Dutch-Flemish research programme for Dutch Language and Speech Technology

STEVIN Final Internal Evaluation

March, 2010

Introduction

The STEVIN Programme Committee (PC) was requested to carry out a self-evaluation as one ingredient for the final evaluation of the STEVIN Programme.

Nine questions were formulated that the PC was requested to answer. The section

Self-evaluation **questions** lists the original questions (in Dutch). The answers the questions, together with English translations of the questions, are provided in separate sections, one for each question.

Self-evaluation questions

The following is the list of the questions for the PC self-evaluation as determined by the TST-Board (in Dutch):

Evaluatievragen STEVIN Eindmeting
1. Accountability • Werden de middelen door STEVIN in de periode 2004-2009 op een correcte en adequate manier aangewend? • Waren de voorziene middelen (op projectniveau) al dan niet voldoende?
2. Effectiviteit • Heeft STEVIN de doelstellingen zoals neergeschreven in het STEVIN-meerjarenplan gehaald? • Welke vooropgestelde activiteiten of doelstellingen werden niet gerealiseerd, en waarom niet? • In welke mate hebben de projecten/flankerende activiteiten/… die STEVIN heeft gerealiseerd in de beschouwde periode bijgedragen tot de realisatie van de vooropgestelde beleidsdoelstellingen zoals bondig omschreven in bijlage 1 punt 3? • In welke mate zijn de projecten/flankerende activiteiten die STEVIN heeft gefinancierd/georganiseerd in de beschouwde periode een concrete en onderbouwde vertaling van de doelstellingen (in het STEVIN-meerjarenplan)? Welke belangrijke doelgroepen heeft STEVIN hierbij geïdentificeerd en op welke manier werden deze benaderd? • In welke mate zijn de doelstellingen van het STEVIN-meerjarenplan adequaat vertaald in de jaarwerkplannen, projectoproepdocumenten e.d. van STEVIN?
3. Meerwaarde/relevantie • In welke mate zijn de projectoproepen/flankerende activiteiten die STEVIN heeft gerealiseerd in de beschouwde periode doorgedrongen tot het TST-veld (onderzoekers en TST-ontwikkelaars) én (specifieke) gebruikersdoelgroepen in het bijzonder. In welke mate leidden de projecten tot bruikbaar materiaal voor het veld en (maatschappelijke) gebruikersdoelgroepen? • In welke mate is de technologische en wetenschappelijke stand van zaken in het TST-domein geëvolueerd door de projecten/flankerende activiteiten die STEVIN heeft gefinancierd/georganiseerd? • Welke meerwaarde heeft STEVIN al dan niet voor de wetenschappelijke wereld, de industrie en de maatschappij opgeleverd? Welke zijn de aanwijsbare verbeteringen in de TST-infrastructuur als gevolg van het STEVIN-programma?

4. Additionaliteit (zowel input-, output-, als gedragsadditionaliteit) • Welke voor de start van STEVIN gesignaleerde knelpunten in het TST-domein werden dankzij het STEVIN-programma aangepakt, en met welke mate van succes? Welke niet?
5. Efficiëntie • Heeft STEVIN de doelstellingen zoals neergeschreven in de overeenkomst EWI-NTU en STEVIN-meerjarenplan gehaald op een aanvaardbare/efficiënte manier?
6. Positionering • In welke mate is er eventueel overlap met andere activiteiten/projecten van andere actoren? • Welke rol speelt STEVIN binnen het domein van het taal- en spraaktechnologieonderzoek in het Nederlandse taalgebied en internationaal? • In welke mate werd taal- en spraaktechnologie op de beleidsagenda van relevante overheden gezet.
7. SWOT-analyse • Wat zijn de sterke/zwakke punten van STEVIN? Welke uitdagingen en bedreigingen dienen zich aan in de nabije toekomst, rekening houdend met het ruime beleidskader (zowel in Vlaanderen en Nederland als daarbuiten)? • Wat is de evolutie t.o.v. de SWOT-analyse uit de M&I Partners-voorstudie (zie bijlage 1 punt 3)?
8. Monitoring • Op welke manier werd/wordt de performantie (zoals efficiënte en effectieve uitvoering) van STEVIN gemonitord? • Was deze adequaat?
9. Governance • Op welke manier werd/wordt het STEVIN-programma georganiseerd? • Was deze manier adequaat (SWOT-analyse), efficiënt, effectief?

In the sections to follow we will provide English equivalents of these questions and the answer to these questions by the PC.

Question 1: Accountability

- **Werden de middelen door STEVIN in de periode 2004-2009 op een correcte en adequate manier aangewend?**
- **Waren de voorziene middelen (op projectniveau) al dan niet voldoende?**

In English, these questions can be formulated as follows:

- **Have the resources been used in a correct and adequate manner by STEVIN in the period 2004-2009**
- **Were the provided resources sufficient or insufficient (on a project level)**

The PC operationalizes the terms "correct and adequate" as "in line with the original multiyear work programme" and/or plans formulated in the initial phase of the programme. With this operationalization it can be stated that the resources have been used correctly and adequately.

The overall STEVIN resources were globally distributed over 5 items, of which R&D projects (>76%) and demonstration projects (>8.5%) are the largest, in accordance with the plans in the multiyear programme. Supporting activities for creating new networks and transferring knowledge occupied almost 6%. The Dutch HLT Agency (>2.5%) and STEVIN managements 6.5%. This distribution is completely in line with the multiyear work programme and the initial budgets, though the costs for supporting activities were originally overestimated and have been reduced to bring those more in line with reality during the programme.

(See Factfile, Summary Dutch-Flemish STEVIN Programme budget, p. 5 for details)

The multiyear programme (section 5, esp. 5.1 and 5.3) aims to cover projects from all 4 layers in the layered Language and Speech Technology (LST) innovation model. And STEVIN indeed has covered all 4 layers:

1. Basic resources for LST (51.1%)
2. Basic Research (23.3%)
3. