovation

ment has adopted a different perspective on science in its stance on innovation. This has become an important issue in the last few years, making it necessary for scientific groups to form coalitions with industry and non-scientific non-profit organizations. A substantial part of financing science through funding organizations such as KNAW, NWO and TNO is related to the so-called innovation platform. A final factor making it desirable to opt for a different perspective on research is related to changes within the KNAW-humanities organization. In a recent position paper (July 2012) the KNAW board chooses to cluster the KNAW institutes in the humanities and urges for closer cooperation in research, collections, technology and back-office activities.

clustering

Taken together, these four developments make it essential for us to reconsider our research priorities and the organization in which our research is embedded. In this research plan we opt for a more flexible and dynamic research organization with which we will be able – more than before – to meet the developments mentioned above. The crucial ingredients of this reorganization of the research organization are:

- the change from structural research groups to temporary research projects;
- the discontinuation of the distinction between ethnology and language variation in the organizational structure;
- a focus on technology and eHumanities within the research projects;
- a focus on valorization within the research projects;
- the collaboration of technologists, documentalists and researchers in the selected research projects;
- a stricter monitoring of the research results;
- a separate group for Collections & Technology;
- representation of Collections & Technology on the Management Team;
- an increase in temporary research staff, implying a decrease in structural research positions;
- a move towards rejuvenating the structural research staff.

Towards temporary research projects

Dutch language and culture

From its foundation in 1930, Diversity in Dutch Language and Culture has been the central theme for research and documentation in the Meertens Institute. In the first decade of its existence, Piet Meertens brought together the Dialect Committee and the Folklore Committee in one organization. Although in those years the focus was on documentation, the leading idea was that the study
of variation in language and culture belongs together. This idea still holds, as is obvious from the fact that the KNAW board has recently formulated ‘Dutch culture, history and language’ as one of the two research areas of the KNAW humanities institutes.

However, from the early days on, the Meertens Institute has kept apart its two major areas of documentation and research, viz. dialectology and ethnology. Documentation and research functioned in separate groups for language and culture. These groups were relatively independent, in organization, in documentation and in their research strategy. They interacted only at management-level of the Institute as a whole. Interaction at the research level was accidental and infrequent. This is also clear from the two preceding research plans “Het oog op de toekomst” (1999) and “Dynamische tradities” (2006). To a large extent, the separation of the two fields is due to a different orientation in the sciences, with different methodologies for the two fields. The dialectology group operates in the context of linguistic theory based on structuralist views, and in later years oriented towards a generative perspective. The study of European ethnology on the other hand started out as a historical discipline and was increasingly influenced by the social sciences, e.g. anthropology and cultural sciences.

This implies that the Institute has operated in an ambivalent situation. From an empirical perspective, language and culture appear to be part of one domain in which they combine naturally. From a scientific perspective, however, the two fields are separated by diverging scientific traditions and by different methodologies. Given this situation, the Institute has up to this point chosen to accept and adopt the distinction by separating the two fields in two relatively independent research groups.

This research plan diverges from this twofold perspective. As of 2013, the two research groups will be discontinued. Research will now be organized in smaller research units and temporary research projects. Jointly, the research projects determine the majority of the Institute’s research activities. Of course, the projects may differ quite substantially in their focus and methodologies, as is to be expected given the variety of disciplines related to the research fields of language and culture. The Institute respects and supports this variety. It will do so by making the strategic research agenda an important topic for ongoing discussion in the Board of Research, which will receive a prominent position in the organizational structure of the Institute.

Below, we approach the research of the Institute from a general perspective – objectives, research domain, organization, quality and relevance. Then we focus on the theoretical perspectives for the coming period (Chapter 2), and the...
way these perspectives are implemented in research projects for the period 2013-2018 (Chapter 3).

### Objectives

The Meertens Institute has a long tradition in the documentation and research of Dutch language variation, in particular dialects, and in the traditions and rituals we encounter in everyday culture. In the previous decades the Institute’s focus has shifted from documentation to research. At present, research is the major component of the Institute’s activities, and documentation and other activities are to a large extent made dependent on the scientific efforts. In this paragraph we describe the activities of the Institute from the perspective of research. There are six characteristics together determining the relevance of research activities at the Meertens Institute. The combination of these characteristics makes our research unique in comparison with research taking place at other institutions.

**shift from documentation to research**

A research Institute of the Royal Netherlands Academy of Arts and Sciences has an obligation to produce excellent research on its field of inquiry, in our case Dutch language and culture. The focus of our research activities is directed towards the understanding of variation and diversity in Dutch language and culture (see below). Moreover, the research is comparative and should be internationally prominent, giving rise to new theoretical insights and perspectives. Its scientific prominence should follow from important publications in a-level journals and books with prestigious publishers, in high-level book publications, in being successful in acquiring funds, in key-note and state-of-the-art-lectures etc.

**intellectual prominence**

The Meertens Institute has systematically built and acquired a large number of collections which provide unique perspectives on our language and culture. These collections and, whenever relevant, new data to be collected, form the starting point for most research projects.

**documentary prominence**

Large amounts of data can be brought together and made relevant for research by well-structured, sustainable, open-access databases with, among other things, standardized metadata and persistent identifiers, and by the development of specialized applications. In recent years, the rapid development of electronic databases and tools has opened up new and fascinating research perspectives, especially with regard to large quantities of data. When-
ever relevant, research projects should have an eHumanities component in which new approaches to the data are exploited and investigated.

Research in the Meertens Institute is organized in groups in which researchers, collection specialists and technologists work together. These teams work on temporary topics with a concrete focus in terms of scientific relevance, explicit goals, concrete workplan and planned output. The shift from individual research to research teams is gradually taking place in the humanities. Given the complexity of our research topics, we expect collaboration in research teams to have a positive influence on the results.

Working in a national and international network is important for creating cooperation between research groups focusing on the same or similar research questions. Moreover, this increases the impact of the projects. In the coming period, we will keep supporting existing international networks such as Edisyn, SIEF, the Diachrony network and the eLaboratory for Oral Culture. We will continue and if relevant, expand our relationships with national and foreign universities. A topic that will receive special attention is cooperation with other KNAW-humanities institutes, and this topic has a special status since it constitutes an important aspect of the KNAW strategy (KNaw Contourenplan 2012). In this context, we will support research themes allowing cooperation with the Huygens Institute, NIOD, KITLV, IISG and the Fryske Akademy. At the moment the Institute collaborates with the Huygens Institute (Nederlab) and the Fryske Akademy (The Language Portal). Additional KNAW-projects are expected to start in the coming period.

In our view, it is important for the success of a research project to have a wider audience than the scientific research community alone. Projects should be active in finding ways of enhancing their impact by connecting to societal demands or interests. In the case of language and culture such connections can often be realized quite naturally. This is also related to the new innovation perspective initiated by the Dutch government and implemented by research organizations such as KNAW, NWO and TNO. We expect the increasing collaboration with partners in commercial and governmental areas and in heritage institutions to lead to applications of scientific results in the fields of language and culture — eg. media, ICT, cultural heritage, gaming — and to products that have a wider relevance than the scientific community alone.
Research area: empirical domains, perspectives and theoretical approaches

In this section we describe the empirical domains, the perspectives and the theoretical approaches determining the research space at the Meertens Institute.

**empirical domains**

The Meertens Institute has always taken Dutch language and culture to be its field of inquiry. However, within the rather broad empirical domain defined by these terms, there should be specialization in order to have sufficient focus in the research programme. This specialization is determined by the relevance of the subfields and by the potentials of the research staff. At present, four domains are assumed to be relevant. They are: oral culture, traditions and rituals, syntactic variation, and phonological variation. These domains have a long history at the Institute, are supported by large constructed corpora, and the Institute’s researchers belong to the international top of their fields.

**perspectives**

The four research fields above all employ a synchronic and a diachronic perspective while focusing on the understanding of the present situation. They study how present-day phenomena came about and how they spread and develop.

**approaches**

Given the variability and complexity of the data, we opt for a number of theoretical approaches optimally complementing each other. First of all, we have the *structuralist approach* which approaches the empirical object by investigating the observable properties and trying to find empirical generalizations that may give rise to an explanatory theory. In linguistics, the generative theory is important at the Meertens Institute. In ethnology, there are various approaches that are reminiscent of the structuralist approach, especially in the study of oral culture. A second option is the *interpretive approach*, which at this point is an important perspective in ethnology, focusing on the meanings and symbols people employ in human interaction and emphasizing agency and performance over structure. The interpretive approach is part of modern sociolinguistics as well, and as such connects to the main perspective in ethnology. Another way to gain insight into the phenomena is a *quantitative approach* making use of statistical techniques. In sociolinguistics, for instance, structuralist approaches are quite naturally connected to quantitative approaches. A quantitative orientation is relevant in some domains in the cultural sciences as well. Finally, we have the rather new and daring *cognitive science approach* to variation in language and culture. Obviously, the selected research domains have a cognitive significance, for instance with regard to the mental organization of the linguistic system or the way music is organized in the brain.
Organization

Until now the research has been organized in two research groups, Dutch Ethnology and Variation Linguistics. These groups are embedded in the Institute’s organizational structure provided below (see p. 14).

Within these groups there are a fixed number of structural research positions based on a division in scientific subfields. The domain of Dutch Ethnology is divided into Feast&Ritual, Religious Culture, Material Culture and Oral Culture. The field of oral culture is again divided into ‘folktales’ and ‘folksongs’ which have a structural research position each. The field of the research group Variation Linguistics is divided into six different subfields with six research positions (cf. research plan Dynamische tradities, 2006). They are divided over two research domains (phonology and syntax) and three empirical fields (geographically, diachronically and socially determined language variation).

Although it provides a relatively good coverage of the fields of investigation, this setup does not create enough flexibility and dynamics in the research programme, given that individual research is determined by the quality and ambition of the individual researcher occupying a specific cell within the structure, and not by the specific project he/she is working on. In order to create greater dynamics and more flexibility, the Institute’s management has decided to change the research organization in such a way that the research topics and the related staff positions will depend on temporary research themes and actual research projects being carried out. This perspective creates a new situation in which the division into two research groups will be replaced by a division into a number of research themes and related research projects and in which staff composition depends to a large extent on the selection of research projects. This perspective requires a stricter management of the research activities as actual projects must be selected and regularly monitored. It also creates a more dynamic situation requiring a constant discussion of the Institute’s research policies (see p. 15).

The Institute has already started implementing the organizational change described above. In January 2011 and May 2012 internal calls for research projects were launched. In the 2011 round six research proposals were submitted, three of which were accepted – The Life of the Liquidae, Oral Transmission and Dutchness – after an extensive, partly external, review and selection procedure. The first two projects started in 2011/2012 and the third will start in January 2013. In the 2012 round three projects have been submit-
Organization structure of the Meertens Institute 2012

**director**
Hans Bennis
(chair MT)

**ARBO**

**MT**
Sjef Barbiers
Willy Jongenburger
Herman Roodenburg
Hetty Garcia (secr.)

**General Affairs**
Willy Jongenburger (chair)
Hetty Garcia et al (secr.)
Marion Haring et al (fin.)
Marc Kemps-Snijders et al (techn.)
Marcia Lucassen (personnel)
Douwe Zeldenrust et al (collections)

**Variation Linguistics**
Sjef Barbiers (chair)
Hans Broekhuis
Leonie Cornips
Ben Hermans
Frans Hinskens
Marc van Oostendorp
Gertjan Postma
Nicoline van der Sijss
Ton van der Wouden

**Projects**
PhD-students
Postdocs

**Projects**
PhD-students
Postdocs

**CSB**
John Helsloot
Ben Hermans
Frans Hinskens
Theo Meder (chair)

**Dutch Ethnology**
Herman Roodenburg (chair)
Louis Grijp
John Helsloot
Peter Jan Margry
Theo Meder
Irene Stengs

**WeCo**
Regina Bendix
Liliane Haegeman
Roeland van Hout
Marc Jacobs
Joep Leerssen
Alexandra Lenz
John Nerbonne

**OC**

**Projects**
PhD-students
Postdocs
ted: Maps & Grammar, Construction of Local Identities and Invariants in Language Change. The first two were accepted and will start in January 2013. The third one is provisionally accepted and under revision.

Board of Research

In order to accommodate the new research structure, we will change the Institute’s organization by setting up a Board of Research, composed of the senior researchers of the Institute. This board will prepare decisions concerning the selection and monitoring of the research projects and the Institute’s research policy. As a consequence of the new policy, the research projects will function more as independent research units – e.g. with respect to finance and hiring personnel – than has been possible in the present organizational structure.

Summary of Changes

To summarize the changes with respect to the present situation:
- discontinuation of the Ethnology and Variation Linguistics research groups;
- the discontinuation of these research groups implies the discontinuation of the leaders of these groups as members of the Management Team;
- introduction of the Board of Research consisting of all senior researchers;
- the Board of Research advises on all research matters such as research strategy, research priorities, selection of research projects etc.;
- the Management Team does not decide on research matters without the advice of the Board of Research;
- the director is a researcher and a member of the Board of Research;
- the appointment of a temporary (3 years) chair of the Board of Research;
- the new Management Team includes two members of the Board of Research: chair and director;
- the director is responsible for the annual job-evaluation talks with the senior research staff. A more detailed protocol of these talks is available;
- largely independent status of research projects with respect to hiring of temporary staff, financial and organizational management of the projects, supervision and evaluation of the staff employed in the research projects;
- separate groups for Collections and for Technology;
- the head of the Technology group is a member of the Management Team;
- the discontinuation of the CSB, the scientific committee of the Institute, which is replaced by the more prominent Board of Research in the new structure.

Collections and Library

In an ideal situation the groups for technology, collections and library would be centralized in one group. In view of the present situation it will take time and effort to realize this. For the coming period, we opt for two groups, one for technology and one for collections. However, these groups are in a situation of change at the level of KNAW-humanities (cf. KNAW Contourennota 2012). It will take time and wisdom to arrive at the situation in which collections and tech-
nology are centralized in a single group in the organization of the Institute. The head of General Affairs, who also leads the department of Collections, and the head of Technology will both be involved in this process in the coming years.

Although the pace and the impact of the KNW-restructuring process are unclear, it is to be expected that the new structure will be subject to change as a consequence of the KNW-policy with respect to the (clustering of) KNW-humanities institutes. The Contourennota states: 'the autonomy of the institutes with respect to back office activities (finances, personnel, ICT and collections) will be discontinued.' These tasks will be delegated to the board of directors of humanities institutes. This may give rise to a situation in which the role of the heads and the departments of General Affairs, Collections and Technology in the structure above must be reconsidered. It is not clear yet how and when the implementation of the Contourennota will take place. What is clear though is that the Board of Research will have a crucial position in the organization, now and after implementation of the KNW-Contourennota.

Scientific quality

As a KNW institute, the Meertens Institute must be a scientific leader with respect to its fields of investigation in the Netherlands. It should also be a key player in the European context. The Institute’s researchers should be recognized specialists in their fields of study. The quality of their output should be high and their work must have a measurable impact on the international research community. Recent evaluation of the Institute has rated the quality of the Institute’s research as 'very good'. We intend to raise the quality assessment of the Institute to 'very good to excellent' in the coming period.

Of course, the most relevant way of positioning the Institute nationally and internationally is by means of important publications (journals and books) and lectures. The Institute will actively promote a substantial and high-level output of its researchers. We will expect a minimum of two substantial international publications in peer-reviewed journals or books per researcher (0.8 fte) per year (cf. section 1.3). As is clear from the Self-evaluation 2006-2011, this number of publications is more or less similar to the average output of articles in the preceding period, but on the individual level substantial differences are found. We will have a stricter procedure (possibly having consequences) for controlling the individual output in the coming period, taking other categories into account.
of output (e.g. books, databases, national publications) into account, in line with the KNAW-brochure ‘Quality indicators for research in the humanities’ (2011).

national and international cooperation

Another criterion for research projects at the Institute – both internally and externally financed projects – is national and international collaboration. The Institute is first applicant or, as the case may be, coordinator of many projects involving collaboration with universities and research institutes in the Netherlands and abroad. For instance, the NWO-groot projects ‘The Language Portal’ and ‘Nederlab’, the NWO-CATCH-projects ‘Fact’ and ‘Cogitch’, the KNAW Computational Humanities project ‘Tunes&Tales’ and the NWO-Horizon-projects of ‘Knowledge & Culture’ crucially involve national and international cooperation.

The idea is that the Institute should have a leading role in its research domains not only by producing outstanding results but also by supporting and initiating cooperation with important research groups or individual researchers outside the Institute. As mentioned above, an important issue for the coming period is collaboration with other KNAW-humanities institutes, in line with the KNAW Contourennota.

guest-researchers

The Institute supports and initiates cooperation with guest-researchers. There are different groups of guest-researchers. First, there is the group of retired or otherwise unemployed scholars and scholars with a temporary freelance position at the Institute. In the coming period this group will harbour among its researchers Har Brok, Ton Goeman, Joep Kruijsen, Piet and Anke van Reenen (Variation Linguistics), Willem de Blécourt and Eveline Doelman (Ethnology).

The second group consists of Dutch and foreign scholars involved in one of the research projects or studying part-time at the Institute to strengthen ongoing research projects. This category contains among others the KNAW visiting professor Andrew Nevins (2012, MIT), the Rubicon postdoc positions of Anna Strycharz (2012, University of Edinburgh) and Francesc Josep Torres Tamarit (2012-2014, University of Barcelona) and the temporary stay of Göz Kaufmann (2012, University of Freiburg). We will encourage initiatives to have more guest-researchers in the categories mentioned above in the coming period.

boards and conferences

Another way of creating a strong position in our fields of research is by holding positions on national and international boards and organizing national and international conferences. To give some examples: since 2002 the Institute has hosted the secretarial office of the international ethnology organization SIEF (secretary Margry); it is represented in the board of the Dutch Unesco-group for intangible heritage (Margry), members of the Institute are on the board of the European organization for generative linguistics GLOW (chair Barbiers), the organization for Language Variation ICLAVE (Hinskens) and the American Folklore Society (Margry). The Institute is also represented nationally and inter-
nationally at the interface of research and technology: CLARIN: Barbiers, Bennis, Kemps Snijders; CATCH: Bennis, Brugman, Alkhoven; SURF: Bennis. In the new platform for innovation the Institute is active in the area of Creative Industry (Bennis, Alkhoven). We will remain active in this area, in such new initiatives as The Language Archive and CLARIAH. As was the case in the previous period, the Institute will actively support the organization of conferences.

One of the Institute’s goals is for each structural researcher at the Institute to hold a professorship (‘bijzonder hoogleraar’, ‘0-hoogleraar’ or ‘deeltijd hoogleraar’) at a university. At this moment, nine members of the research staff hold such a position, at six different universities: University of Utrecht (Barbiers, Grijp), Free University Amsterdam (Hinskens, Roodenburg), University of Amsterdam (Bennis, Margry), University of Leiden (Van Oostendorp), Maastricht University (Cornips), Radbout University Nijmegen (Van der Sijs). This amounts to about 50% of the present structural research staff. There are concrete plans for at least one more position.

The Institute aspires to a strong position in the field of eHumanities. It has achieved an important position in this domain in the past few years, as evidenced by its role in CLARIN, CATCH, The Language Archive and CLARIAH, and by projects such as The Language Portal, Nederlab and Tunes & Tales. A driving idea behind the turn towards technology is that new technological developments appear to make it possible (and in the long run inevitable) to change the field of humanities studies, in particular in areas that make use of large amounts of data. The Institute has invested a great deal of energy and manpower in turning a fairly modest technological support department into a department of technological innovation occupying a leading position in our country and a strong position in the international context (together with the Max Planck Institute Nijmegen) in the fields of data management and infrastructure. Based on these innovations, we expect fascinating new possibilities for our research in the near future.

**Earning capacity**

The previous period saw a large number of externally financed projects. The number and size of these projects was fairly stable during the period 2006-2012, amounting to around 35% of the research/technology budget (see chapter 3.2). Funding comes from competitive programmes of large public research foundations such as the Dutch Science Foundation (NWO-
programmes, NWO-(middel)groot, NWO-VENI, NWO-CATCH etc.), the European Science Foundation (EURYI, ESFRI/Roadmap) and the American Andrew W. Mellon Foundation, but also from smaller foundations such as the Prins Bernhard Cultuurfonds, the Swiss Van Riemsdijk Stiftung, or Digitaliseren met Beleid. For the coming period we expect to be able to extend this percentage to 40% of the research/technology budget, based on the fact that we recently succeeded in acquiring a number of positions in the new large projects Nederlab (NWO, KNAW, CLARIN/CLARIAH) and Knowledge & Culture (NWO-Horizon).

KNAW-grants

The Institute also receives grants from competitive KNAW-funds. Recent examples are the KNAW Visiting Professor grant for Andrew Nevins in 2012, the Computational Humanities project Tunes & Tales that started on January 1, 2012, and the NIAS Theme Group ‘Global Society and Identity’ (2013-14).

increasing the number of PhD-students

Although the amount of external funding is substantial, we would like to have a steadier inflow of PhD-students paid by external grants. We will keep trying to increase the number of externally funded research projects in the coming period. As for technology, we have limited capacity to accommodate new externally financed projects until the end of 2013, especially since the hiring of specialized personnel with the required capacities is quite complicated.

Societal relevance

societal vs scientific relevance

The Institute is vividly aware of the role it can play in society in topics such as national identity, immaterial cultural heritage, multiculturality, the status of language varieties and language change. There are various domains in society (other than the scientific community) that are relevant for us. However, in all cases the activities directed towards societal issues are derived from our primary goals of research and documentation.

societal issues

The Institute is present and active in discussions on politics and policy, for example in the discussion surrounding the implementation of the Unesco convention on intangible heritage (collaboration with the National Unesco Committee). The Institute also welcomes cooperation with professionals in the fields of language and culture. It collaborates with the Netherlands Open air Museum (NOM) in Arnhem and the Zuiderzeemuseum in Enkhuizen, with various cultural organizations in CATCH (such as Sound and Vision, Royal Library, the National Archive, various museums), with cultural heritage organizations such as the Dutch Centre for Popular Culture (VIE), its Flemish
counterpart (FARO) and the Reinwardt Academy. In the context of the innovation platform (Creative Industry), the Institute will continue its role as the centre for the Cultural Heritage network (in collaboration with Sound & Vision, the National Service for Cultural Heritage (RCE) and the city council of Amsterdam).

Another group to be mentioned here is the layman interested in language and culture. The databases of the Institute are primarily directed towards research activities within and outside the Institute. However, we try to make these data available in open access to a wider audience, employing attractive user-interfaces allowing people to profit from our activities. These investments are extremely successful, given the enormous interest in our website. In 2010, the website of the Institute had almost two million visitors (unique IP-addresses) with more than 47 million page views (excluding search engines). We are in the process of building a new website and new webpages, among which Kew letters, Meertens questionnaires and Maps produced by the Meertens Institute.

Related to this are our recent efforts to let the interested outsider participate in our research through crowd sourcing. The first results of crowd sourcing will come out in the annotation of Dutch songs and the 17th and 18th century Kew letters. The results so far are very encouraging. We are at present engaged in making our large collection of questionnaires accessible for research purposes to the general public by way of crowd sourcing. In this way the scientific relevance of our collections is enhanced substantially and public interest in our data and our research is growing simultaneously.

Finally, the Meertens Institute will again show society the relevance of its research activities in a wide range of activities. Here we list, among others, popular books, articles and lectures, a digital newsletter, participation in the Science Week and the yearly publication of a booklet as a present to our relations. In the coming period we will keep looking for opportunities to present the results of our research and our digital databases to a wider audience in open access.
2

2.1 The targets
2.1.1 Linguistic targets
2.1.2 Ethnological targets
2.1.3 Overview of targets
2.2 Methods
2.3 Research themes in linguistics
2.4 Research themes in the field of ethnomology
2.5 Cognitive science approaches to language and culture
2.6 The external influence on the properties of language and culture
2.7 Research themes for the period 2013-2018
The diversity and variety of Dutch language and culture have always been central topics in documentation and research at the Meertens Institute. In the two earlier research plans – “Het oog op de toekomst” from 1999 and “Dynamische tradities” from 2006 – the red line has also been that at the Meertens Institute both fields of study are combined. However, from the very beginning the study of language variation in the Dutch language area (dialectology, as it used to be called, now variation linguistics) and the study of everyday culture (earlier folklore studies, now ethnology) have been separate entities, content-wise and organization-wise. In this new plan, we make a break with this tradition: all research on language and culture at the Meertens Institute will from now on be organized in temporary projects with a thematic design. In this chapter we will define the six research themes that have been selected for the coming period. In the next chapter we present the research projects that together constitute the thematic research of the Meertens Institute for the period 2013-2018.

This new research plan links up with a recent KNAW memorandum ‘De contouren van een Vernieuwings- en Stimuleringsprogramma van de Geesteswetenschappelijke Instituten van de KNAW’, from July 2012. An important dimension of this policy outline memorandum involves cooperation among KNAW institutes and a sharper characterization of research, one of the cornerstones being research on Dutch culture, history and language (Nederlandse cultuur, geschiedenis en taal). In this chapter we discuss in greater detail how the Meertens Institute approaches this field of investigation via its central research theme: research into the internal and external properties of the diversity and dynamics in the language and culture of the Netherlands.

The targets

Language and everyday culture are singularly complex objects of research, especially because of the variety of forms they display and the inherent variability they possess. The goal of research at the Meertens Institute is to contribute to an understanding of language and culture, in particular their diversity and dynamics. To achieve this goal, research is being done on a well-defined area – language and culture in the Netherlands – with the help of an extensive documentation system. The multifaceted object of research is approached in a comparative perspective from different angles which together aim at providing a balanced interpretation contributing to the general and international scientific discussion on human behaviour.
Language and culture can be studied using an internal or external approach. On the *internal* approach the question is to what extent language and culture are determined by structural properties of the object of study itself. This perspective entails that language and culture are studied without straightaway appealing to the social context in which they play a role. Continuing this line of argument, one might conclude that these properties can be taken to be, at least partly, the result of a specific, human, cognitive aptitude. After all, we are dealing here with components of human behaviour at least partly unique to the human species, often referred to as the ‘nature’ aspects of language and culture.

As far as research on language is concerned, this line of inquiry has a long tradition, in which the American linguist Chomsky in particular has given an important impetus to the study of the so-called ‘internal language’ or ‘I-language’. The central hypothesis is that general properties of the linguistic system are part of the innate competence of native speakers. Given these properties – universal grammar – and their interaction with other cognitive modules (speech, for example), a child acquiring language can pick up its native language. The internal system defines at the same time what is a possible language and what is not.

More indications are emerging that cognitive capacities play a decisive role in other aspects of human behaviour as well, and there is a growing awareness that these can be investigated by studying the internal structure of cultural phenomena. The main thrust is to gain a clear understanding of the dependence of cognitive functions like memory, ethics and emotions on neural mechanisms that are genetically encoded. Research by scholars like Dehaene, Damasio, Changeux, Hauser, Rizzolatti and Spelke shows that there are strong indications that also parts of human cultural behaviour are related to cognitive functions (knowledge modules) which belong to our ‘nature’, and therefore represent specific dispositions. This view partly stems from an older approach to culture, viz. the structuralist tradition of which the cultural anthropologist Lévi-Strauss is an important representative.

The opposite of the internal approach to language and culture is the *external* approach, in which attention is focused on specific linguistic and cultural practices which are studied in order to gain insight in the influence of external factors on language and culture. This tack is commonly referred to as the ‘nurture’ side of language and culture. The sociolinguist William Labov is an important scholar whose research is determined chiefly by factors governing ‘external language’ or ‘E-language’, while at the same time he takes a Chomskyan perspective on the complementary ‘I-language’ as his point of departure. The interdisciplinary tradition of ethnology has drawn its inspiration from a
number of scientists as far as the external approach to culture is concerned. Important representatives of this contextual approach are Foucault, Elias, Turner, Goffman and Bourdieu. Proceeding from an interest in the works of these scientists, comparatively much attention has been paid in ethnology from the beginning of the nineteen-nineties to the way in which groups of people appropriate or even literally ‘incorporate’ cultural repertoires in specific contexts.

Language and culture are both characterized by a complex interaction of human nature and social environment. Jointly, these act as vehicles for and exponents of everyday human behaviour, within which the large range of diversity offers to individuals and groups the possibilities to construct their own social identities. The diversity manifests itself in three different dimensions: social, geographical and historical. Given the similarities noted above, a programme in which both fields of study, language and culture, are linked should provide fascinating possibilities for research that is being carried out hardly anywhere else at the moment. For the coming period four projects are envisaged in which linguists and ethnologists collaborate: Dutchness (chapter 3.I.VII), Local Identities (chapter 3.I.VIII), Knowledge & Culture, subproject 1 (chapter 3.I.IVa) and Knowledge & Culture, subproject 4 (chapter 3.I.IVc).

**Linguistic targets**

In the previous decades *internal linguistics* was geared in large measure towards the discovery of the grammatical systems and parametrized universal principles making up the core of human language competence. Moreover, the focus was on the idealized language user whose grammaticality judgements in particular were considered scientifically relevant. This approach appeared to be a long way off from the detailed study of observed language variation which occupies centre stage at the Meertens Institute.

*External linguistics*, especially traditional dialectology and historical linguistics, has long been determined by the attention it pays to an explanation of the observed variation in terms of historical processes. Sociolinguistics, which came up in the second half of the twentieth century, was chiefly motivated by the desire to understand language variation as a process in which internal factors and social variation come together. In Labov’s classic model, language variation is analyzed quantitatively as a reflex of social categories that speakers belong to, such as age and sex. Recent developments (inspired by Bourdieu among others) do no longer take speakers to be passive ‘carriers’ of language varia-
tion, but rather see them as actors using language items for the expression of ethnic, regional or local identities. On this view, social and geographical categories are not clear-cut ‘entities’ (as in a Labovian framework and in traditional dialectology) but time-bound, layered, ever shifting, relational and situational phenomena.

Methodological differences
Methodologically speaking, there was a big difference between internal and external since external linguistics often availed itself of quantitative methods by means of which specific linguistic phenomena are correlated with geographical or social variables. Internal linguistics was qualitative in the sense that the idealized user’s internal grammar was best analyzed on the basis of non-quantified judgements of native speakers. The frequency of occurrence of a phenomenon is not relevant; what is relevant is whether or not a phenomenon is found to be a representation of the internal system of language.

microvariation
This situation has seen a gradual change. Also as a consequence of incentives emanating from the Meertens Institute – such as the international GLOW conference on microvariation organized in 2002 – generative linguistics came to address itself in increasing measure to microvariation in language and with that, to an ever more accurate empirical approach to variable and non-variable properties of a linguistic system. In this view, comparatively invariant properties can be seen as resulting from the underlying internal system of language, while other aspects of variation must be approached also from an external perspective. This perspective now underlies most of the linguistic research carried out at the Institute. It implies that the internal and external perspectives on language can be viewed as complementary aspects of variation linguistics that jointly yield an explanatory model for the occurrence of language variation and language change in all its manifestations.

general questions
The actual questions about language variation are:

Internal - ‘how is the system of language put together, such that it allows for the occurrence of language variation and language change?’
- ‘how does the system determines the limits of language variation and the direction of language change?’

External - ‘which historical, geographical and social factors play a role in the distribution and dynamics of language variation?’
- ‘how do people construct and represent geographic, social and ethnic identities in their language?’

Now a further question can be added:
- ‘what is the interaction between internal and external factors in language variation?’
Ethnological targets

Usually, ethnology does not draw sharp or explicit distinctions between internal and external perspectives. However, when we compare the older, philological approaches in ethnology – those dealing with the transmission of oral culture (partly going back to the brothers Grimm) – with the younger, more anthropology-oriented approaches – those dealing with for instance material culture, religious culture and feast and ritual – we may note fairly similar distinctions. Although the philological approaches have shown an interest since a couple of decades in issues of performance and appropriation, they have always cherished internal approaches such as tracing narrative or melodic motifs, the smallest structural elements in tales or songs.

In contrast, the anthropology-oriented approaches have mostly favoured external approaches, a tendency strengthened by the so-called ‘cultural’ or ‘linguistic turn’ of the 1970s and 1980s, inspiring many ethnologists to study such central notions as ‘everyday culture’, ‘heritage’ or ‘national (or local, regional and ethnic) identity’ as social constructions and to emphasize dimensions of agency, ways of seeing, and appropriation. Since the 1990s, however, ethnologists recognizing various shortcomings of the cultural turn (and often inspired by Bourdieu’s notion of the habitus and similar practice oriented views) have addressed issues of culture and nature. More recently, these developments have kindled a new interest in the corporeal, sensory and emotional dimensions of culture, in which also issues of cognition, comparable to those addressed in contemporary linguistics, have come to the fore.

Interestingly, such cognitive questions are now also being addressed in the most recent philological approaches. For instance, in studying songs the sensory and emotional effects of musical ‘hooks’ are being investigated. With its many national and international networks, ethnology at the Meertens Institute is eager to contribute to the various developments summarized here, and has already started to do so over the last few years. It sees the cultural and the present bodily (or phenomenological) turn as largely complementary, and will most definitely continue its investigations informed by the former turn. But it also aims at enriching such investigations with its research pertaining to the latter turn, thus further exploring the complex relations between structure and agency, nature and nurture and external and internal approaches to culture.

The actual questions about cultural diversity and change are:

Internal - ‘how does the (Bourdieuian) habitus function, such that it allows for the occurrence of diversity and change in everyday culture?’
How does the habitus also determine the limits of this diversity and the direction of this change?

External
- ‘which historical, geographical and social factors play a role in the diversity and dynamics of everyday culture?’
- ‘how do people construct and represent geographic, social and ethnic identities in their everyday culture?’

A further question will be added:
- ‘what is the interaction between internal and external factors in cultural change?’

Overview of targets

In the coming period the focus of research in the Institute will not only be on the internal and external properties of language and culture, but also on the relationship between the internal and the external, between nature and nurture. Besides, research on the relationship between language and culture will be in the offing. The current research in the area of everyday culture can then profit from the internal orientation of linguistic research, while the linguistic research may profit from the insights developed in recent ethnological research. Our conviction is that the issue of the diversity and dynamics of language and culture can only be understood from a multidisciplinary perspective. It is necessary therefore that research should not restrict itself to disciplinary research proceeding from a particular theoretical orientation, but that in addition to this, by collaboration of researchers from different disciplines, it should lead to a comprehensive and differentiated analysis of the subject of investigation.

Multidisciplinary perspective

Central research theme

Summarizing, we define the very general central theme of research of the Meertens Institute as:

Research into the internal and external properties of the diversity and dynamics in the language and culture of the Netherlands
Methods

Empirically-based research into the interaction between internal and external factors for the understanding of the diversity and dynamics of language and culture also requires the availability and verifiability of large quantities of data and methods of analyzing these, both quantitatively and qualitatively. The Meertens Institute possesses relevant and unique collections on different subfields. In the coming period some of the collections will be put into service once again for research. This also calls for developing new techniques and methods.

Collecting data as part of a research project will remain an important task, providing current research with an empirical basis. The techniques used for this are modern techniques using ICT, in addition to the usual fieldwork, archival research, study of literature, and questionnaires. The well-known written Meertens questionnaire has now been replaced by a digital questionnaire which is put before a Meertens panel. Likewise, web 2.0 techniques, e.g., crowd-sourcing, are deployed in collecting and enriching relevant data.

As set out in the chapter on technological innovation, new developments also make it possible to mobilize the available data fruitfully in current research by using a digital infrastructure. This involves database techniques and tools for making these databases accessible. To illustrate: it is important for research into the geographical diversity of everyday language and culture to make the data visible by means of cartographical applications, in which not only the distribution of the data is shown on a map, but also the possible correlations between phenomena can be represented, or in which a large number of phenomena can be assembled to make geographical differences visible in an aggregated manner. It is also important for quantitative research to be able to sort the available data with a statistical programma and uncover connections between variable phenomena. Especially in the case of large quantities of data a statistical method is crucial in order to be able to make the empirical generalizations serving to test and accentuate the developed hypotheses. It is necessary for the study of historical data that well-structured data collections are developed in which an ordered search can be conducted and which can be linked to cartographical and statistical applications.

Therefore, various research methods will be put to use in the coming period: participant observation, archival and library research, structured questionnaires, interviews, crowd sourcing, and experimental laboratory-like settings. More often than before, computational methods will be used. Finally, theoretical and
mathematical (statistical) modelling will be employed. These recent developments fit in with new programmes such as the KNAW programme Computational Humanities.

uncovering patterns

It is expected that these methods will help uncover structures and patterns in the language and culture of everyday life which would remain out of range in traditional research methods. We are dealing here with, on the one hand, patterns in cultural phenomena themselves, and on the other hand, with the distribution of cultural phenomena in geographical and social space and in time.

Research themes in linguistics

linguistics at the Meertens Institute

Variation linguistics at the Meertens Institute investigates the Dutch language in all its manifestations and diversity. As a research subject it is considered the result of the interaction between language-internal factors (language in the mind) and language-external factors (language in the world). A unique feature of the linguistic research at the Meertens Institute is that it highlights this interplay and thus sees language variation as the key to uncovering the language-internal as well as the language-external system. The ultimate goal is understanding language variation in all its facets. With this, the linguistic research at the Meertens Institute sets itself apart from, on the one hand, theoretical linguistic research that looks for the properties all languages have in common and to a great extent abstracts from variation and language-external factors, and, on the other hand, from sociolinguistic and language-in-use research that to a great extent abstracts from language-internal factors. In taking this stance, linguistic research at the Meertens Institute associates itself with one of the great debates in linguistics: the question as to which properties of natural language are determined by internal cognitive factors and which properties are determined by language-external factors. Moreover, in this way the Institute links up with the central scientific discourse on the question as to which parts of human behaviour are determined by human nature and which by the environment.

grammatical variation

A further restriction in circumscribing the research subject is motivated by the phrasing of the central question above. As the Institute is interested first and foremost in the system of language, the research is restricted to the grammar of Dutch (syntax, morphology and phonology) but emphatically from a comparative perspective in which other languages, if relevant, are taken into consideration. Diversity and dynamics in lexical variation and pronunciation are less systematic than grammatical diversity. These instances of language variation
will be investigated only to the extent that they are important for their insight into grammatical variation.

Research into the interaction between language-internal and language-external factors requires the availability of large quantities of data on language variation and tools for analyzing them qualitatively and quantitatively. This is another point where linguistic research at the Meertens Institute, with its phonological, morphological and syntactic databases, is unique (see chapter 5 on Technological Innovation).

Putting the interplay between language-internal and language-external factors at the forefront leads to two research themes:

(i) **boundaries of language variation**

Language spreads in space. Geographical space is an aspect of social space (language among people). During the past decades the Meertens Institute has produced phonological, morphological and syntactic databases and atlases depicting the distribution of language phenomena in geographical and social space. Our point of departure in the study of distributional patterning is that there are language-internal as well as language-external regularities.

The first task is determining which of the properties we have come across belong to the internal system and which to the external system. This is not a simple chore. Generally speaking, we may say that the properties belonging to the internal system result in relatively invariant phenomena while the external system manifests itself in the more variable properties of the language. Whether or not a particular property is variable can only be established after thorough investigation and on the basis of an abstract analysis, however.

An example may be helpful. Looking at the famous Dutch/German verb clustering at the end of a (sub)clause in 'Ik geloof niet dat hij dat [had willen zeggen] (I don’t believe that he [had want say] had wanted to say that), we note that the three verbs are placed in a specific order. Dialect data show that four of the six logical word order possibilities actually occur in the Dutch-speaking language area. The language system must therefore at least allow for four possibilities. The question is whether the two other orders [wollen had zeggen] and [zeggen had willen] are excluded by the language-internal system or happen not to occur on account of external reasons. Besides, the question remains as to whether the four attested orders are a consequence of the variability of verbs in a so-called verb cluster or whether there are independent variables interacting in such a manner that there appear to be four possibilities, while in reality there are fewer. Recent research does point in the direc-
tion of the latter option and provides an analysis in which only two of the six possibilities are yielded in the ordering process of three-part verb clusters. This gives us a window on the organization of the internal language system. Next, of the two orders generated by the internal system, it should be investigated in what ways they manifest themselves in time and space (properties of the external system).

**research questions**

This research theme leads us to the following research questions:
- Which properties of Dutch are variable and which are not?
- Which properties belong to the internal system and which to the external system?
- How are the variable properties distributed in time and space?
- Some variable properties are more resistant to variation than others. For example: (nearly) all Dutch dialects are V2 and sov (contrary to English, for instance), but they differ substantially in their verb conjugation. To what extent can these differences in variation be attributed to language-internal factors and to what extent to language-external factors? More generally: is there a qualitative difference between micro-, meso- and macrovariation?
- What is the difference between linguistic variables showing a clear geographical distribution and those that do not?
- What is the difference between linguistic variables showing a clear social distribution and those that do not?
- Why is there a sharp geographical boundary (isogloss) for specific phenomena and a shifting one for other phenomena?
- How do the various linguistic levels (sound, word, phrase, sentence) differ with regard to geographical and social spreading?
- To what extent is diffusion of linguistic variables determined by language-external factors (geographical/social/economic/political), and to what extent by language-internal factors (e.g. the operation of analogy, paradigmatic gaps).

±realized

Slightly imprecisely speaking, we may posit that the language system, the l-language, creates a dichotomy between grammatical phenomena which are possible or impossible in natural language – [±grammatical] – and phenomena which occur or do not occur in social space – [±realized]. This gives us four logical types of language phenomena:

A [±grammatical, +realized]
B [±grammatical, -realized]
C [-grammatical, -realized]
D [-grammatical, +realized]
A is the class of occurring phenomena admitted by language-internal factors, B is the class of phenomena that, although admitted, are not realized in varieties of the language due to language-external factors, C is the class of phenomena that are excluded by language-internal factors and do not occur. Class D is interesting since it concerns phenomena that are excluded by language-internal factors, but nevertheless do occur.

Gaining insight in the differences between A and B is important: why are certain possible phenomena not realized in a variation space, and why are others realized? Category D at first sight looks inconsistent since an occurring phenomenon implies grammaticality. Recent research, however, suggests that phenomena of type D occur but only as temporary variants constituting a transition from partial grammar $x$ to partial grammar $y$, where $x$ and $y$ are both part of the category [+grammatical].

(ii) the relationship between diversity and dynamics

As is implicit in the line of research set out in (i), we might take for granted that diversity and dynamics of language are likewise dependent on internal and external factors. There would then be no essential difference between synchrony and diachrony. Labov’s work in particular argues for such an approach. Phenomena often cited as relevant for language change, such as grammaticalization — the phenomenon that content words fade to function words and lose their autonomy as words — and language contact — the phenomenon that contact between languages is a driving force in language change — are not only relevant to studying the various stages of a language, but also have repercussions on the geographical and/or social variation space.

However, there are at least two other theories on language change in the international debate:

Internal - language change is determined by the way in which language is structured (structuralism, generative linguistics);

External - synchronic language structure is accounted for by historical processes (Junggrammatiker, Evolutionary Phonology, Grammaticalization processes).

It will be clear that we will be guided by the view that language change as well as language variation are determined by an interplay of internal and external forces. In this context, research into the relationship between synchrony and diachrony is important: especially which factors change simultaneously must be investigated as this may shed light on an underlying common factor of change. A special instance is the notion ‘failed change’: the rise of a new variant goes hand in hand with the rise and fall of a different variant.
Finally, it seems to be the case that the direction in which languages change is determined by an algorithm related to theory of evolution. The grammar (the internal system) generates relatively arbitrary variation; the context (the external system of a specific language variety) selects specific variants from the offer (e.g. on the basis of social or regional prestige) while other conceivable variants disappear from the system or are selected by other language varieties; finally, language acquisition takes care of the transfer of the selected variant from mother to child. There appear to be indications that the kind of acquisition can have an influence on the process of language change. There are clear differences between languages in a linguistically stable situation with little or no second language acquisition, and languages in an unstable situation in which many adult speakers learn the language as a second language (language contact). Research will have to show to what extent dynamics in dialects and other language varieties (e.g. the process of regionalization) support such a view of language change.

Research themes in the field of ethno

research themes in the field of ethnology

Ethnology at the Meertens Institute investigates everyday culture in the Netherlands in all its diversity and dynamics. In keeping with recent discussions in the social sciences and humanities it employs an internal, to some extent structuralist, as well as an external, contextualizing, approach. The latter approach, in which culture is considered from a context-bound perspective in such a way that the role of external factors is central in modelling cultural categories, is the most generally accepted discourse within the discipline. But culture can also be understood as structure, where each element derives its specific meaning from contrasts with elements elsewhere in the system, and this system can be the result of common human competences. Both perspectives are relevant to ethnology at the Meertens Institute.

Cognitive and contextual factors

The interaction between the internal and the external, between cognitive and contextual factors, plays a central role in ethnological research in the areas of folktale, song, everyday culture, heritage and national identities. Ethnologists emphasize that what is referred to as ‘popular culture’, ‘heritage’, or ‘national identity’ are always constructs. They are ‘inventions’, even though their appeal is based on the opposite, i.e. on the promise of age-old essences, of timeless individuality (‘eigenheid’). This does not explain, however, how such phenomena, in spite of their constructed character, are still convincing, how they often manage to appeal to the deepest feelings of the parties concerned. On the
external approach, then, we will also in the future be concerned with processes of appropriation and meaning assignment, including the underlying processes of assimilation and demarcation, inclusion and exclusion. But the dramatic power of cultural phenomena is also at stake here, its reverberation in the concretely physical, sensory and emotional experiences of the groups in question – ‘sensory ethnology’. Bourdieu’s notion of habitus is important here: how culture becomes ‘second nature’, is literally incorporated by groups, whether we are dealing with songs and tales, the daily contact with the objects surrounding us, or with parties, rituals, movements and gestures.

The domain of Dutch everyday culture is too extensive to be covered by a single research group. It is necessary therefore to structure the research by means of lines of approach and concrete research questions. Basing ourselves on the Institute’s past (databases, lines of research), we will put the emphasis on folktale and folksong, ritual and religion, and material culture. For one thing, we will be concerned with understanding the diversity and dynamics of everyday culture in the context of our present-day multicultural, medialized society. For another, a focus will be introduced in the research which will link the research oriented towards external factors with research into internal factors.

Putting the interplay between internal and external factors in research into everyday culture at the forefront leads to two research themes:

(iii) the relations between social, geographical and diachronic diversity in culture
Cultural phenomena migrate in time. This is a matter of change, but also of continuity. The question is how these two aspects relate to each other. Additionally, there is the question to what extent the existing diversity and the specified dynamics of cultural phenomena are dependent on internal and external factors. In other words, can diachronic diversity and diversity in social and geographical space be seen as essentially identical processes where the difference lies in the nature of the central variable (i.e. time or space)? Or are there factors that necessitate a principled distinction between social, geographical and historical diversity?

Just as in research on variation linguistics, the initial hypothesis here is that dynamics and diversity are determined by an interplay of internal and external factors. In this connection, research into the relation between synchrony and diachrony is important. In studying change it is obvious that we need to investigate which factors change simultaneously as this may reveal an underlying common factor of change.
Central research questions are:
- Some cultural phenomena spread more rapidly or ‘survive’ longer than others. To what extent can these differences in resistance be attributed to the phenomena themselves and to what extent to contextual factors?
- In what way are diachronic processes boosted by synchronic processes?
- Does contact among cultures constitute a factor in the nature or pace of cultural change, and if so, in what way?
- To what extent is contact between groups conducive to cultural change?
- Does contact among cultures lead to forms of interculturality, and if so, how do these forms of cultural dynamics manifest themselves? Questions about creolization, hybridization, ‘cultural translation’ etc. arise at this point.

(iv) boundaries of cultural diversity
The constructivist approaches in the social sciences and the humanities that came up in the seventies of the previous century were often connected with optimistic ideas about the manipulability of culture. Research into cultural diversity, in whatever society, often appeared to be a goal in itself. In our plan this approach is again adopted, but it will be complemented with more recent phenomenological or cognitivist approaches, focusing on the performativity of culture and on issues of body and mind, including the senses and the emotions. Central to these latter approaches are notions of habitus, which may be viewed as an incorporation of structure, making it to a certain extent static and binding and thus setting certain limits to cultural diversity, to the actual resonance of cultural constructions. Some forms or elements of everyday culture or heritage manage to be convincing, but others do not. A classic case in ethnology is the phenomenon of ‘folklorism’. When certain cultural phenomena are thought to be on the way out, attempts at preservation are often undertaken. But in the process the old still disappears, since as a rule only the elements experienced as colourful or dramatically appealing are preserved or, if necessary, enlarged. Research on the properties of this process yields insights into the place and meaning of constituent elements of the cultural phenomenon under investigation.

internal coherence of culture
Ethnology at the Meertens Institute therefore sees it as its brief to have an open eye for not only the diversity and dynamics of culture, but also for the internal coherence of culture, and from this perspective to contribute to the interdisciplinary scientific debate on the relation between structure and agency, ‘nature’ and ‘nurture’ continuity and change.

research questions
This issue leads to the following research questions:
- What are the geographical or group-dependent differences in the diffusion of specific traditions and rituals?
- How do groups mobilize traditions and rituals in order to establish or contest geographical or social dividing lines?
- What is the difference between cultural phenomena showing a clear geographical and social spread and others that do not?
- To what extent is the diffusion of cultural phenomena restricted by contextual factors or sometimes even advanced by them?
- What is the role of cognitive competences, senses, emotions, ‘ingrained habits’ (habitus) in the diffusion and transmission of cultural phenomena?

Cognitive science approaches to language and culture (v)

There is increasing evidence that human cognition consists of various interacting modules organized for a specific part of human behaviour. These core modules of our knowledge system involve cognitive capacities in the areas of e.g. language, number, geometry, music, visual perception, taste, emotion and ethics. The assumption is that such modules are genetically encoded and so are an inherent part of the human genome. Thus, on the basis of innate, language-specific information, DNA provides the child with the possibility of developing itself in these areas so as to become a full member of society.

Although there are enough indications for the cognitive capacities mentioned above that it is a matter of mental predisposition, there is ample discussion on the point and often fierce and acrimonious debate. For a long time language and culture have been regarded as external properties chiefly conditioned by the environment. On this view, culture and nature are opposite entities and culture is regarded as a product of ‘nurture’. The nurture view of language and culture is subject to pressure from biological, neurological and cognitive-psychological research. The neuropsychologist Dehaene for instance, shows that mathematical reasoning has a neural foundation; Rizzolatti describes empathy as a neural disposition; the behavioural biologist Hauser argues that morality contains a nature component, and the developmental psychologist Spelke argues that the interaction between innate knowledge modules allows children to organize their perceptions.

Starting from the hypothesis that language and culture are founded on related but specifically designed knowledge modules and that these knowledge modules are instantiated in interaction with variable external factors, the two
research fields of the Meertens Institute form an attractive, coherent research field. The internal line of approach focuses on the question as to what the properties are of the internal modules, viewed from a microperspective: to what extent can parts of the language system and of a number of aspects of everyday culture, which are investigated from a comparative perspective within a restricted geographical and temporal space, throw light on aspects belonging to the internal system of human cognition and to what extent are there similarities between the knowledge modules determining everyday culture and knowledge modules for language?

Language module

Linguistics at the Meertens Institute has for some considerable time been based on the notion that the grammar of a language must be viewed as a complex interplay between internal and external factors. This idea is also the basis of the current research programme of variation linguistics, as described above. Because language represents a unique human property and given the complexity of the language system, the uniformity with which children acquire their language and the low quality of the language input in the language acquisition process, it may be assumed that language is not solely determined by general cognitive strategies, but that the acquisition of language is based on a specific cognitive language module enabling the language acquisition process. It does not alter the fact, however, that it is quite possible and even likely that this module – often referred to as the language organ – bears a resemblance in its properties to other knowledge modules. It would lead to a perspective in which at the most abstract level of analysis there are general cognitive principles which are integrated in a specific manner for different cognitive aptitudes.

The internal system of culture

Contrary to research on the internal system of language, research on the internal system of (everyday) culture can only fall back on a limited amount of actual research. Research has been done in the area of music (e.g. Jackendoff, Honing), rituals (e.g. Handelman), visual culture (e.g. Freedberg) and corporality (e.g. Rizzolatti), but for other subdivisions of culture research is still in its infancy. Nevertheless, the Institute will make a part of its research capacity available in the coming period for providing further content of the perspective sketched above.

Research questions

Relevant research questions are:
- Which properties of language and culture belong to the internal system?
- To what extent are these internal properties parametrized?
- Which logically possible means of expression of language and culture are excluded by the internal system?
- Are there similarities between the internal properties of language and those of everyday culture?
- In what way can internal properties of language and/or culture be considered to be the result of one or more knowledge modules?
- Do the various knowledge modules influence each other, and if so, what are the properties of the interfaces between these modules?

The external influence on the properties of language and culture (vi)

Language and everyday culture are aspects of human behaviour; salient features of which are a large measure of diversity and hybridity that can be employed by groups in society to shape identities. It is for good reason that immigrants are supposed to learn Dutch as quickly as possible and settle down in new cultural surroundings. Being a member of Dutch society evidently implies that one has mastered these two characteristic aspects of our society. The research on the interplay of language and culture, on the one hand, and Dutch society on the other, gives an idea of and provides an insight into the ways in which groups in society manifest themselves culturally by the appropriation and incorporation of specific variable aspects of language and culture.

Important societal changes go hand in hand with changes in linguistic and cultural behaviour. In this day and age we are seeing a movement from local to regional organization and from a more monocultural to a multicultural society. We are also witnessing a continually shifting attitude vis-à-vis a construct notion such as national identity. The continually shifting social organization of society is reflected in language and everyday culture.

We note three processes in language: (i) regionalization, in which specific features of local language varieties disappear and are replaced by more general features of regional dialects; (ii) the development of ethnolects and other foreign varieties of Dutch; (iii) an extension of the ‘range’ of Dutch, i.e. substandard or ‘wrong’ language variants become part of common Dutch, enlarging the difference between Netherlands Dutch and Belgian Dutch.

Each of these three linguistic processes leads to the following research questions:
- In which linguistic features is the process visible, and in which components of the language system (word, sound, grammar)?
- What is the linguistic direction and extension of this process, and how is this accounted for by the linguistic model?
- What is the linguistic relationship between the new variety and the current standard language?
- Does the standard language change under the influence of the new variety and vice versa?
- Which external factors have an impact on the process (e.g. geography, age, gender, social group, prestige, power, religion, economy, media), how do these relate to each other, and to which distribution of the new variety in society does this lead?
- Which of the shifting aspects in this process are seen by the speakers as typical of the new situation, the new regional, social or national identity, and how are they mobilized in the construction of this identity?

For ‘foreign’ Dutch we can add the following question:
- Are there conspicuous differences between the Dutch spoken by different groups of ‘new Dutch citizens’?

The changes in standard Dutch also raise the following question:
- How does the change in standard Dutch relate to the public view on language norms?

In these three processes of change (from local to regional, multiculturality, changes at the national level) language change and changes in everyday culture go hand in hand. These two domains will therefore be investigated in their mutual relationship and interaction. The research on linguistic phenomena will take place in projects in which comparable aspects of everyday cultural behaviour are investigated. This implies that the research questions for linguists and ethnologists are formulated in such a way that the results of the subprojects can be interpreted in a common perspective. In this way this line of research sheds light on the way in which social changes in society become visible in linguistic and cultural behaviour.

For the investigation of the three types of cultural processes (regional, multiculturality, national) this approach leads to research questions which are highly parallel to the questions about language change:
- In which cultural aspect is the process visible?
- In what ways are these changes designed, consciously or unconsciously?
- What is the direction and extension of this process, and how can this be accounted for from an ethnological perspective?
- What is the relationship between new and current forms of everyday culture?
- Do the current forms of culture change under the influence of the new forms and vice versa?
- Which external factors have an impact on the process (e.g., geography, age, gender, social group, prestige, power, religion, economy, media), how do these relate to each other, and to what kind of social distribution of cultural variety does this lead?
- Which of the shifting aspects in this process are seen by the speakers as typical of the new situation, the new regional, social or national identity, and how are they mobilized in the construction of this identity?

A further question can be added for the multicultural process of change:
- Are there striking similarities and differences between the everyday culture of different groups of ‘new Dutch citizens’?

The changes in standard Dutch also raise the following question:
- How do these developments relate to the public view on national identity?

Given the similarities in phrasing the research questions on linguistic and ethnological research it is desirable to design research of this kind in common research projects in which both linguists and ethnologists are involved. A good example from the past was the project ‘Talen en Culturen in the Utrechtse Lombok en Transvaal (TCULT)’ (Languages and Cultures in the Utrecht districts Lombok and Transvaal) (1998-2005). The goal of these projects is to get a better scientific picture of the way in which linguistic and cultural phenomena take shape in a changing society.

A research plan to be developed in the coming period focuses on the comparison of the varieties of Dutch spoken in the former colonies Surinam, the Netherlands Antilles and Indonesia with the varieties of Dutch spoken by the corresponding immigrant groups in the Netherlands, in relation to the different histories and cultures of these groups and from a synchronic and diachronic perspective. The Dutch language and culture of these groups will also be compared with postcolonial immigrant groups. This research would involve cooperation between linguists and ethnologists and potentially also researchers in colonial and immigration history (KITLV). The text corpora to be collected for the historical part of the research will be integrated into the NWO groot project Nederlab.

Research themes for the period 2013-2018

As was concluded at the end of section 2.1.3 the research at the Meertens Institute is concerned with: Research into the internal and external properties of immigrants and emigrants changing society
coopration with other KNAW-institutes
of the diversity and dynamics in the language and culture of the Netherlands. Since dynamics presupposes a diachronic perspective, this description of the research links up directly with the first main area mentioned in the KNAW-Contourennota: Nederlandse cultuur, geschiedenis en taal (Dutch culture, history and language). In the coming period consultations will be started on how cooperation with other KNAW-institutes on this research theme can be realized, and which KNAW-institutes will participate in this. This might lead to adaptations in our research programming and changes in the prioritization of research lines.

**six research themes**

In the period 2013-2018 an important part of the research will be organized in the framework of the six research themes described in the previous sections:

i. Boundaries of language variation

ii. The relationship between diversity and dynamics in language

iii. The relations between social, geographical and diachronic diversity in culture

iv. Boundaries of cultural diversity

v. Knowledge and cognition: language and culture from an internal perspective

vi. External aspects of language and culture

These themes have been selected as the Institute’s research themes on the grounds that they emphasize an internal or external dimension and/or highlight the interplay between internal and external.

**thematic organization**

In the coming period research will no longer be determined individually on the basis of the capacities of a researcher who has been appointed in a specific position. In the previous period researchers were appointed as experts in a particular (sub)field, e.g. a phonologist working within the framework of sociolinguistics or an ethnology researcher specialized in folktales. As of January 1, 2013 the majority of the research will be thematically organized, with the consequence that researchers are appointed on a temporary basis in one or more ongoing projects in one or more research themes. These themes will have a theme leader who will take responsibility for continuity and consistency of the research that is carried out within the projects in as far as it is relevant for his/her theme.

**research projects**

As will become clear in the next chapter, the suggested research lines are outlined as coherent research projects with a substantive description and a concrete result, with, among other things, a description of subprojects, a time scheme, a staffing plan, a budget, and a description of the expected output.
(publications, databases, acquisition of external funding, etc.). All of this holds for externally financed projects as well as internal projects.

The external projects were submitted among others to NWO and European funds, given the relevance of their contribution to answering the central research questions discussed in earlier sections. As for their content and organization, the projects are directed by the project submitter/leader.

The internal projects accepted by the Institute for a maximum of five years after submission, evaluation and discussion have a project leader who is responsible for progress, is a member of the Board of Research, and is the contact person for the Institute’s management.

The transition has come into effect in 2012. The Institute has gone through a development from a structurally defined organization to a thematically defined organization. In this period teams of researchers worked out thematic projects and submitted them to the management for acceptance. The management designed procedures for the way in which evaluation and selection of projects take place – with the help of external experts- and for the way in which current projects are monitored and evaluated.

In the next chapter we provide an overview of ongoing and starting research projects. These projects are directly related to the research themes that have been introduced in this chapter. Together they constitute 85% of the research activities in the Meertens Institute for the coming period.
List of projects

1. Oral Culture
2. Maps and Grammar
3. Life cycle of liquids
4. Knowledge & Culture
5. Nederlab
6. The Language Portal
7. The 'Dutchness' of the Dutch
8. The construction of local identities through cultural practices
9. Invariants in Language Change

Research projects

3.1 2013-2018
3.2 Structural research staff/formation/research in projects
3.3 Research themes and research projects
In this chapter we provide a list of ongoing projects that presents a more precise overview of the Institute’s research activities in the coming period. In addition, we will indicate how the research staff is divided over these research projects.

List of projects 2013-2018

Below we present a list of ongoing projects and projects that will start at or before 1.1.2013. The list combines projects accepted through the internal procedure (see above; projects Ia, II, III, VII, VIII and IX) and externally financed projects ( Ib-e, IV, V and VI). It will be seen that at least 85% of the research volume of the structural staff will be invested in these (internally and externally financed) projects in the coming three years. However, there is room for one or two additional projects. We expect a new project to start as the consequence of the collaboration of KNAW-institutes, and there are several grant-applications under review. Together these projects constitute the Institute’s research programme for the coming years.


In the field of oral culture – folk songs and folktales – a number of coherent projects are brought together. Tenured Meertens Institute staff take part in the overarching internal project Oral Transmission accepted in 2011. Furthermore, there are three externally financed research projects employing temporary staff and in collaboration with external partners: Tunes & Tales (KNAW-Computational Humanities), fact (NWO-CATCH) and cogitch (NWO-CATCH). Together these four projects constitute the centre of expertise eLab Oral Culture (see http://www.elab-oralculture.nl). The oral culture group is also involved in various smaller projects and projects mainly dealing with the digitalization of collections of songs and stories, among which Bellenspeelklokken (The music of eighteenth-century bell-playing clocks in the Netherlands; Meertens Institute and Museum Speelklok, 2009-2013), Dutch Songs online (NWO-middelgroot, 2009-2013) and TINPOT (KNAW, UVA, VU Center for Digital Humanities, 2012-2013). These projects are not included in the list below. A related activity involves participation in a subproject of Knowledge & Culture (NWO-Horizon; see project IVa below). In addition, the oral culture group maintains two large digital collections: the Dutch Song Database (Nederlandse Liederenbank; http://www.liederenbank.nl) containing metadata of 150,000 Dutch songs,
including an increasing number of song texts, music notation and audio, and the
Dutch Folktale Database (Nederlandse Volksverhalenbank; http://www.verhalenbank.nl) containing the texts of 42,000 Dutch and Frisian folktales including metadata. These databases are used for all research projects described below.

la Oral Transmission (1.9.2011 - 1.9.2016)
   ○ project leader: Louis Grijp (0.6)
   ○ research staff: Theo Meder (0.6)
   ○ support staff: Martine de Bruin (1.0), Maarten van der Peet (0.4), Ellen van der Grijn (0.6), Marianne van Zuijlen (0.8 until 1.4.2013)
   ○ internal project

The central problem in the overarching project on oral culture – both songs and folktales – is the mechanism of oral transmission as it works in artefacts such as folk songs and folktales. Our questions deal with

• Variation:
  - To what extent can a tune or tale vary due to oral transmission?
  - Are there stable unvarying elements?
  - Are there convergent forces at work compensating for the divergence caused by variation?
  - Can we build a model for the process of oral transmission accounting for the variability?

• Recognizability:
  - How is it possible that we still recognize a song or tale after several generations of oral transmission?
  - Are all descendants of one original tune or tale recognizable?
  - What are the criteria for recognizing a tune or tale?

We expect to find structural and cognitive elements (such as motifs, memes and ‘hooks’) that can both explain the nature of variability and the recognizability within variation.

From the central problem of the overarching project Oral Transmission we have deduced partial questions to be answered in the projects Tunes & Tales, FACT and COGITCH described below. In our approach, computational methods will be combined with predominantly structuralist musicological and narratological methods.

The Dutch Song Database (especially the field recordings of folk songs by Ate Doornbosch and others) and the Dutch Folktale Database will be the fore-
most suppliers of data and metadata. For the ICT technology see the projects Ib-d below.

Ib Tunes & Tales (1.1.2012 - 1.5.2016)
- project leaders: Louis Grijp, Theo Meder
- temporary research staff: Peter van Kranenburg (postdoc, 1.0), Berit Janssen (PhD, 0.8), Folgert Karsdorp (PhD, 1.0)
- support staff: Martine de Bruin, Ellen van der Grijs, Marianne van Zuijlen
- temporary technical staff: vac.
- external staff: Henkjan Honing (UVA, 0.1), Frans Wiering (UU, 0.1), Remco Veltkamp (UU, 0.05), Mariët Theune (UT, 0.05), Franciska de Jong (UT, 0.05), Antal van den Bosch (RU, 0.1), Arjen Versloot (FA, 0.1)
- project: Computational Humanities (KNAW), €27.000
http://www.elab-oralculture.nl/tunestales

Tunes & Tales is a project on oral culture, using both Dutch songs and Dutch and Frisian folktales.

Central research questions are:
- How can we model the process of oral transmission?
- Can we analyze folk tunes and folktales as motif sequences?
- Can motifs be used for recognition and retrieval of artefacts?

The hypothesis is that tunes and tales both consist of motif sequences acting as vehicles of oral transmission. We expect to be able to develop a model based on these motif sequences, which can predict the occurrence of variable motif patterns in oral tradition over time.

First, tunes and tales will be formalized as layered sequences of motifs. This task will be carried out by the two PhD students (one for tunes and one for tales), and will result in prototype systems for the automatic recognition of motifs in tunes and tales. This important first step will enable us to analyze the large amounts of available data for creating models of the mechanism of oral transmission. Additionally, automatic recognition of motifs can be used for automatic classification of the corpora.

On the basis of the automated analysis of motif sequences, the postdoc will create generative models simulating the oral transmission including the inherent variation. These models will be tested on the oral material at the Meertens Institute in cooperation with domain experts of the Institute. The analytical methods to be pursued include process mining, association rule mining, gram-
mathematical inference, genetic algorithms, case-based reasoning, iterated learning and constraint-based modelling.

The basis for our analyses are transcriptions of field recordings by Ate Doornbosch from the Dutch Song Database and all Dutch and Frisian tales from the Dutch Folktale Database. The application of ICT technology will include techniques from information retrieval such as tf-idf, natural language processing, machine learning, and named entity recognition.

**FACT (1.1.2012 - 1.5.2016)**
- project leader: Theo Meder
- temporary research staff: Dolf Trieschnigg (postdoc, 1.0*), Dong Nguyen (PhD, 1.0*)
- support staff: Marianne van Zuijlen
- temporary technical staff: Iwe Muiser (1.0*)
- external staff: Mariët Theune (UT), Franciska de Jong (UT), Antal van den Bosch (RU)
- project: NWO-CATCH, €55,000
  http://www.elab-oralculture.nl/fact
  *seconded to the Meertens Institute by UT

The central concern of CATCH-project FACT is the analysis of folktales. FACT (Folktales as Classifiable Texts) aims at developing software for automatically enriching a corpus of Dutch folktales with metadata such as language, names, keywords, genres, summaries and catalogue numbers (or folktale types). These tools will speed up the digitization process, turning the Dutch Folktale Database into a sophisticated online heritage archive as well as a research tool.

Important research questions involve recognition in oral transmission:
- Which tools work best in folktale annotation as far as languages, names, keywords, genres, summaries and folktale types are concerned?
- How do people recognize tales that have undergone variation during oral transmission, and will the computer be able to simulate this in any way?
- Which strategies are used in human and computational classification (or: which criteria do people and computers use to recognize specific tales)?
- Will computer clustering be able to improve human classification of folktales?

We expect automatic corpus analysis to offer new possibilities in the classification of folktales by means of clustering. In this way we hope to find new and surprising relationships between groups of tales – either in structure or content.
The algorithms developed in the project will be integrated into user-friendly tools in order to support annotation and exploratory research of the folktale corpus. With these tools we can research variability in oral and written transmission, as well as the pros and cons of human classification and computerized clustering.

The data to be used for training purposes will consist of the 42,000 folktales present in the Dutch Folktale Database today. In order to evaluate the results on new folktale material, new collections of folktales are being gathered and digitized by Marianne van Zuijlen. Furthermore, the metadata in the Dutch Folktale Database have to be standardized, and the possibilities to add metadata need to be more uniform in the future. Hence, the database is in need of spring-cleaning and updating to ensure consistent metadata generation and more accurate analysis in the future.

The ICT technology used in FACT will include information retrieval, natural language processing, machine learning, named entity recognition, document indexing, text summarization, document classification, clustering and visualization tools.

Id  cogitch, project on musical hooks and interoperability (1.1.2012 - 1.5.2016)
  ◦ project leader: Meertens Instituut: Louis Grijp
  ◦ temporary research staff: Ashley Bugoyne (postdoc, UVA), Jan van Balen (PhD, 1.0*)
  ◦ temporary technical staff: Dimitrios Bountouridis (1.0*)
  ◦ support staff: Martine de Bruin, Ellen van der Grijn
  ◦ external staff: Henkjan Honing (UVA), Frans Wiering (UU), Remco Veltkamp (general project leader; UU)
  ◦ project: NWO-CATCH, €55.000
    http://www.elab-oralculture.nl/cogitch
  *seconded to the Meertens Institute by UU

COGITCH (COgnition Guided Interoperability beTween Collections of musical Heritage) is a CATCH-project on traditional music/songs and early 20th century popular music. COGITCH’s general objective is to develop generic techniques for indexing distributed music sources by means of an interoperable system. As this is about audio, a search engine such as WITCHCRAFT, based on music notation, cannot be used. Instead, we investigate the cognitive notion of ‘hook’ (the most salient, memorable, and easy-to-recall moment of a musical phrase or song) as a means of recognition/retrieval. The hooks must be annotated manually, a task to be performed by crowd sourcing. Thus, COGITCH involves three intertwined strands of research objectives:
- We design and develop a novel infrastructure to let listeners collectively provide annotations and to derive cognitively relevant features.
- Based on these cognitively relevant features we invent and implement new music thumbnail extractors and music similarity methods.
- Using content-based retrieval methods, a generic interoperable search infrastructure is designed and implemented for accessing the audio collections of the Institute for Sound & Vision in Hilversum (popular music) and the Meertens Institute (folk music).

We assume that for audio retrieval, cognition may provide an alternative for the usual data-based approach. In order to use hooks for retrieval ends, we assume that the cognitively relevant features can be mapped onto computational models. We expect that connecting the audio collections in Hilversum and Amsterdam will make it possible to study the complex relationship between traditional (folk) music and modern (popular) culture in the Netherlands in the first half of the 20th century.

In developing retrieval methods, we will take a top-down approach, working from musical knowledge and cognitive psychology towards the identification and processing of audio features. A web-based, so-called ITCH environment (Identification, Tagging and Characterisation of Hooks) will be designed and constructed to obtain large amounts of judgements from the lay audience on what makes a fragment of music easy recognizable and/or stick in one’s mind (annotation of hooks by crowd sourcing). The collections of Sound & Vision and the Meertens Institute will be made operable. An important problem is that very few of the early 20th century recordings in Sound & Vision have been digitized. A number will be selected and digitized as part of the project.

The Institute for Sound & Vision (Hilversum) possesses a unique audio collection of early 20th century popular Dutch music, which is complementary to the Meertens Institute’s collection of traditional Dutch folk songs. Both institutions wish to make these musical archives accessible in an integrated way to the general public as well as to scholars.

II Maps and Grammar (1.1.2013 - 1.1.2017)

- project leader: Sjef Barbiers (0.4)
- research staff: Hans Bennis (0.2), Gertjan Postma (0.3), Marc van Oostendorp (0.1)
- temporary research staff: 1 PhD (if NWO-programme is granted, an additional postdoc and 2 PhDs)
In recent years we have observed an increasing interest in the linguistic properties of language varieties and dialects and a fair number of large scale dialect databases has become available for syntactic and phonological variation research, including some atlasses showing the geographic distribution of linguistic features.

However, a crucial issue remains to be investigated, namely the question as to whether the concepts ‘dialect’ and ‘regional language variety’ are contentful from a grammatical perspective, and if so, in what way. Is there any linguistic substance to the division of a language area into smaller components?

Four ways are used to approach this issue:
- First of all, we will aggregate the existing empirical material computationally to distinguish relatively stable, homogeneous linguistic regions.
- Secondly, we will approach the topic from a more qualitative linguistic angle by trying to establish inclusion relations in the linguistic system which will give clues about cooccurrence patterns that might be typical for determining linguistic regions.
- Thirdly, we will look for geographical areas showing a relatively unstable and heterogeneous linguistic situation. This instability may indicate that these areas are transition zones between language varieties.
- Fourthly, we will investigate whether the different regions identified in this research correspond to different grammars, i.e. abstract coherent rule or constraint systems underlying the attested (morpho-)syntactic and (morpho-) phonological patterns.

Together, these four perspectives should enable us to uncover the grammatical landscape underlying the observed variation. The divisions of the map based on inclusion relations, transition zones and grammatical systems are expected to differ from the divisions resulting from the quantitative (or aggregational) method.

We will use, among others, the large datasets on dialects collected recently, such as SAND and GTR for Dutch, and the databases made accessible with the Edisyn search engine for other dialect families. The aggregational method and the discovery of inclusion relations require cluster analysis and advanced mapping algorithms, many of which are already available in modern dialectometry. One of the goals is to develop these tools further; in particular we would like to develop a tool to visualize three-dimensional linguistic landscapes.

- external research staff: Roberta d’Alessandro (UL), John Nerbonne (RUG)
- internal project (external grant proposal submitted (NWO-programme))
III  Life cycle of liquids (1.1.2012 - 1.1.2016)

- project leader: Marc van Oostendorp (0.2)
- research staff: Ben Hermans (0.3), Frans Hinskens (0.4)
- temporary research staff: Etske Ooijevaar (PhD)
- internal project

This project studies the sociolinguistic and phonological make-up of the consonants /l/ and /r/ (together known as liquids), both in contemporary varieties of Dutch and in its diachronic development. In coda position the liquids of Standard Dutch and many Dutch dialects are undergoing rapid changes, a process that has been underway for a several decades already. Moreover, centuries ago the vocalization of /l/ following non-high back vowels has been lexicalized in most varieties (English old corresponding to Dutch oud, cold to koud).

In this project we aim at investigating how these changes can be understood in terms of the interaction between E-language (phonetics, sociolinguistics, language transmission) and I-language (phonology). In particular, the following two questions are central: (a) how can the life cycle of changes like these be understood (moving from blind phonetics to the lexicon via productive phonology), (b) what can we learn from these processes about the phonetic grounding of phonological primitives: are these primarily acoustic/perceptual or articulatory?

The main hypotheses for each of the two research questions just mentioned are:
(a) We expect to find confirmation for Kiparsky’s (2010) ‘amphichronic’ model of phonology in which diachronic processes and synchronic phonological constraints have separate, clearly distinguishable functions (and where phonetic naturalness and sociolinguistic factors play a clear role in the former; with phonological theory playing a role in the latter). Furthermore, we also expect to find confirmation for Labov’s (2007) distinction between transmission and diffusion.
(b) At the beginning of the project, there seems to be evidence pointing both in the direction of an articulatory and of a perceptual representation of liquids. It can only be expected then that the ideal model should somehow combine the best properties of both, although how exactly this will work out is as yet unknown.

This project will use a modern type of fieldwork that combines advanced phonetic measurements (such as ultrasound) with sociolinguistic methodology (specifically for the speaker design and the techniques used for collecting data). So far, these combinations have only been tried out experimentally in the framework of usage-based approaches (such as sociophonetics), but we
intend to interpret the data instead in a more rule-based approach to phonology or in a hybrid model, dovetailing either; which is yet to be developed. Furthermore, we will use electronic sources, including those made accessible in Nederlab, for studying the historical sources, which will get an interpretation in the same framework.

The collected data will be of three types: phonetic measurements (both acoustic and articulatory), sociolinguistically rich and diverse data (partly collected with the aid of questionnaires), and historical corpus data. We will need to put together a database to make the first two types of data, and interface them with the third type.

IV Knowledge & Culture (1.1.2013-1.1.2017)

- NWO-Horizon
- project leader: Johan Rooryck (Leiden University), in collaboration with Edward de Haan (University of Amsterdam) and Hans Bennis

In various domains of cognitive science – psychology, mathematics, cognitive neuroscience, linguistics, music cognition, and biology – a new paradigm is being developed that will profoundly affect research in the humanities. This emergent view holds that humans and non-human animals are born with core knowledge systems: a small set of hard-wired cognitive abilities that are task-specific, language-independent, and non-species-specific. The ‘core knowledge’ paradigm challenges scholars in the humanities to ask the question which parts of human culture belong to nature, and how nurture and culture build on nature. In this research program, four domains of the humanities will be investigated from the point of view of core knowledge: music cognition; language and the core knowledge of number; visual arts and geometry; and poetry, rhythm, and meter. Two of these projects (Language & Number and Poetry, Rhythm & Meter) will be led by the Meertens Institute. In a third subproject (Music cognition) researchers from the Meertens Institute are involved.

IVA Knowledge & Culture: Subproject 1 – Music cognition, language, and core knowledge systems (1.1.2013 - 1.1.2017)
- subproject of NWO-Horizon project Knowledge & Culture
- project leader Meertens Instituut: Louis Grijp
- research staff: Marc van Oostendorp (0.05)
- temporary staff: 1 postdoc, 1 PhD
- external staff: Henkjan Honing (UvA), Johan Rooryck (UL)
Music and language have a lot in common. Both are rule-governed and subject to recursion, the ability to indefinitely lengthen sentences or musical pieces. Infants acquire both music and language effortlessly, suggesting an innate ability. Predispositions for music are related to those abilities that distinguish humans from other animals and emerge prior to birth (Trehub 2003; Zentner & Kagan 1996).

In this subproject of Knowledge & Culture we raise the following central questions:
- What is shared (and what is unique) in music and language?
- How can the cognitive mechanism of relative pitch throw light on the relation between music and language and musical ability?

The first question will be the subject of a postdoc-project which will investigate the nature of the relationship between music and language. This researcher will attempt to integrate the insights from both research traditions introduced above, the theoretical-linguistic tradition and the developmental-experimental studies, and combine this with recent insights from the field of music cognition (cf. Honing 2011).

The second question will be addressed in a PhD-project and raise the following research questions:
- What is the evidence for relative pitch as an innate or at least a spontaneously developing skill?
- Is relative pitch shared with language, or do we have two different pitch perception systems for music and language?
- What is the role of relative pitch in cultures with a tonal language? How does the semantics marked by pitch interact with the constraints of music?
- Is there a relation between relative pitch and the core knowledge systems of number or geometry? If so, what is its nature?

**IVb Knowledge & Culture: subproject 2 – Language and the core knowledge system of number (1.1.2013 - 1.1.2017)**

- subproject of NWO-Horizon-project Knowledge & Culture
- project leader subproject: Sjef Barbiers (0.1)
- temporary research staff: 1 postdoc, 1 PhD
- external staff: Pierre Pica (CNRS, Parijs), Fred Weerman (UVA), Johan Rooryck (UL)

This project investigates the mapping between the core knowledge system for number and the faculty of language. The central research question is how the
linguistic system and the number systems determine each other. Recent research shows that the core knowledge system of number consists of two nonverbal subsystems present in both newborn infants and animals (Spelke 2011).

The first subsystem, the *Approximate Number System* (ANS), compares the numerosities of distinct sets without individuating their members. This system is imprecise and ratio-sensitive. Initially, it only distinguishes between sets with a ratio of 2 or higher (e.g. 4 vs. 8), the ratio decreasing with maturation (Dehaene 1997). The second subsystem, the *Object Tracking System* (OTS), yields representations of small numbers of objects, i.e. 1 to 3 (perhaps 4). It does not work on sets larger than 4. OTS is also different from ANS in that it is precise, able to distinguish between 2 and 3, and sensitive to adding or subtracting one individual.

The project addresses the following research questions:

(1) a. Does the distinction between OTS and ANS manifest itself in the numeral systems of the languages of the world, and if so, how?
   b. What does the special linguistic status of 1 as opposed to all other numerals (cf. Borer 2005, Barbiers 2007) tell us about OTS and ANS?

(2) a. What is the sequence of acquisition for (the morphosyntactic properties of) 1, the other numerals, and quantifiers such as MANY?
   b. How are words for ordinals acquired by children?
   c. How is the numeral for 2 acquired by children?

The primary data source for the first study is the *World Atlas of Linguistic Structures* (Dryer & Haspelmath 2011, Comrie 2011; http://wals.info/), which currently contains 2678 languages in 510 genera and 212 families, including relevant features for ordinal numerals (321 languages) and order of numeral and noun (1154 languages). Starting from these data we will select at least one language from as many genera as possible, completing the inventory with reference grammars and data elicitation. The data thus collected will be made available to WALS and other systematic collections of linguistic properties such as Syntactic Structures of the World’s Languages (sswl.railsplayground.net). The data will be analyzed for their contribution to the theory of morphosyntax.

The data for the second study will be collected in mixed longitudinal and cross-sectional elicitation, perception, and judgment experiments involving children from age 2 until (at least) 8.

IVc  Knowledge & Culture: subproject 4 – Poetry, rhythm, and metre
(1.1.2013 - 1.1.2017)
- subproject of NWO-Horizon-project Knowledge & Culture
- project leader subproject: Marc van Oostendorp (0.1)
- research staff: Louis Grijp
Rhythm is a form of temporal organization that plays a role in any systematic human activity involving sound: poetry, non-poetic language, and music. In music, rhythm is considered one of the dimensions of musical structure, along-side melody and colour (Honing et al. 2009; Honing 2009). In natural language, rhythm is a property that distinguishes types of languages (French is 'syllable-timed', English 'stress-timed'). In poetry, the metre of a verse can be described as a regular alternation of prominent and less prominent syllables, where prominence is defined differently in different poetic traditions.

The questions to be addressed in this subproject are the following:

a. What are the invariant, universal aspects of poetic and phonological metre, once we abstract from language-specific variables?

b. To what extent is metre a rhythmic ability that is properly linguistic or part of a more fundamental human ability that is shared with music?

c. How can the universal aspects of metre throw light on a possible cognitive ability of rhythm? Is the connection to the core knowledge system of number, as outlined above, real?

V  Nederlab (1.1.2013 - 1.1.2018)

Language and culture are inherently dynamic phenomena. Scholars in the humanities – linguists, literary scholars, historians – try to understand processes of change and variation and to uncover their internal and external causes. Large quantities of historical data are needed to answer the research questions relevant to these fields of study. Until now, this research has necessarily consisted of case studies of limited scope, covering relatively short time periods and single disciplines. Nederlab will develop a comprehensive research corpus in the form of an infrastructure for the humanities in which Dutch
texts are brought together in digitized form. It will allow humanities researchers to formulate new research questions, and to set up systematic research into the interaction of changes in culture, society, literature and language across the full history.

A group of experts in diachronic research, corpus and infrastructure construction, search technologies and tool development collaborate in the development of a user-friendly web interface enabling a distributed search of available diachronic corpora, at text and metadata level, which will lead to better access to existing corpora, as well as to quality improvement and standardization of data and metadata. The infrastructure will be set up in close consultation with the research community, and will be tested in concrete research pilots.

VI The Language Portal (1.1.2011 - 1.1.2016)

- project coordinator: Ton van den Wouden (0.8)
- research staff: Hans Broekhuis (0.8), Marc van Oostendorp (0.1), Hans Bennis (0.1)
- temporary research staff: Björn Kohnlein/Kathrin Linke (postdoc)
- external staff: Geert Booij (UL), Arjen Versloot (UVa/FA), Carole Tiberius (INL), Norbert Corver (UU), Jenny Audring (postdoc UL), Crit Cremers (UL), Eric Hoekstra (FA) a.o.
- project: NWO-groot, €2.600.000

The Language Portal develops a virtual language institute in which all available grammatical knowledge of Dutch and Frisian is brought together and presented in a coherent manner. It includes sources such as the Modern Grammar of Dutch (Syntax of Dutch) that have been completed and adapted to the portal. It develops new grammatical devices based on existing sources (Morphology of Dutch and Phonology of Dutch), and it creates related devices for Frisian (and Afrikaans). These subsystems of the grammars of Dutch and Frisian are brought together in an innovative virtual portal that allows researchers to assess all aspects of these grammars within a well-organized and consistent digital system. It will substantially further research on the Dutch and Frisian languages. The Meertens Institute is the centre of this project in which other institutes (INL, Frisian Academy) and universities (Leiden, Utrecht) are involved.

- project leader: Herman Roodenburg (0.4)
- research staff: John Helsloot (0.3), Peter Jan Margry (0.3), Irene Stengs (0.3), Marc van Oostendorp (0.05)
- temporary research staff: 1 PhD
- external staff: Hester Dibbits (Reinwardt Akademie), Jan Willem Duyvendak (UvA), Paul Mepschen (UvA), Piek Vossen (VU), Marleen de Witte (VU)
- internal project

This project investigates the rediscovery of ‘Dutchness’ – all the recent constructions (political, intellectual, artistic or commercial) in which what is felt as ‘typically’ or ‘authentically’ Dutch is again promoted, celebrated or commercially exploited. Partly related to the initiatives of several Dutch governments to strengthen the nation after ‘nine eleven’ and the political murders of Pim Fortuyn and Theo van Gogh, many new constructions of ‘Dutchness’ have emerged in political, intellectual, artistic and commercial circles, including those of the fashion industry and the creative industry. Interestingly, similar developments have been observed in, for instance, Germany, England, Denmark, and a number of other European countries. In this interdisciplinary and comparative project questions will be asked about how these recent constructions actually emerged, about the constructions’ grammars of identity/alterity (Bauman), their problematic aspects of autochtony and nativism (Geschiere, Duyvendak), and how they are actually related to more general processes of globalization, mediatization and the present emotion industry. Finally, questions will be asked about the emotional and sensory hold of the constructions. How are they appropriated, often literally incorporated? Here notions of habitus (Bourdieu) may be helpful but also the notions of ‘style’ and ‘aesthetics’ as recently developed in religious studies (Meyer).

Most of the project will be based on fieldwork, archival research and interviews/questionnaires. Part of the project will also consist of ‘sentiment mining’ in the records of the Dutch parliament over the last ten years. The goal is to semi-automatically create wordnets through pattern recognition for the parliament’s vocabulary expressing emotions related to ‘Dutchness’.
The proposed research focuses on cultural practices as the locus of identity formation, because language and culture are indispensable resources for people to construct social relations and project social identities. The project addresses the interrelated questions as to how local identities are shaped through language and cultural practices and how processes of local identity formation are driven by power asymmetries between people living in the centre and periphery. These power asymmetries can, but need not be, ‘real’ in the sense that they can be measured in, e.g., economic strength. What is important is that people experience a difference between centre and periphery and that those inhabiting the (perceived) periphery feel marginalized economically, politically, culturally, and linguistically by the (perceived) centre. A basic concern is to investigate how different actors (individuals actors as well as collective actors) engage with these power dynamics and how they make use of linguistic and cultural resources in the formation of local identities.

Specific research questions that will be addressed are:
- How is a ‘peripheral’ place such as Limburg imagined and constructed as different by its inhabitants and ‘outsiders’ and what is the role of old and new media in this process?
- Which linguistic and cultural resources do speakers draw on to express and/or construct self and other?
- Which language varieties and linguistic elements are perceived as uniquely local, and are valued as quintessential features of local identities?
- What are the intertwined perceptions on local languages as threatened and treasured heritage, as vehicles of oral traditions and expressions, and as markers of identity?
- When does ordinary language become a performance of linguistic uniqueness?

We expect that the social significance of the centre-periphery dynamics can be identified not only at the national level but also, recursively, at ‘lower’ levels, within regions of the nation (Limburg, the Randstad), and even within cities and smaller local entities (cf Eckert 2011, Gal & Irvine 1995).

In order to fruitfully address these research questions, a corpus of ethnographic audio and video recordings of linguistic and cultural performances will be assembled (and has been collected in the recent past). In addition to products of local print media, radio and television networks, the project will also study relevant practices of computer-mediated communication.

The data should be stored and archived but also be made searchable. The linguistic data should be developed such that linguistic analysis can apply to it (tagging, searchable on linguistic elements (lexicon, phonology, morphology and syntax).

IX  Invariants in Language Change (1.1.2013 – 1.1.2016)

- project leader: Gertjan Postma (0.3)
- research staff: Ben Hermans (0.3)
- temporary research staff: 1 PhD
- external staff: Andrew Nevins (UC London), Tobias Scheer (U-Nice), Edward Keenan (UCLA)
- internal project under review

The quest for invariants is pursued in Keenan and Stabler (2005) from a cross-linguistic point of view and is focused on the semantic-syntax interface. The advantage of this choice is that invariants are easily accessible. However, it also presents a problem as the invariants may be derived from the universal properties of the semantic module alone, rather than from invariants in the computational system of the Grammar. In this project, we approach the quest from a diachronic point of view and concentrate on the (morpho)syntax-phonology interface, a territory in which the Meertens Institute is specialized. At the same time, our choice provides a more challenging perspective since at this interface variation is notorious and omnipresent.
The available research formation indicated in column I below does not include (a) the 0.2 fte position for the extraordinary professors Barbiers (UU), Bennis (UVA), Cornips (UM), Grijp (UU), Hinskens (VU), Margry (UVA), van Oostendorp (UL) and Roodenburg (VU) and (b) 0.2 fte for all the members of the research group for tasks not or indirectly related to the research projects (committees, meetings, technology, collections, reviews, valorization etc.).

<table>
<thead>
<tr>
<th>researcher</th>
<th>I</th>
<th>projects</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sjef Barbiers</td>
<td>0.6</td>
<td>Maps &amp; Grammar</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Knowledge &amp; Culture</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nederlab</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Hans Bennis</td>
<td>0.4</td>
<td>Maps &amp; Grammar</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nederlab</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Language Portal</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Hans Broekhuis</td>
<td>0.8</td>
<td>Language Portal</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Leonie Cornips</td>
<td>0.6</td>
<td>Local Identities</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Louis Grijp</td>
<td>0.6</td>
<td>Oral Transmission</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>John Helsloot</td>
<td>0.6</td>
<td>Dutchness</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Ben Hermans</td>
<td>0.8</td>
<td>Liquidae</td>
<td>0.3</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nederlab</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Invariants</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Frans Hinskens</td>
<td>0.6</td>
<td>Liquidae</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Peter Jan Margry</td>
<td>0.8</td>
<td>Dutchness</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Theo Meder</td>
<td>0.8</td>
<td>Oral Transmission</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Marc van Oostendorp</td>
<td>0.6</td>
<td>Knowledge &amp; Culture</td>
<td>0.15</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maps &amp; Grammar</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquidae</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Language Portal</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dutchness</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Herman Roodenburg</td>
<td>0.6</td>
<td>Dutchness</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Irene Stengs</td>
<td>0.8</td>
<td>Dutchness</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local Identities</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Gertjan Postma</td>
<td>0.8</td>
<td>Maps &amp; Grammar</td>
<td>0.3</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nederlab</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Invariants</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Nicoline van der Sijs</td>
<td>0.8</td>
<td>Nederlab</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Ton van der Wouden</td>
<td>0.8</td>
<td>Language Portal</td>
<td>0.8</td>
<td>0.8</td>
</tr>
</tbody>
</table>

I = available research time
II = research time invested in specific projects
III = total research time invested in projects
Structural research formation = 11.0 fte
In research projects = 9.5 fte

Temporary research formation
(estimate for 1.6.2013), excl. 20% for non-research tasks:
- internally financed PhDs = 3.2 fte
- externally financed PhDs = 4.6 fte
- externally financed postdocs = 3.2 fte
  total = 11.0 fte

Research projects 2013 - 2018:

<table>
<thead>
<tr>
<th></th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>la</td>
<td>Oral Transmission</td>
<td></td>
<td></td>
<td></td>
<td>(1.9.2016)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tunes &amp; Tales</td>
<td></td>
<td></td>
<td></td>
<td>(1.5.2016)</td>
<td></td>
</tr>
<tr>
<td>lb</td>
<td>Fact</td>
<td></td>
<td></td>
<td></td>
<td>(1.5.2016)</td>
<td></td>
</tr>
<tr>
<td>lc</td>
<td>Cogitch</td>
<td></td>
<td></td>
<td></td>
<td>(1.5.2016)</td>
<td></td>
</tr>
<tr>
<td>id</td>
<td>Maps &amp; Grammar</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2017)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>Life cycle of liquids</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2016)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Knowledge &amp; Culture</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2017)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>Nederlab</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2018)</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>The Language Portal</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2016)</td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>Dutchness</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2017)</td>
<td></td>
</tr>
<tr>
<td>VII</td>
<td>Local identities</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2017)</td>
<td></td>
</tr>
<tr>
<td>VIII</td>
<td>Invariants in Language Change</td>
<td></td>
<td></td>
<td></td>
<td>(1.1.2016)</td>
<td></td>
</tr>
</tbody>
</table>
Internal vs external means for research & technology

In the period 2010-2012 the percentage of externally financed projects for research and related technology is in between 35% in 2010 and 54% in 2012. For 2013 we expect the percentage to be 45%. Our target for the coming period is to be able to finance at least 40% of our research activities and technology by external means.

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2011</th>
<th>2012 (prognosis)</th>
<th>2013 (budget)</th>
</tr>
</thead>
<tbody>
<tr>
<td>internal (lumpsum) financing of research &amp; technology</td>
<td>1.439.000</td>
<td>1.374.000</td>
<td>1.410.000</td>
<td>1.475.000</td>
</tr>
<tr>
<td></td>
<td>65%</td>
<td>47%</td>
<td>46%</td>
<td>55%</td>
</tr>
<tr>
<td>external financing of research &amp; technology</td>
<td>780.000</td>
<td>1.541.000</td>
<td>1.692.000</td>
<td>1.212.000</td>
</tr>
<tr>
<td></td>
<td>35%</td>
<td>53%</td>
<td>54%</td>
<td>45%</td>
</tr>
<tr>
<td>total R&amp;T</td>
<td>2.219.000</td>
<td>2.915.000</td>
<td>3.102.000</td>
<td>2.687.000</td>
</tr>
<tr>
<td>management, collections &amp; general affairs</td>
<td>1.577.000</td>
<td>1.468.000</td>
<td>1.482.000</td>
<td>1.336.000</td>
</tr>
</tbody>
</table>

Research themes and research projects

Although each project is related to different research themes (see chapter 2), the table below lists them in a single theme-category, the major focus of the projects in question.

<table>
<thead>
<tr>
<th>theme</th>
<th>projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>i boundaries of language variation</td>
<td>II, VI, IX</td>
</tr>
<tr>
<td>ii relation between dynamics and diversity in language</td>
<td>III</td>
</tr>
<tr>
<td>iii relation between social, geographic and diachronic diversity in culture</td>
<td>VII</td>
</tr>
<tr>
<td>iv boundaries of cultural diversity</td>
<td>Ia, Ib, Ic, Id</td>
</tr>
<tr>
<td>v knowledge and cognition: language and culture</td>
<td>IVa, IVb, IVc</td>
</tr>
<tr>
<td>vi external aspects of language and culture</td>
<td>V, VIII</td>
</tr>
</tbody>
</table>

An overview of themes and projects is given below.
Overview of themes and projects

**Theme 1**
Boundaries of language variation
theme leader
Sjef Barbiers

**Projects**
Maps and Grammar (II)
Sjef Barbiers
The Language Portal (VII)
Ton van der Wouden, Hans Broekhuis
Invariants in Language Change (IX)
Gertjan Postma

**Theme 2**
Relation between dynamics and diversity in language
theme leader
Frans Hinskens

**Projects**
Life cycle of liquids (III)
Marc van Oostendorp

**Theme 3**
Relation between social, geographic and diachronic diversity in culture
theme leader
Herman Roodenburg

**Projects**
The ‘Dutchness’ of the Dutch. Constructions of the ‘national’ in a neo-liberal world (VII)
Herman Roodenburg
<table>
<thead>
<tr>
<th>Theme 4</th>
<th>Theme 5</th>
<th>Theme 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Projects</strong></td>
<td><strong>Projects</strong></td>
<td><strong>Projects</strong></td>
</tr>
<tr>
<td>Oral Transmission (Ia) Louis Grijp</td>
<td>Music cognition, language, and core knowledge systems (IVa) Louis Grijp</td>
<td>Nederlab (V) Nicoline van der Sijs</td>
</tr>
<tr>
<td>Tunes &amp; Tales (Ib) Louis Grijp Theo Meder</td>
<td>Language and the core knowledge system of number (IVb) Sjef Barbiers</td>
<td>The construction of local identities through cultural practices (VIII) Leonie Cornips</td>
</tr>
<tr>
<td>Fact (Ic) Theo Meder</td>
<td>Poetry, rhythm, and metre (IVc) Marc van Oostendorp</td>
<td></td>
</tr>
<tr>
<td>Cogitch (Id) Louis Grijp</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Knowledge and cognition: language and culture**

**External aspects of language and culture**

**Boundaries of cultural diversity**

**theme leader**

Marc van Oostendorp

Irene Stengs

Louis Grijp
4

4.1 Organization
4.2 Acquisition policy
4.3 Aims and strategy
4.4 Library
4.5 Paper collections
4.6 Audio collections
4.7 Digital collections
The Meertens Institute provides scientific insights into the nature and organization of Dutch society, focusing on everyday behaviour in language and culture and approaching this task from a comparative point of view. These insights are often founded on extensive documentation to some extent available in digital format. The Meertens Institute started collecting data in the early 1930’s, so documents have been collected over the past eighty years. Collecting relevant data is still an important part of the Meertens Institute’s research projects, although the way in which data are harvested has changed substantially: we have gone from paper questionnaires to a digital panel, from field work by bike to analyzing the web and from Meertens informants to crowd-sourcing. Documents are placed in several types of collections. Among them we find a large number of paper data collections, digital databases, audio collections and specialized book collections in the library. These collections and their position in the near future will be the topic of this chapter. Some of the matters discussed here will have a direct bearing on issues raised in the next chapter on technology.

Organization

As of 2013, the organizational structure of the Institute will contain four divisions: Research, Collections, Technology and General Affairs (see chapter 1). The Collections division includes the library, and other research collections (including audio digitalization) and focuses on digital developments (digital access to books and journals, digital (audio)collections, digital archiving). The Collections division will be part of General Affairs until January 2013, but in the new organizational structure it will have an independent position. The head of Collections is also the head of General Affairs. Given the relevance of digital collections and the importance of optimal accessibility to the collections with state-of-the-art digital services, it stands to reason that there should be close cooperation between Collections and Technology.
We expect that the clustering of KNAW-Humanities institutes will have an effect on the positioning of the Collections division. In the ongoing discussion in the Academy an explicit goal is to arrive at an integrated collections policy of the institutes that will be part of the KNAW Humanities Centre. This will also involve a new governance structure of the Collections divisions involved and a new location for the physical collections of the various institutes (see ‘Implementatie Contourennota 2012-2016’).

### Acquisition policy

| relevance to research purposes | The collections acquisition policy is closely linked to the research agenda of the Meertens Institute. The material on language variation and everyday culture to be collected by the Institute is selected on the basis of its relevance to research purposes, its quality, its relationship to existing collections, and its (digital) state. This policy implies that we generally do not strive for completeness in the fields of Dutch language variation and everyday culture. We aim instead at a high-quality digital set of collections for research purposes, as much as is legally possible available in open access (for the library collection, the open-access criterion is not a leading principle). For the coming period we intend to remain active in developing and acquiring collections directly related to our research topics. |
| related collecting institutions | The Institute has excellent connections with related collecting institutions in The Netherlands and Flanders. Dutch and Flemish Universities, professional organizations such as the Dutch Centre for Intangible Heritage VIE, and museums such as the Netherlands Open Air Museum quite often offer collections of papers, photographs, recordings and/or books to the Meertens Institute. Individual collectors and researchers consider it an honour when the Meertens Institute is prepared to accept their collections or archive as part of the Meertens collection. |
| national heritage | In particular domains we intend to strengthen our position as an institution taking responsibility for parts of the Dutch national heritage. Over the last few years our collections on folktales and folksongs have increased considerably, mostly through gifts from collectors or institutions not having the means and the knowledge to take care of their collections. The fact that our collections in the domain of oral culture (Liederenbank, Verhalenbank) constitute an important aspect of recent cultural heritage projects (e.g. the NWO CATCH projects Witchcraft, Fact and Cogitch) testifies to the opposite character of this policy. |
Similar domains in which research and the preservation of cultural heritage go hand in hand are, among others, the collections on pilgrimages, state inventories and the audio collections of dialects dating from the middle of the last century, and more recent dialect databases such as GTR and SAND. The huge collection of Meertens questionnaires (1931-2012) may also be considered to be part of the national cultural heritage. Acquisition in these cases is not based on research considerations only, but is also determined by the wish to keep these collections relevant, up-to-date and as complete as possible.

Aims and strategy

Given the relevance of data collections to our research projects, the Institute’s researchers invest considerable time in creating new or innovating existing collections, in close cooperation with members of the groups for collections and technology. A second indication of the importance of our data collections is their popularity for various types of extra-mural users. In order to accommodate them, a great deal of energy is invested in making the collections digitally available in open access with the help of a user-friendly interface on the Institute’s website. The website is consulted quite frequently by researchers and professionals. The popularity of the site (2010: 1,971,213 hits [unique IP-addresses]) shows that such investments also pay off for the public at large.

Continuing our policy of the last few years, we aim at a further digitalization of our collections, at the development of tools making the collections available for research purposes, at enriching the collections with standardized metadata, and at guaranteeing the sustainability of digital data through curation and archiving and through interoperability with relevant collections outside the institute, both nationally and internationally. One of the problems we have to face here is finding the financial means needed for digitalizing important collections. To achieve these goals, we have been – and will be – active in acquiring external funding. Given our large number of externally funded projects in this area, we have already been proved to be quite successful.

For the coming years we will set our goals even higher: we aim to become one of the leading institutions in The Netherlands in the field of digital humanities, in particular in our fields of inquiry. Our focus in this domain is the development of a stable data infrastructure with innovative tools for integrating large datasets in humanities research. Moreover, we intend to form new international alliances in order to extend our international network of humanities
centres with digital research collections. We have already managed to achieve this in the context of CLARIN (the Institute is a prospective CLARIN-NL-centre) and the European dialect group Edisyn (www.dialectsyntax.org).

**paper collections**

As for the non-digitalized collections of the Institute, we are aiming at a more professional location and an explicit collection policy with respect to issues such as conservation, maintenance, storage, deaccessioning and availability. The discussion on the clustering of KNAW-humanities institutes in an Amsterdam KNAW Humanities Centre will lead to the formulation of a more explicit collection policy in the light of possibilities that may arise as the consequence of this clustering. It may also give rise to alliances with collections of other KNAW institutes and with the library of the University of Amsterdam.

**Library**

**representative survey**

In today’s situation, the library is primarily a paper library and its collection is determined by professional considerations. The library aims at offering a representative and as complete as possible a survey of Dutch ethnology and variation linguistics and is unique in this respect in The Netherlands. There is no other institution that offers such a comprehensive survey. The importance of the library for research in the Institute is quite considerable.

**online and open access**

Given developments in the digital field, it will be necessary to reconsider our policy with respect to the library. Large amounts of literature are or will become digitally available, increasingly in open access as well. The DBNL (Digital Library of Dutch Literature) has been active in this field for years. Extensive projects such as Google Books and Libratory, in which the Royal Library and university libraries are also involved, take care of the digital availability of older books, magazines and journals especially. Magazine publishers are increasingly turning to offering magazines online, mostly for a fee.

**e-books and e-journals**

These developments have consequences for the library of the Meertens Institute in several ways. Firstly, acquisition policy needs to be reconsidered. It will be focused on buying e-books and licences to e-journals, rather than obtaining material in print. Magazines and books determining the identity of research at the Meertens Institute will continue to be bought in the future. Books and magazines that are available digitally need not be bought in print unless there are good reasons. In these cases the library might make do with offering a list of titles of relevant electronically available publications with links to these publi-
cations. In addition, the library needs to ensure good access to the publications (among other things through licences). Should any doubt about this arise, potentially important texts can be included in the repository of the Institute. As from 2011 the Academy provides access to content from Elsevier Publishers, Springer and Wiley and an agreement has been concluded with the Royal Library for, among other things, JSTOR, LexisNexis and the project MUSE. Via intranet the library offers a survey (through licences) of the available relevant publications with links to the (full text) publications.

Secondly, it is desirable in this context that specific parts of the library be included in existing digitalization programmes. For example, the unique collection of dialect grammars would have to be considered for inclusion. To achieve this, the Institute would have to be in touch with these programmes or would have to generate means itself to carry out this task.

Thirdly, the library will have to take charge of the increasing number of digital articles and data to be stored in the repositories of the Institute (see chapter 5.5). It stands to reason that the library should mediate in including publications and databases in these repositories and that it should host a good search engine for searching the Institute’s repositories.

Fourthly, the library will become more efficient once the cataloguing system is connected to the Union Catalogue: it will then be possible to extract title descriptions instead of creating them. The unique library collection will become more visible and be consulted more frequently after it has been hooked up to Worldcat. Realising these connections will have the highest priority in the coming period. In line with the substantial changes in the media landscape, the Institute’s library will then change from a collection of books and journals to a gateway to information.

**Paper collections**

In the coming years we will keep investing in maintaining our paper data collection in an optimal condition for research purposes; we consider digitalization the best option for conservation, especially for vulnerable research collections of which the paper quality is deteriorating rather quickly, such as the collection Boekenoogen (oral culture). This type of large paper collections is rather expensive to digitize but the Institute is continuously looking out for funds to safeguard our collections and to make them available in open access on the internet.
One way of keeping existing paper collections relevant and finding the financial means to digitalize them is to make them part of ongoing or future research. A good example is the questionnaires collection. The Meertens Panel (www.meertens.knaw.nl/meertenspanel) is a source of new data by virtue of internet questionnaires. By continuing collecting data via questionnaires, the numerous paper questionnaires that have been collected as of 1931 retain their relevance, providing researchers with a diachronic window of 80 years. A crowd-sourcing project has recently started with the aim to make all paper questionnaires digitally available.

Several of the paper collections important for research are not yet or insufficiently available in the desired digital format. A large part of this material cannot automatically be transferred into a digital format by means of OCR (Optical Character Recognition). This holds, for example, for hand-written answers to questionnaires or for the collection of field names, which consists of hand-drawn maps with the field names put in by hand. For research that investigates older texts – such as diachronic language research or research on song material from an older stage – an adaptation by hand of the digitalized material will also be necessary to make them suitable for research. There are indications that these expensive operations will be made easier in the future by technological developments like the improvement of character recognition, automatic recognition of handwriting, standardization of spelling and conversion of speech to text. In the coming period tools like these will not yet be so perfect that they can be used on a large scale in the basic process of digitalizing research data. This implies that for the time being considerable funds will have to be invested in making serviceable research data available digitally. Priority will be given to the Institute’s own collections that link up with current research or that offer concrete possibilities for future research.

The audio studio of the Meertens Institute offers optimal facilities for digitalizing audio material from various sources. Several of the audio collections belong to the national cultural heritage and play a role in research programmes and in valorization projects. If intellectual property rights (IPR) allow it, we will try to make the digitalized material available via the Institute’s website in open access. This holds in particular for those collections that are used in ongoing research projects. 85% of the audio material that is present at the institute has already been digitalized in AIF or BWIF-format, 3.3 TB in total. The remaining 15%
consists of older material that has not yet been digitalized due to a low priority – relatively high quality of the analogue tapes and/or relatively low relevance for research purposes – or new material collected recently. Most of our audio material is now digitalized on CDs, which are kept – together with the original tapes – in a climate test chamber in the archive. We are also looking for ways of putting the audio material on the web, similar to for instance the Soundbites collection, the dialect material collected in the 1960s by Jo Daan and her colleagues (www.meertens.knaw.nl/soundbites).

Our target is to digitalize most of the non-digital collections in 2013-2018. In addition to the required metadata description and the addition of persistent identifiers, it will be necessary to provide automated support for parts of the work processes in archiving and publication. We expect that long-term persistency procedures and automated support can be handled by collaborating with the Max Planck Institute and DANS in The Language Archive initiative.

**Digital collections**

Digital research data are primarily made available through research-related projects. It is important that the data on which research is based are electronically available and online checkable. Grant-giving organizations often demand lasting open access to data. Open access to all kinds of research data enables us to develop new research plans and projects in which researchers and technology developers cooperate on methodological and theoretical innovation. Although the Institute maintains a sizeable repository of autonomous documentation, i.e. documentation no longer used in ongoing research projects, all efforts are directed to supporting current research projects, either contributing directly or by making research data available for relevant research in the near future.

Generally, all digital research data are made publicly available to both the research community and the general public. The Meertens Institute will deliver all metadata documents in standardized format and make them available for harvesting in accordance with CLARIN infrastructure guidelines. The metadata documents and digital objects are currently stored on local servers but the Meertens Institute will collaborate closely with The Language Archive (TLA) and KNAW to deliver long-term persistency guarantees for its data and metadata.
persistent identifiers

Where appropriate, persistent identifiers are assigned to all resources in line with current thinking on digital archiving guidelines. For persistent identifier management SARA’s Handle Service is used, which is part of the EPIC consortium. It is the intention to re-use, wherever possible, existing systems and technology. TLA already houses a large digital archive and related technology that will be rebuilt to match CLARIN infrastructure requirements. The Meertens Institute will ensure that relevant procedures fit the Meertens collections and procedures, thus creating a joint archiving solution.

resources

Metadata linking to digital research data is by way of standard harvesting protocols (OAI-PMH). Research data are made available using standard web server technology. The data may be protected where appropriate. In addition to opening up archived resources via the internet, the Institute also maintains a large number of ‘operational’ resources. These services are often made available through domain-specific web applications or websites, depending on the research project’s demands. The list of publicly available ‘operational’ resources can be found at http://www.meertens.knaw.nl/cms/nl/databanken. For researchers and the general public the Meertens website acts as a central portal through which these data can be assessed. Available metadata of archived resources are centrally collected and a user environment providing search and browse options is currently under development.

international infrastructure

The Institute is active in making the data collections available to different groups of users, of which the research community is the most important one. We are involved in the development of the technology required for research purposes in national and international projects and networks, such as CATCH and CLARIN (see chapter 5). In our view, one of the most important technological tasks concerns the development of an optimal international infrastructure for digital collections. In the coming period the Institute aspires to becoming one of the leading institutions in this venture.
5

5.1 The relationship with research
5.2 Documentation
5.3 Projects
5.4 Communication
5.5 Archiving and open access
5.6 Technical equipment: system management and helpdesk
5.7 Clustering of academy institutes in the humanities
5.8 Agenda for technological innovation for the period 2013-2018
The last few decades have witnessed a radical change in society as a result of the technical revolution in digital and computational resources. Twenty-five years ago e-mail was in its infancy and the world-wide-web had yet to be invented. Nowadays life without e-mail and the internet is virtually unthinkable. These changes have also had an impact on the Meertens Institute with respect to the organization of research, documentation and archiving, as well as internal and external communication. To give an example: an Institute website has been set up, which went online in 2000, and has now developed into a much-frequented and popular website, with over a million annual visitors. In 2009 1,494,959 people (unique IP-addresses) accessed the site and looked at more than 45 million web pages; in 2010 the totals were: 1,971,213 visitors and more than 54 million web pages.

It would seem that these developments do not mark the end of a technological era, but rather the beginning of a new period in which the digital humanities (e-humanities) will develop much further with the help of new digital technology. These developments open up new prospects for research and documentation at the Meertens Institute with its large number of databases and collections in the field of Dutch language and culture. Naturally, the Institute cannot afford to lag behind, or sit on the fence, as developments in research will avail themselves of new technological possibilities in ever-increasing measure. Indeed, it is the Institute’s ambition to play a prominent role in these developments. Inevitably, this will have consequences for the organization of research and documentation, and for the composition of the Institute’s personnel. A prominent place in this research plan is reserved for this aspect.

We will distinguish between different types of technological innovation in this chapter, specifically with regard to the different ways in which new technology is developed and employed in research, documentation, archiving and internal and external communication. Finally, a separate section is devoted to the technical equipment and technical support necessary to achieve these goals. These topics return in the individual sections of this chapter.

The relationship with research

The nature of research in the humanities will undoubtedly change considerably as a result of the possibilities offered by new technology. A digital revolution in the humanities is likely to take place as ever-increasing quantities of data...
will become available to the researcher by means of improved technological resources. Especially in those subjects in which empirical processes play a central role, technological innovation will change the character of the field of study. Given the availability of much larger data sets and standardized digital research methods (see section 5.2), existing hypotheses can be tested quantitatively, and research questions can be asked which are now outside the reach of the researcher. It will also become possible to subject hypotheses to computer-modelling.

**melody recognition**

To give an example: digital research on variation space in melody based on large data collections may lead to research on the cognitive representation of music, given that music-theoretical models of variation only partly represent melody recognition by humans. Adding insights from cognitive sciences may lead to a deeper, more empirically reliable, insight into our human ability to recognize melody, and thus to a deeper insight into our human capability of comparing melodies. Research within the framework of Witchcraft (NWO-CATCH), now completed, and its successor's COGITCH (NWO-CATCH) and Tunes & Tales (KNAW-Computational Humanities), which addresses melody recognition computationally, provides grounds for the assumption that such an approach to melody recognition can yield new insights into the way in which music is stored and processed in the brain. One of the projects of Knowledge & Culture (NWO-Horizon) takes this issue as its research topic. A European visual database of iconic gestures, recently started at the Meertens Institute, opens up similar possibilities concerning recognition of gestures.

**language variation**

Another example relates to the fact that we would like to be able to study the phenomenon of variation in the position of the finite verb in the Germanic and Romance languages in detail (e.g. down to the level of the dialects) on the basis of an unprecedentedly rich database composed of data collections from the whole of Europe. Current technology makes it possible to put together a relevant database to study structural language variation in detail. The empirical generalizations that might follow from the availability of such an overarching database may lead to drawing up and testing new hypotheses which are consonant with the digitally compiled corpus. This might eventually lead to drawing up a new set of testable hypotheses on the relationship between language variation (synchronic and geographical) and language change (diachronic). Such an enterprise might then lead to new hypotheses regarding the nature of linguistic structure, based on a deep insight into the position of the verb in Indo-European languages.

**Computational Humanities**

The Royal Netherlands Academy of Sciences (KNAW) has recently paid a great deal of attention to developing a programme specifically directed at
innovation in research with the help of innovative technology. In 2011 the special programme Computational Humanities has embarked on a number of projects funded by KNAW-central. The Meertens Institute participates in this programme, which is directed at “discovering high-level concepts, patterns and motifs across media” with its project Tunes & Tales.

In the coming period the majority of the Institute’s researchers will spend part of their research time on developing and implementing plans in the area of the technological innovation of research in the humanities. This will be done within the framework of the research projects they are working on: firstly, to ensure that the data on which their research is based are electronically available and checkable online, and secondly, since this investment will lead to new plans and projects in which researchers and technology developers cooperate on methodological and theoretical innovation of research. Additionally, it should lead to a greater participation in the discussion concerning digital technology in the humanities and to leading publications. It is hoped and expected that the Meertens Institute will play an active role in this area in the (inter)national field.

Documentation

Documentation at the Meertens Institute in the area of everyday culture and language varieties has been a central part of the Meertens Institute’s mission. The primary goal of documentation has been to provide the Institute’s scientific research with an empirical basis. The Institute has a large collection of databases at its disposal, which are only to a limited extent digitally available. The size of these research databases increases all the time as new research regularly yields new databases and as the Institute regularly acquires new collections (see chapter 4).

This indicates the line of thinking about collections and databases at this moment: the collected data are available for research; these data are collected on the basis of research questions. Although the Institute offers a substantial amount of autonomous documentation – i.e., documentation that is not or no longer used for current research – the efforts in the area of documentation are in virtually all cases directed at scientific research, either by contributing to current research or by making relevant research possible in the near future.

In the coming period new research data will often be collected by making use of new techniques in which informants will be asked online to fill in question-
A problem that will require a great deal of attention in the coming period is that much of the existing documentation of the Institute is not yet digitally available. And once the material has been digitalized it is often insufficiently enriched by metadata and tags to make it serviceable for research purposes. Below we describe the different phases that databases have to go through before they are optimally ready for research purposes.

In the first instance, we are dealing with the time-consuming and therefore expensive operation of digitalizing large amounts of material: scanning of paper information carriers / audio digitilization of older sound tracks, and making this information available in a searchable text format. The Institute owns a large number of important collections for research that are not yet or insufficiently available in the desired digital format. A large part of this material cannot automatically be transferred into a digital format by means of OCR (Optical Character Recognition). This holds, for example, for hand-written answers to questionnaires (see chapter 4.6 for a new version of the questionnaire) or for the collection of field names, which consists of hand-drawn maps with the field names put in by hand. Automatic speech to text conversion techniques are not yet available for our type of audio data. As a tekst-based format is imperative for research purposes it will be necessary to provide the audio material with a transcription. For research that investigates older texts – such as diachronic language research or research on song material from an older stage – an adaptation by hand of the digitalized material will be necessary to make it suitable for research. There are indications that these expensive operations will be made easier in the future by technological developments such as the improvement of character recognition (OCR), automatic recognition of handwriting, automatic standardization of spelling or automatic conversion of speech to text. In the coming period tools like these will not yet be so perfect that they can be deployed on a large scale in the basic process of digitalizing research data, which implies that for the time being considerable funds will have to be invested in making research data digitally available. In making funds available in this area, priority will be given to the Institute’s own collections that link up with current research or offer concrete possibilities for future research. It is worth mentioning in this respect that the Institute is also investi-
gating methods of applying crowd-sourcing techniques to assist in the labour-intensive tasks of digitizing research data and providing metadata descriptions. Examples of this are the ongoing ‘Gekaapte Brieven’ project and a crowd-sourcing project involving the transcription of the handwritten answers to the questionnaires sent out from the 1930’s onwards.

Digitalized materials must be provided with metadata – information on properties like the place where and the time when the data was recorded, information on the author, speaker, informant and the person recording the data and on the contents of what has been recorded. Metadata of this kind make it possible to search for the desired information in large quantities of digital data. This information must be added to the data in such a way that the metadata are searchable by search engines. A specific problem here is the fact that the metadata need to be standardized in such a way that a researcher can find various relevant data corpora with the help of the metadata. Work will be necessary in the coming period on assigning metadata to research data insofar as this has not yet taken place – e.g. the collection of more than a thousand hours of dialect material from the middle of the 20th century, the Soundbites, has only received a limited amount of digital metadata (most of the material has not even been transcribed). The Institute will also focus its efforts on a good metadata system in close cooperation with national and international partners (in 2010 the Meertens Institute, together with the Max Planck Institute in Nijmegen, worked on a metadata standard (CMDI) in a project within the framework of CLARIN-NL), allowing researchers to distill the information relevant to their research from the enormous mass of digitally available data. A prototype of this search engine based on the CMDI framework has been made available and will continue to be developed within the CLARIN-NL research programme.

The relevance of digital material for scientific research increases proportionally with greater enrichment. For the collection of song material – the Song Database – it will be necessary to link the song texts in the corpus to a form of musical notation. For the collection of folk tales (the Folktale Database) it is necessary to link individual stories to thematic information via the so-called Aarne-Thompson index. For research into the structure of language it will be necessary for grammatical information to be added to the text. Here, the main enrichment of the material will be information on word classes. This process can be simplified considerably by employing Parts-of-Speech tagging, but even here the process cannot run fully automatically, especially in those cases where the language material deviates from the present standard language, as is the case in dialects or historical stages of Dutch. Greater enrichment of language material requires considerable effort: we are dealing here with, among other
things, morphological and syntactic annotation, phonetic transcription, recognition of phrases, syntactic parsing via a so-called tree bank, etc. Within the projects running at the present moment, the goal is to enrich the material as much as is possible in accordance with internationally adopted guidelines. Enriching the material collected in the framework of the project Syntactic Atlas of Dutch Dialects (SAND; 2000-2009) has borne out that this is a supremely labour-intensive operation.

**development of tools**

Furthermore, it will be necessary to open up the data through ‘tools’, i.e. applications making it possible to represent, order or edit the body of research data in various ways. A well-known application of this kind concerns the use of existing statistical applications in ordering research results and representing them in graphs and diagrams. Besides these standard applications, there are several applications to the development of which the Meertens Institute contributes. To illustrate: a geographical application has been developed in the framework of the linguistic atlases which enables language phenomena to be shown on a map of the Dutch language area. At this moment work is being carried out on another application that makes use of GIS (Geographical Information Systems) in order to project field names onto a Google-type map. One more application that is being worked on is the development of a melody-recognition mechanism that would enable the researcher to automatically recognize similar or comparable melodies, thus allowing the ordering of data in the Song Bank. It will turn out to be necessary to keep developing applications in tandem with ongoing research in order to open up new dimensions in computational research in the humanities.

**digital infrastructure**

Finally, the enriched data and their applications must be embedded in a digital infrastructure. The goal is that the available databases can not only be accessed individually but also that the collection of (inter)national databases on a specific research area can be interlinked and investigated together. To achieve this, the metadata of the digitalized research data are made available for central harvesting and the internal search mechanisms are adapted so that the research data can also be considered for their content.

**workflow**

Moreover, the idea is to generalize existing applications, which as a rule have been developed for a single database, in such a manner that they may have a much broader diffusion. In this regard not only the distribution of the technology developed but also the harmonization of the methodology developed plays a role. Furthermore, it will be necessary for applications to be connected with each other in a so-called workflow. As an example we note that with linguistic enrichment, the output of transcription tools must serve as the input for a PoS-tagger, and its output must in turn serve as input for a syntactic parser.
Given that linguistic and ethnological research is international by nature, it will be necessary to work at an international level on a reliable and sustainable infrastructure. The Meertens Institute intends to develop into a national centre for the infrastructure in the humanities during the coming period, focusing on data and applications of a linguistic and cultural nature. In 2010 the Meertens Institute had already presented itself as an (inter)national centre for an European infrastructure within the framework of the European infrastructure project (ESFRI) CLARIN – Common Language Resources and Technology Infrastructure – , which addresses itself mainly to language sources. In the coming period the Institute will continue to develop towards becoming a centre for digital infrastructure for the humanities, among others in the context of the new infrastructure initiative CLARIAH.

The Meertens Institute has a reputation to uphold in the area of unique collections and data in the field of language and culture. As appears from the enormous public interest in the various websites, these data are a valuable source for research in and outside the Institute and form the basis for an (inter)national knowledge centre in the area of language variation and everyday culture. As the Strategic Agenda 2010-2015 of the Royal Netherlands Academy of Sciences puts it: “The Academy Institutes function as national knowledge centres and/or are in charge of national facilities which are internationally accessible as well. These facilities cover the management of and access to unique infrastructural provisions and world-famous collections”. It is exactly this objective that the Meertens Institute considers important for the coming period. The expectation is that the road already taken in the direction of an internationally operating institute in the area of digital humanities will take us to the point where the Meertens Institute, possibly together in a joint venture with other KNAW humanities institutes such as the Huygens Institute, NIOD, KITLV and DANS, becomes a national agenda-setting institute and an international point of contact for parts of the e-humanities. In so doing, the Institute will be able to extend its prominent position from a collection-housing institute to a position where it becomes an innovative knowledge centre in the area of language and culture.

Projects

In the past and coming periods the Meertens Institute has been involved or will be involved in a fair number of extensive national and international projects in the field of technical development. These projects together provide the
substance of the developments mentioned in the previous section. To illustrate, the infrastructure to be developed in CLARIN (project A) will yield a number of technical specifications (among other things with regard to persistent identifiers and metadata) that resources must comply with. CLARIN is working on this in cooperation with, among other things, CATCH+ (project B) and Alphalab (project C). Below, the most important projects will be mentioned.

A) CLARIN-NL (ESFRI/NWO; 2009-2015)

The object of CLARIN is to develop an international infrastructure for the humanities (and the social sciences) in the area of language, speech and text, in which the databases and applications of the organizations involved are made accessible and available for research in a more or less uniform manner.

The European programme CLARIN-EU is set up in such a way that the moneys for developing such an infrastructure are awarded at the national level. Within this framework, The Netherlands, the principal mover of this programme, has made available an amount of upwards of 9 million euro to CLARIN-NL. A large number of projects are funded with this money, the selection having been achieved by way of calls for proposals. From the very beginning the Meertens Institute has taken an active position within CLARIN-NL.

At the management level the Institute is represented in the Executive Board (Bennis is chairman of the Executive Board and contracted by CLARIN-NL for one day per week) and in the Advisory Board of CLARIN-NL (Barbiers, Brugman, and Kemps Snijders) and on an international level in the Scientific Board of CLARIN-EU (Bennis, nominated by NWO), and in the DCR-Board in the area of ISOCAT (Kemps Snijders). Meanwhile, the Institute has been accepted by the Board of CLARIN as a foundational CLARIN centre, together with the Max Planck Institute in Nijmegen (MPI), the INL in Leiden and DANS in the Hague. The Meertens Institute is also represented in the application CLARIAH (September 2011), Common Lab Research Infrastructure for the Arts and Humanities – to some extent a continuation of CLARIN-NL – as a co-applicant (Bennis, Kemps Snijders) and within this programme as coordinator of the programme ‘Visualisation & Presentation of Information Content’.

The Institute is involved in a number of CLARIN projects:

larger CLARIN projects

Larger projects, financed in CLARIN-NL:

- IPP (Infrastructure Implementation Plan; coordinator MPI; Kemps Snijders; 2010-2013);
• Search & Develop (coordinator Meertens Institute; Kemps Snijders; 2010-2013);
• The Dutch-Flemish project TTNWW (coordinator Meertens Institute; Kemps Snijders; 2010-2013);

**Smaller projects, financed in CLARIN-NL:**

• The Meertens Institute was involved in the project Metadata in 2009-2010 (coordinator MPI; De Vriend);
• The project MIMORE (Microcomparative Morphosyntax Research Tool; coordinator Meertens Institute (Barbiers, Kunst), in collaboration with the University of Utrecht) was finished in 2011 and the envisaged tool was delivered;
• The project C-DSD (De Bruin) is a project that aims at curating the Dutch Song Database (coordinator University of Utrecht); the project was finished in 2012;
• The project COAVA (Cornips) is a project in which a tool will be developed for easy exploration of the linguistic characterizations of objects from two databases of language variation (historical dialectology) and three databases of language acquisition (coordinator Meertens Institute); the project started in 2011;
• The project INPOLDER (Postma) is a project that aims at providing syntactically analysed corpus material for historical Dutch (coordinator University of Nijmegen); the project started in 2011;
• The project CLAVAS (Brugman) aims at developing a vocabulary tool usable by CLARIN to increase the quality of metadata (coordinator Meertens Institute); The project started in 2011;
• The project PILNAR (Kemps Snijders) intends to establish a core corpus of modern pilgrimage narratives from various sources (Coordinator: University of Tilburg); the project started in 2012;
• The project MIGMAP (Kunst) will develop a web application where the migration patterns between municipalities of the Dutch population in the 20th century are presented over three generations (Coordinator: University of Utrecht); the project started in 2012;
• The FESLI project (Snijders) aims to deliver qualitative and quantitative comparison methods for the study of Specific Language Impairment (SLI) (Coordinator: University of Amsterdam); the project started in 2012.

B) **CATCH and CATCH+ (NWO et al.; 2005 - …)**

The CATCH (Continuous Access To Cultural Heritage) programme initiated by NWO develops generic methods and techniques at the intersection of humani-
ties, information science and cultural heritage, for the use of and in interaction with collection-managing institutions. Technical innovation, cooperation and transferability are central parts of the programme. CATCH+ continues to build on the result of CATCH-projects and intends to develop widely usable software as a result of research in completed CATCH-projects. At the management level the Meertens Institute is represented in CATCH (Supervisory Board: Bennis) and CATCH+ (Steering Committee Bennis; projectleader Van Alkhoven; technical leader Brugman) and has been involved in the CATCH-project Witchcraft (Melody Recognition; implemented at the Meertens Institute in cooperation of the University of Utrecht; 2006-2009; Grijp, Van Kranenburg) and is involved at this moment in the CATCH-project LINKS (personal entries; implemented at the iiSG; 2009-2013; Bennis, Bloothooft) and the CATCH+ project Witchcraft+ (Meertens Institute; 2009-2011; De Bruin). As of October 2010, the Institute has been administrator of CATCH+ and the central office of CATCH+ has been taken over from Sound & Vision and incorporated in the Meertens Institute.

In 2011 two new CATCH-projects have been accepted (see Chapter 3):

- **cogitch** (coordinator University of Utrecht; 2011-2015):
  The Meertens Institute and the Netherlands Institute for Vision & Sound manage large collections of folksongs and popsongs. The researchers aim at developing a generic tool to open up the two collections (De Bruin, Grijp).

- **fact** Folktales as Classifiable Texts (coordinator Meertens Institute, Meder: 2011-2015):
  A collection of Dutch folktales will be provided with metadata with the help of software still to be developed. Moreover, for the first time automatic document-clustering techniques will be applied within the framework of fact to a very large and varied collection of folktales. New relations between folktales might be discovered with this tool.

**C) European Dialect Syntax (ESF; EURYI grant; 2005-2012)**

One of the goals of the European Dialect Syntax Project (Edisyn) is to bring dialect syntax databases together online to allow the researcher to search these databases with a common search engine, using tags and English glosses, to visualize these data on Google maps and to export them to statistical tools. The dialect syntax databases include both systematically collected data and corpora of transcribed speech. Currently, databases from Scandinavian, Dutch, English, Italian, Portuguese, Estonian and Slovenian dialects are searchable. In the future, databases from currently running projects on Hessian, Alemannic, French and Spanish Basque, Breton, Spanish, American English and Malagasy
dialects will be added. The Meertens Institute (Barbiers, Kunst) was the coordinator of the project and responsible for software development, technological advice and maintenance of the infrastructure. Although the project ended in 2012, the databases and the international network will be continued under the supervision of the Meertens Institute.

D) Dutch Songs online (**NWO-middelgroot 2009-2013**)

In Dutch Songs online an extensive database containing 100,000 Dutch song texts until 1900 is being developed. Using the new search possibilities of the database, an interdisciplinary working group plans to investigate the role of songs in the construction and transformation of group identities in Dutch and Flemish society from the Middle Ages to 1900.

As a collaborator in this project, the Meertens Institute (De Bruin, Grijp) is responsible for the selection and description of the material and for the development and hosting of the research interface. Among the other partners are The University of Utrecht as administrator and the Digital Library of Dutch Literature (Leiden).

E) Language Portal (**NWO-groot; 2011-2016; see Chapter 3**)

The Language Portal is an extensive project (c. 3.5 million euro) which aims at bringing together existing information on the grammar of Dutch and Frisian (phonology, morphology and syntax) and making this electronically accessible and searchable for scientists (English version) and professionals (Dutch version).

The Meertens Institute is involved as administrator (Bennis, Van der Wouden) and as collaborator (Van Oostendorp, Broekhuis, Linke) in the development of this wide-ranging portal. Among the other partners are the Fryske Akademy, **INL** and several universities. Cooperation with Flanders is envisaged.

F) Tunes & Tales (**Computational Humanities; Royal Netherlands Academy of Sciences; 2011-2015; see Chapter 3**)

The Academy subsidizes the project Tunes & Tales: modeling Oral Transmission, of which the Meertens Institute is the chief applicant (Meder, Grijp). In this
project research will be carried out in the oral transmission of melodies and folktales. A basic assumption is that melodies and folktales are construed from regular building blocks or motives.

G) **CODA (Andrew W. Mellon Foundation (USA); 2011-2012)**

CODA

The CODA proposal (Brugman) centres around two main use cases concerning transcription of scanned handwritten historical documents. It builds on several software components implemented for CATCH and CATCH+: an annotation repository, document annotation tool, workspace services, entity detection software and image processing software for detecting bounding boxes for written lines.

H) **Hijacked Letters from the 17th and 18th centuries (Prins Bernhard Cultuurfonds; 2011-2012)**

Hijacked Letters

This project has produced a digital edition/transcription of letters that have already been scanned (Kew letters) from the collection Hijacked Letters, with the assistance of volunteers who will use an interactive website to be developed in this project (projectleader Van der Sijs). In the coming period new initiatives will continue the transcription of the remaining material not digitalized yet through crowd-sourcing.

I) **TINPOT (Taal, Identiteit, Netwerken en Produktgeruchten Op Twitter; Digital Humanities knaw; 2012-2013)**

TINPOT

The TINPOT (project leader Meder) project will develop software for analyzing the language used in tweets in order to automatically determine gender and age group of twitterers. Furthermore, software will be developed for determining the network of twitterers, their activity, and whether they are opinion leaders or followers. Finally, temporary rumours and hoaxes will be monitored from start to end. The software will analyze the characteristics of twitterers who tweet and retweet, will follow the paths the tweets take and monitor whether or not the tweets change during the process. The TINPOT project is funded by the Public-Private Center for Digital Humanities (KNAW), The software development is supported by the University of Twente. The tweets from Twitter are being supplied by Teezir BV in Utrecht.
J) Transcription of Soundbites (KNW Open-Access fund; 2012-2013)

The goal of this project (coordination Van Oostendorp, Zeldenrust) is to relate existing (paper) transcriptions of dialect recordings from the period 1955-1973 to digital soundfiles, which will be available on the website of the Institute (Soundbites). A second objective is adding metadata in CMDI-format and persistent identifiers in accordance with CLARIN-specifications.

K) Nederlab (NWO-groot, 2013-2018; see Chapter 3)

On the initiative of the Meertens Institute, a NWO-groot proposal has been handed in (coordinator Meertens Institute; Barbiers, Bennis, Van der Sijs) which aims at developing a platform for researchers in the humanities (among them historians, literary scholars and linguists) working on historical Dutch texts. Corpora and tools are brought together in this platform and a future user-friendly web interface will ensure that an important instrument is developed for research on historical texts. The project has been accepted (4 m€) and will start in January 2013.

Communication

Communication with researchers, professionals and interested people outside the Institute increasingly takes place through the Institute’s website, making databases and their technical applications available in this way. Recent instances of databases available through the web are, for instance, the collection of audio material containing dialect recordings from the fifties to the seventies of the previous century – Soundbites – and the revised databanks of saints, christian names and family names. The number of hits indicates that this is an extremely attractive way to make data available to the general public, but also to researchers in the area of language and culture in the Netherlands and abroad. The website also serves to inform visitors of ongoing research and activities, such as lectures, congresses and open days. News on press releases and special events can be found here as well. As of 2011, the Institute distributes an electronic newsletter informing those interested of the most recent developments. The institute is also active on Facebook and Twitter, with more followers every day.
In 2012/2013 the website will undergo revision, partly in order to better reflect the new organization structure. More information will become available in English and search options and navigation will be improved.

Part of the website is in interactive form ("web 2.0") where website visitors not only consume data but also react and contribute to research. In 2010 the famous questionnaire of the Meertens Institute began a new lease of life: the paper questionnaire used until 2009 was replaced by a digital inquiry form where informants answer questions at their computers. In February 2010 a questionnaire on language change in Dutch showed that this type of inquiry can make a successful contribution to the Institute’s data collection. In this way the questionnaire as a means of collecting data can be maintained and the extensive collection of lists from 1931 remains valuable, among other things as a means of gaining insight into a changing society.

The Institute intends to extend interactive internet communication in the coming period. By the middle of 2010 the web-based ‘Meldpunt Taal’ (Contact Centre Language) was founded on the basis of an idea of Ewoud Sanders in the newspaper NRC/Handelsblad and on the initiative of the Meertens Institute. The Institute for Dutch Lexicology, the periodical Onze Taal (Our Language), Van Dalen Lexicography and the Meertens Institute cooperate in the Contact Centre. The goal is to develop and maintain a website where visitors can ask questions about language, and exchange and share their language findings (new words, odd constructions, funny expressions, etc.) with experts.

Another example of reaching a wider audience is via the website ‘The Language Detector’ (2012), which acquaints the visitor with the rich linguistic variation in The Netherlands. On the basis of a number of multiple-choice questions on word formation and word order, the detector determines the speaker’s regional background. The information is based on data from the Syntactic Atlas of the Dutch Dialects (SAND) and Morphological Atlas of the Dutch Dialects (MAND).

Not only communication with outsiders but also internal communication is extremely important for the running of the Institute. As of 2009, the Meertens Intranet has been in existence, where information on activities, meeting reports, regulations, introduction of new members, etc. can be found. Successfully running an organization also presupposes a good exchange of information. In the coming period the Intranet will turn out to be an important means of exchanging news on formal and informal matters.
Archiving and open access

For digitally archiving publications and databases of the Meertens Institute a repository should be used in which publications and databases are stored in a secure and reliable manner.

As for a repository for publications, the Institute finds itself in a situation in which researchers enter their publications in the research information system METIS used by universities and Academy institutes. They enter a description of their publications detailing author, title, year, publisher, etc., while a link should make it possible to request the full text from NARCIS – National Academic Research and Collaborations Information System (to the extent that copyrights allow this and that NARCIS provides the relevant service optimally). In September 2012 almost all the articles of researchers as of 2000 have been scanned and introduced into this repository. The coming period will be devoted to introducing book publications and to retrodigitalizing.

As far as storage of data is concerned, the Institute will introduce the research databases present in a data repository in 2011/2012. The storage facility will have to satisfy the criteria developed by DANS, a so-called Trusted Digital Repository having a Data Seal of Approval. This implies that the data must be visible on the internet and be accessible in accordance with current legislation, be available for a long time in a workable data format and be reliable and checkable, among other things by the assignment of persistent identifiers (PIDs). The data will be recorded under our own control. At this moment discussions are going on with the Academy and the Max Planck Institute Nijmegen (The Language Archive) about a long-term guarantee for language data and, in the same framework, possibilities are being investigated for a similar guarantee for data on culture.

The Institute has the policy that publications and data produced by the Institute and deposited in a repository should be, to the extent that this is possible, openly accessible to others. This open-access policy has a number of, above all legal, restrictions, for example with respect to copyright and privacy laws. In this respect we stick to the ruling preferred by the Academy (Code of conduct for the use of personal data in scientific research, KNAW 2003 and the Strategic Agenda KNAW 2010-2015, p.33, memo Open Access KNAW 2010). Existing rules for free access to publications and research data are interpreted in such a way that publications and data are accessible freely unless there are solid legal or scientific arguments not to do this (“yes, unless”). Besides evident
exceptions such as publications in journals where the publisher prevents open access on the grounds of copyright, there may also be cases in which pro tem research data are not accessible. Such a situation may occur when a group of researchers or an individual researcher has brought data together on which they are writing a publication. As soon as the publication based on these data is complete, the collection of research data financed by public means is made available to others in open access through the website, i.e. without previous permission or approval and without charging costs. Even if research data can only be made public with the greatest difficulty on account of privacy considerations, it will be attempted to make the data on which the research is based openly available by anonymization or otherwise, so that the research concerned is checkable with respect to the underlying data (among other things by way of an access management system).

open-access journals

In the coming years the Institute will contribute to the discussion on open access by insisting on maximally open access in a teamwork context and by taking part in initiatives promoting open access of publications. Initiated by staff members of the Institute and financially supported, the electronic ethnological journal Quotidian began its life in 2010 (Amsterdam University Press) and the journal Taal en Tongval has been converted into an online journal with a grant from NWO.

Technical equipment: system management and helpdesk

I&A

It stands to reason that a basic condition for technological development is the possibility to have modern, state-of-the-art, equipment. This equipment needs to meet the needs of all technical requirements allowing optimal research in the humanities. The main issue here is desktops or end-user workstations, servers, printers and scanners. After a long period of discussions the management of the equipment has been transferred to I&A, a department of the Academy Office. Until 2010 we kept control of management ourselves. An important reason to outsource this task is that the Institute is too small to have the required knowledge and expertise at various levels adequately represented in the staff. Scaling up as regards contents and continuity will be the result of outsourcing. In the spring of 2010 it was decided to transfer maintenance to I&A. One of the consequences is that the helpdesk has been put at a distance (telephone helpdesk) and now operates on stricter procedures.
Enlarging possibilities for teleworking is also envisaged. The coming period will have to show whether a new organization of system management and helpdesk will come up to expectations. Above all, the rapid development of the Institute in the area of technology makes it imperative to follow recent developments in the field of storage and server maintenance. To achieve this we are in contact with SURF, SARA and BigGrid.

Clustering of academy institutes in the humanities

In the Academy, discussions have been going on for some time on clustering Academy Institutes in the area of the humanities at a university campus. Because of the relatively small scale of these institutes such an operation is partly motivated by the wish to create a greater critical mass and at the same time to create a sharper focus in research, documentation and technology. Furthermore, the ambition is to accommodate such a cluster in a university, so that there will be direct contact between the research institutes and research groups in the university and the talent blossoming there.

The Meertens Institute supports this view, but does not believe that a sharper focus and cooperation with other institutes and faculty research groups automatically emerges from joint research programmes, since the research themes of the various institutes are far apart from each other. Instances of cooperation as regards content are being realized but an enforced adjustment of research programmes based on a wish to scale up does not seem the proper way of arriving at a sharper focus on research in the Academy. In our opinion it is much more interesting to find the commonality in working together on collections and technological innovation in humanities. For institutes harbouring collections (IISG, KITLV, NIOD and the Meertens Institute) it stands to reason that the technological innovation in the area of documentation discussed above in section 5.2.iii constitutes the basis of a physical or virtual clustering of institutes. There are also chances for achieving a greater commonality in the area of applications (section 5.2.iv) and infrastructure (section 5.2.v). For example, DANS, Fryske Akademy, The Huygens Institute, DANS and the Meertens Institute are closely involved in a number of (inter)national activities in these areas, as becomes clear from the Academy programmes Alfalab and Computational Humanities and the ESFRI programmes CLARIN and DARIAH (CLARIAH). The Meertens Institute will support a greater commonality in this field and
will argue that these aspects, jointly with the e-humanities programme of the Academy, constitute an excellent basis for a clustering of humanities institutes.

### Agenda for technological innovation for the period 2013-2018

**Positioning the Institute as an (inter)national digital knowledge centre for Dutch language and culture**

- A central position for the relationship of research and documentation through technological innovation;
- Participating in discussions on technological innovation of research in humanities, among other things in the Academy programme Computational Humanities;
- Developing and executing projects in the framework of technological innovation (for example, modelling);
- Digitalization of as many collections as possible that are not yet available online within the framework of research and documentation projects;
- Contributing to discussions on clustering of Academy-GSW institutes based on a common technological agenda for the innovation of research in the humanities.

**Developing a digital infrastructure**

- Further development into a national Clarin centre for the infrastructure in the humanities (“trusted centre”);
- Participating in new infrastructure initiatives such as CLARiAH;
- Participating in (inter)national discussions on achieving standardization of metadata and data;
- Introducing metadata into datasets as part of research and documentation projects;
- Enriching databases as part of research and documentation projects;
- Making database sets available in accordance with (inter)national directives (among other things, the attribution of persistent identifiers and making resources available for harvesting through OAI-PMH);
- The development and generalization of applications that can be employed in research in the Institute (including arranging access rights to individual resources).
Enlarging access to publications and data (open access)

- Continuing the operation of inserting full data on publication in METIS;
- Storing all full-text publications in a repository;
- Linking full-text publications to METIS through NARCIS;
- Making available the databases of the Meertens Institute in a portal to ensure user-friendly access;
- Promoting open access to digital publications and databases.

Enlarging possibilities for research and documentation by using interactive forms (web2.0)

- Developing new, interactive parts of the website;
- Further development of the ‘crowd sourcing’ paradigm as a method for involving the general public in data acquisition tasks;
- Further revision of the digital questionnaire of the Meertens Institute.

Working on innovation of the humanities by means of technological innovation

- Carrying out innovative research programmes in the area of language and culture in which new technological possibilities are exploited;
- Cooperating with groups in The Netherlands and outside which are active in this field.
6

6.1 The scientific community
6.2 Students and teaching
6.3 Politics and policy
6.4 Professionals and cultural heritage
6.5 The industry and the innovation platform
6.6 The interested layman
6.7 Agenda for valorization of research and documentation (for the period 2013-2018)
‘Valorization’ at the Meertens Institute

A recent much-discussed topic is the issue of valorization of research. This term, which has come to replace the earlier notion ‘relevance for society’, is explained and understood in different ways. The Advisory Board for Science and Technology Policy (AWT) defines valorization as follows:

Valorization is a process that sees to it that scientific knowledge can be used in daily practice. More precisely, valorization is about making research results suitable and available, so that the chances that third parties can use them increase.

[Alfa en Gamma stralen (Alpha and Gamma rays), AWT-report 70, March 2007, p.17]

Although the notion is often interpreted narrowly as the contribution of technical, medical and natural-science research to the economy by way of patents and technology, the AWT argues for a broader view of the term, extending valorization to Alpha- and Gamma-scientific research, referring for example to “the preservation of our social connections, culture and democracy.” In the strategic agenda of the KNAW (Royal Dutch Academy of Sciences) 2010-2015 it is also posited that it is of great importance for the Academy that the discussion on utilization of knowledge “encompasses the full width of science, in which especially the humanities and the social sciences should also be involved” (p.32).

The Meertens Institute also opts for a broad interpretation. Valorization is about the utilization of acquired knowledge by others and is not in the first instance about the question as to whether utilization yields economic profit. Taking this interpretation for granted, valorization is an important part of the tasks of the Meertens Institute. The research carried out is not only legitimized by its scientific value, but for an extra-university research institute like the Meertens Institute the value for society produced by its research is also important.

Several groups in society can make use of the Institute’s results in science and documentation. We distinguish the following groups: the research community, students, politics and management, professionals in the area of language and culture, the industry and the interested layman. In the following sections the role of the Meertens Institute vis-à-vis these different
groups will be discussed. A central notion in this discussion will be that the Meertens Institute is not merely a purely scientific organization but that, in addition to its important scientific mission, it has a social responsibility which derives from it, and which is turned into policy aimed at these different target groups.

The scientific community

An important target group for the Meertens Institute is the scientific community. For this group, valorization mainly deals with making available research data in the area of language and culture. The documentation task has a direct bearing on this. As described in chapter 5 'Innovation and Infrastructure', the Institute robustly invests its resources in putting the available research data and collections at the disposal of researchers, digitally and in open access.

A good example of research valorization is the ESF project Edisyn (European Dialect Syntax 2005 - 2012), which plays an important role in research on the syntactic properties of European dialects. In the previous years – in addition to a theoretical syntactic perspective – this project has managed to set up a European network of researchers on microvariation in syntax. Within this network, which is coordinated by the Meertens Institute, a common methodology for data collection, data storage, data retrieval and cartography has been brought about. Moreover, work has been carried out on setting up an online network of databases for syntactic dialect data, using a shared search engine. These activities are not part of research sensu strictu, nor are they activities which are fully within the compass of a specific research project. This is why we use the term valorization. It is our objective to keep supporting this network and the tasks going with it in the coming years, and preferably extend it.

For other projects in which databases play an important role it will be a point of close attention as well to produce the data with their applications in such a way that they may also be used by other researchers and other research projects. Examples of projects in which this matter plays a central role are the Music-Retrieval Project (Liederenbank) and the Estate Inventories Project (Boedelbank). This is how research acquires a surplus value. From the perspective of its documentation task, the Meertens Institute takes one of its responsibilities to lie in playing a leading role in bringing about a national centre for research data and applications in the humanities. In this context it is fitting that the Institute should seek a role as a knowledge centre for language and culture.
Besides digital documentation, there are other ways of facilitating utilization of knowledge for the research community. Various researchers of the institute are editor of journals and book series: Etnofoor (Balkenhol), Nederlandse Taalkunde (Broekhuis), Volkskunde (Elpers), Quotidian (Elpers), Taal & Tongval (Hermans), Studies in Language Variation (Hinskens), Newsletter SIEF (Margry), Meertens Ethnology Lectures (Margry), Neerlandistieken (Van Oostendorp), De Achtziende Eeuw (Roodenburg), Nederlands Kunsthistorisch Jaarboek (Roodenburg), Studies in Netherlandish Art and Culture (Roodenburg), Trefwoord (Van der Sijs).

Institute staff members are involved as editors in wide-ranging reference works in their research fields. Marc van Oostendorp was editor-in-chief of the Phonological Companion, a 5-volume reference work offering a detailed international survey of the state-of-the-art in phonology (The Blackwell Companion to Phonology, 2011). Frans Hinskens, in cooperation with Johan Taeldeman (Professor emeritus, University of Ghent), is engaged in putting together a comprehensive survey of articles on variation linguistics in the Dutch language area (Handbook Language and Space: Dutch; Mouton De Gruyter, 2012/2013). In 2012 the first volume has appeared of a survey edited by Irene Stengs on traditional Dutch festivity celebrations. Similarly, Herman Roodenburg is editor of one of the six volumes of A Cultural History of the Senses (Berg Publishers, 2013/2014), and also of an Oxford Handbook of Cultural History that is expected to appear in 2015/2016. In the coming period activities such as these will continue to be supported, albeit on condition that they are strong in quality, are prestigious and do not obstruct the course of research.

The development of the Language Portal Dutch/Frisian (NWO-groot 2011-2015) should be seen in a similar light (see also chapter 3.VI). Among other things, it is concerned with the development of a detailed description of the grammars of Dutch and Frisian written in English and based on sound science. Existing grammars, such as the Modern Grammar of Dutch, which is in development, (senior author Broekhuis; two volumes have appeared in 2012; more than 2300 pages are completed) and the two books The Morphology of Dutch (Booij 2002, O.U.P.) and The Phonology of Dutch (Booij 1995, Clarendon) are taken together, are related to each other and are remodeled into a coherent, digital unity in order to offer to the international research-community an internationally unique and complete survey of the grammatical properties of Dutch and Frisian. There are serious plans to extend the Language Portal with a new digital version of the standard grammar of Dutch, the Algemene Nederlandse Spraakkunst (ANS). The Nederlandse Taalunie has made serious funds available (€700,000) to rework the electronic version of ANS to a digital grammar written in Dutch – starting in 2013 – that will be included in the Language Portal.
In 2013 the NWO-groot project Nederlab (2013-2017) will start, of which the Meertens Institute is the coordinator (see also chapter 3.V). This project aims at the creation of an infrastructure for historical Dutch texts for linguistic, historic and literary research. It will develop a corpus, search technology and tools that are brought together in a user-friendly interface enabling a distributed search of available diachronic corpora, at text and metadata level. We expect that Nederlab will become a vital portal for the research community that is making use of Dutch historical data.

The documentation which the Meertens Institute harbours in the area of Christian names, family names and place names (field names among others) remains a difficult point. The Meertens Institute has – on quality grounds and after evaluation by a (partly external) committee – terminated research in the area of onomastics. Also outside the Institute there is no or hardly any research in the Dutch research community that justifies substantial investments in the documentation of the study of names. The question is now what should happen to the Institute’s extensive documentation store on the study of names and consonant activities in the area of documentation. The Meertens Institute has been entrusted with tasks in the study of names since 1948. The fact that research on names is being terminated might be an inducement to also terminate related documentation tasks, in line with the Institute’s general policy that documentation should be linked to current or future research.

As of 2012 the Dutch Family Name Bank has been transferred to to Central Bureau of Genealogy (CBG) in Den Hague, for the first 5 years as a joint enterprise with the Meertens Institute; after that, the CBG will take over. In the coming period it will be investigated what we should do with other documentation in this field. A discussions on the tranfer of the Dutch Christian Name Bank to the CBG has been started. This point clearly demonstrates that the main objective of the Meertens Institute lies in its research activities: although the name banks are among the most popular parts of the website, the Institute is in the proces of transferring these documentative data banks to other organizations, on the condition that the documentation of relevant data is being continued by the new owners.

As a research and documentation centre the Meertens Institute has no primary task in higher education. Nevertheless, it is important to be involved in teaching. As Robbert Dijkgraaf (former Academy president) properly
observed at the end of his annual address of 2010: “Educate young people who have woken up to life, who wish to understand the world and wish to solve big problems using their knowledge, some of whom can take over the torch of research and carry it forward, that is what it is all about for us.” (The climate for science; annual address 2010; Royal Dutch Academy of Sciences, p.18). For several reasons it is important for the Meertens Institute to be in a feeding relationship with higher education. Firstly, it is important for the future to share the developments in research taking place at the Meertens Institute with the university research community, of which notably graduate students — and research MA and PhD students in particular — constitute a prominent group. This would lead to a greater spread of acquired insights. Secondly, it is important to be involved in the development of talent, not only as a goal in itself, but also for taking care that there are young people who can take over the torch of research. Thirdly, young researchers often provide the impetus for innovation in research. Their input is extremely important, and for that reason it is desirable to awaken and develop their interest in the research areas of the Meertens Institute during their schooling. Fourthly, the databases of the Meertens Institute offer vastly more possibilities than the staff can avail themselves of. Teaching (classpapers, among others), MA-research and research for the doctoral dissertation can provide a direct contribution to the description and analysis of the material present in the Institute.

The Meertens Institute is involved in the education of young researchers in several ways. Firstly, the Institute is actively engaged in appointing professors with endowed chairs. At the moment nine Institute staff members occupy endowed chairs: Barbiers and Grijp at the University of Utrecht, Cornips at Maastricht University, Van Oostendorp at Leiden University, Hinskens and Roodenburg at the Free University of Amsterdam, Bennis and Margry at the University of Amsterdam, and Van der Sijs at the Radboud University Nijmegen. These researchers are allowed to be active in different Faculties for 0.2 of their appointment, specifically with respect to teaching. These positions are also intended to give shape and substance to cooperation with university research groups. Now that universities are introducing a more stringent and more principled policy with regard to appointing professors with endowed chairs, it is important to let the research qualities of the Meertens Institute be reflected in the number of these positions. The ambition is that all tenured researchers should be professors with endowed chairs. Secondly, in addition to the endowed chairs, the Institute is also involved in other activities in university teaching. Institute researchers are seconded to different universities on a regular basis. Most researchers are regularly involved in graduate teaching in graduate schools, nationally and internationally (among others, in the Netherlands Graduate School of Linguistics and the Huizinga Institute). It is precisely
at the highest level of teaching that researchers of a scholarly organization like the Meertens Institute should be invited nationally and internationally with a degree of regularity, in order to share their results and insights with future researchers.

trainees

Students often get involved in research at the Meertens Institute while doing a work placement. The Institute frequently offers placement positions to students from different universities. These positions are very much in demand given the fact that the ‘placement room’ in the Institute (eight work stations) is continually occupied and trainees need to be placed with current research projects elsewhere. The past had shown that the Institute’s trainees regularly return as PhD students, and this is one of the reasons that makes offering placement positions in research projects a useful investment.

secondary education

The Institute further invests in the ‘Language Quiz’ prepared by the Language Studio (Taalstudio) by supporting it financially and by producing questions for the quiz. The Language Quiz is distributed mainly among pupils in the higher years of secondary education. The Institute also contributes financially and content-wise to the column ‘Language and Speech’ of Knowledge Link (Kennislink), a website also set up for secondary school pupils.

student meetings

Finally, the Institute organizes annual instruction meetings on current research for students from various universities. Staff members are also regularly invited to give lectures at student meetings or to sit on panel discussions on language and/or culture organized by students. These meetings are important for students to acquaint themselves with the Institute’s research.

Politics and policy

societal relevance

Research at the Meertens Institute finds itself located in an area that at various points has a great and direct relevance for important issues in our society. In public debate we have seen a revival of the interest in popular culture over the last few years, even in Dutch Parliament. Directly related to this is the issue of the Dutch identity in a world of ever-increasing complexity in which frontiers are more and more losing their former significance and do not offer any certainty. The often emotionally laden issue of national or regional identity is closely connected with processes of migration and cultural and linguistic integration. It follows that studying topics like language change, second language acquisition, cultural assimilation, identity processes, religious and cultural rituals,
Multi-cultural society and popular culture are closely related to current central issues in our society. Therefore, it is important that Meertens researchers contribute to the discussion on these questions in order for acquired and stored knowledge to play a greater role in political and social discourse.

In the previous period the Institute was involved in the political arena. An example: the discussion in Parliament on the notion of ‘popular culture’ was partly based on the definition provided by the Meertens Institute. The ensuing cultural-political debate arising from this has resulted in the book *Splitsen van Knopen? Over volkscultuur in Nederland* (2009), to the realization of which the Institute has made a significant contribution. Connected with this is the book *Immatereel Erfgoed en Volkscultuur; Almanak bij een actueel debat*, which appeared in 2011, of which the major part was written by members of the Institute. It provides entries on all relevant concepts related to intangible cultural heritage.

The Institute also plays a role in the discussion on the Dutch contribution to the Unesco Treaty on Intangible Cultural Heritage, which was ratified in 2012. This deals with the question as to how to register Dutch intangible cultural heritage – *Sinterklaas* is the stereotype example – with Unesco for protection and what the consequences of this are. The Nederlands Centrum voor Volkscultuur en Immaterieel Erfgoed (Dutch Centre for Popular Culture and Intangible Heritage (vE)) has put itself forward as ‘competent body’, the organization which takes care of the implementation of the Unesco Convention in a national context, while the Meertens Institute acts as its scientific conscience as member of the steering committee. In its upcoming project on “Dutchness” (see Chapter 3, VII), the Institute will also investigate this registration.

As for the multicultural society: here the Institute fairly regularly steps outside the groves of Academe. The Institute takes part in social and political discussions on the role that immigrants and their descendants play with respect to Dutch language and culture and on the way in which a multicultural society can be organized as responsibly as possible. Research in these areas provides an inducement to allocate acquired and stored knowledge to the debate. An example is the popular publication “Eigen en vreemd; meertaligheid in Nederland” (2012) by Cornips.

Other fields in which the Institute joins in the political discussion are the efforts of provinces and regions to get their dialects accepted as regional languages or minority languages, in keeping with the rules of the Council of Europe. The Institute will also, via the Language Portal, join in the discussion on grammatical norms for standard Dutch, an extremely sensitive social and politi-
The Language Studio

A special case of political involvement is the cooperation with the Taalstudio (Language Studio) in the area of language tests for foreigners set by the Immigration Service (IND) in order to determine the origin of asylum seekers (LADO). The Language Studio looks after the counter-assessment of the sometimes rather negligent reports drawn up by the Immigration Service. The Language Studio makes use of meeting rooms in the Meertens Institute and a while ago an investigation was started into the ways in which these language tests can be set and adjudicated in a maximally trustworthy manner. In 2011, an MA-course on LADO was taught at the University Utrecht by researchers of the Meertens Institute, in cooperation with the Taalstudio, the Immigration Office (IND) and the Netherlands Forensic Institute.

scientific conscience

Although the Institute participates in political debate on an incidental footing and joins in issues of policy, it is still insufficiently visible in the discussions. Given that the research subjects of the Meertens Institute – Dutch language and culture – have a prominent place in social discourse, it stands to reason that the Institute should manifest itself actively as the scientific conscience of the nation, following and criticizing the policies and central notions involved, and bring to the front knowledge present in the Institute with a degree of authority. In the coming period it is the Institute’s ambition to take on this role more clearly.

Professionals and cultural heritage

The group of people making use of the result of scientific research in the field of language and culture in their daily work is rather diverse. It is an important group since the contents of their activities are partly determined by underlying insights from science. In the field of ethnology the Meertens Institute works closely together with the Dutch Centre for Popular Culture and Intangible Heritage (VIE) in Utrecht. The goals of VIE are to make accessible and to promote (historical) backgrounds of daily life and intangible cultural heritage. VIE supports organizations in the areas of popular culture, regional history, folklore and folk art. The intention is that in the coming period there will be further discussions on cooperation between the Meertens Institute and VIE.

Centre for Popular Culture

In Belgium the Institute joins forces with FARO, a Flemish point of support for cultural heritage. This is a service organization playing a mediating role.
The direct connection of the Meertens Institute with cultural heritage organizations in the country is self-evident. From early days onwards, there have been good contacts with the Dutch Open Air Museum in Arnhem (NOM) and, to a slightly lesser extent, with the Zuiderzeemuseum (ZZM) in Enkhuizen. Jointly with the NOM, plans are being developed to consider to what extent the available (digital) documentation in the Institute (dialects, songs and stories) can be deployed as an enrichment of the Arnhem museum collection. In a way, the Meertens Institute itself is considered to be an organization in the field of cultural heritage since the Institute harbours large national collections in the area of language and culture. Although there are no special objectives with respect to extending these collections, other than those emanating from research projects, the total of collections steadily increases, frequently because of bequeathments (for example, the private collections of the dialectologists Daan and Weijnen), but also otherwise. Recently acquired collections are, among others, the archive of a circus (Maurice Veldkamp) and a collection of stories (Lai Fan Kwik). The decision as to accepting a collection is based on...
criteria such as relations with current research, potential interest for research in language and culture in The Netherlands, association with or extension of existing collections, the state of the collection and its uniqueness.

**CATCH-projects**

The Institute presents itself as a cultural-heritage organization in other ways as well. The NWO program CATCH (Continuous Access To Cultural Heritage) aims at entering into relations with heritage organizations on the one hand and research in the humanities and information science on the other. In the CATCH-projects Witchcraft, Fact and Cogitch, the Meertens Institute appears as a heritage organization and work is carried out on the basis of the collection of songs and stories maintained by the Institute.

**Camerata Trajectina**

The Dutch Song Database has been important for the repertoire of the well-known, early music group Camerata Trajectina (specialized in Dutch music from the Middle Ages and the Golden Age), led by Louis Grijp. In reconstructing songs for their performances and recordings (e.g. in *The seven sins of Jeroen Bosch* (2009) or *Theater music from the Dutch Golden Age* (2007), *De triomferende min* (2012)), they make use of the collection and the advanced search options that have been included in this digital database.

**eLab Oral Culture**

Finally, the centre of expertise eLab Oral Culture (cf. chapter 3.1) must be mentioned in this section: this centre combines documentation and research on the basis of collections of an immense size. The digitally accessible databanks belonging to this centre – the Dutch Song Bank and the Dutch Folk Tale Bank – provide a good survey of the oral heritage of The Netherlands. In the near future this centre will continue to research and make available tales and songs. The book *Verhalen van stad en streek. Sagen en legenden in Nederland* (Bert Bakker, 2010), edited by Theo Meder and others, was one such example. These databanks constitute an important part of the Dutch cultural heritage.

**The industry and the innovation platform**

In 2012 a new perspective on the relation between humanities research and economic innovation has been introduced in the discussion on the innovation platform, driven by the Ministry of Economy, Agriculture and Innovation. A number of so-called topsectors have been selected by the government, such as water, agri&food, and high tech. The ministry supports these nine sectors by additional funds and requires funding organizations (such as NWO and TNO) to invest substantial parts of their funding means into projects that fall within...
the bounds of these sectors. For the Meertens Institute, the topsector Creative Industries is relevant. This sector is divided into seven networks of which Cultural Heritage is one. Projects that combine research and innovation and that consist of coalitions of research organizations, heritage organizations and industry may apply for substantial funds. The Institute is the coordinator of the network Cultural Heritage and is represented in the steering group, together with the institute for Sound & Image and the Cultural Heritage Agency of the Netherlands. We expect the coming period to be engaged in a number of projects in which research, heritage and commercial institutions are together involved in the creation of new products on the interface of science and industry.

The interested layman

The items on the agenda of the Meertens Institute are not solely important to researchers, students, policy makers, professionals and heritage experts. The public at large also strikes up an acquaintance with and makes use of the data and research of the Meertens Institute. It is important that the Institute should pay attention to this group. A good example of activities in this area is the column of Nicoline van der Sijs in the Wetenschapsbijlage of the newspaper NRC-Handelsblad. But it is not just the readers of the science contributions in NRC-Handelsblad. For instance, research on language deterioration or the Dutch identity is interesting and important for a much broader public. Contributing to public opinion is a task which by its very nature belongs to an institution like the Meertens Institute.

A further public task concerns the dissemination of scientific insights in areas where the average Dutchman thinks that they can/must have an opinion. Dutch non-linguists entertain all kinds of staple ideas on good or bad Dutch, on language norms and on the status of dialects, which only marginally correspond to a scientific perspective on these issues. The same holds for notions like (intangible) Dutch cultural heritage, the Dutch identity and multiculturalism. It is the task of the Meertens Institute to provide good and goal-directed information that puts the scientific perspective on these and similar issues on stage in an accessible and attractive manner.

Finally, it is also possible for the Meertens Institute to inform the public at large about developments in the area of popular culture and language by producing attractive publications, web pages and activities. In the coming period the Insti-
Institute will keep attempting to produce such results as a derivative of research and documentation.

**new-year publications and website**

The last few years the Institute has been successful in the public arena. To illustrate: the annual new-year publication has proved to be a success for a wide audience (some examples are *Hollandser dan Kaas* (More Dutch than Cheese) in 2009, *Een echte Amelander spreekt dialect* (A true inhabitant of Ameland speaks dialect) in 2010, *Gelukkig Nieuwjaar* (Happy New Year) in 2011, and *Eigen en Vreemd* (Native and Foreign) in 2012. The Institute’s website is quite popular with many levels of the population (the recently revised web pages on family names and Christian names drew millions of visitors; it looks very much like it that the Christian-name site may influence the way in which parents choose names for their children). The Institute is often in the news (newspapers, radio and television) and nearly always in a positive way. The introduction of a new website on the Hijacked Letters (http://www.gekaaptebrieven.nl/) as a contribution to the national Kennismaand Oktober 2012 (Science Month October 2012) was very well received in the media and the website has been visited frequently.

**Academic Annual Prize**

In 2009 the Institute managed to reach the final round of the Academic Annual Prize (*NRC, NWO and KNAW*) and the Institute’s researchers are regularly invited to deliver lectures to a wide audience or to sit on opinion panels.

**NWO-meerwaarde**

In 2011 the Institute has managed to rope in two grants from a new NWO fund – *NWO-Meerwaarde* (*NWO-Surplus-Value*). *NWO-Meerwaarde* is geared towards making results from scientific research accessible to a wide audience. The applications within the framework of the Language Portal (*Utilization of Knowledge Language Portal*) and the Institute’s linguistic atlases (*‘De Taaldetector’* (The Language Detector) http://www.meertens.knaw.nl/taaldetector/) have been implemented in 2012.

**PR-plan**

The Institute PR policy is not based on a PR-plan but has more of a made-to-measure character: Every researcher should with a certain degree of regularity ask themselves to what extent their research could be made interesting for a wide audience. The Institute has a PR-staffmember who can advise and stimulate researchers to contribute to this task. The Institute’s Board considers this fundamentally important, and this is why time is made available to spend on this kind of activity. Since this policy has proved to be successful, it will be continued in the coming period.
Agenda for valorization of research and documentation (for the period 2013-2018)

**General**
- contributing actively to the valorization of research and documentation from the perspective of a broad interpretation of valorization;
- contributing actively to valorization as a social responsibility derived from the Institute's scientific mission;
- reconsideration of the Institute's role with respect to onomastics.

**Valorization and research**
- growing into a national (digital) knowledge centre for language and culture;
- supporting the centre of expertise for Dutch Songs and Dutch Tales;
- building a representative corpus of historical Dutch texts suitable for research (Nederlab; NWO-groot 2013-2017);
- maintaining and extending the European network for research in syntactic and phonological variation of language;
- developing the Language Portal Dutch/Frisian.

**Valorization and teaching**
- extending the Institute's role in PhD-teaching;
- maintaining or extending the number of placement positions.

**Politics and policy**
- contributing to the political debate and to discussions on policy in a more prominent and authoritative manner;
- taking the initiative in the discussion on grammatical norms in Dutch;
- participating in the discussion on the Dutch role with regard to the Unesco convention on intangible cultural heritage.

**Professionals and cultural heritage**
- cooperating with VIE and FARO on the basis of a clear position and concrete agreements;
- maintaining/extend cooperation with the museums NOM, ZZM etc.

**The Industry**
- further developing the network Cultural Heritage in the context of the topsector Creative Industry (Innovation Platform).

**The interested layman**
- maintaining, and if possible, extending the Institute's role for the public at large.
The Meertens Institute is a research institute of the Royal Netherlands Academy of Arts and Sciences.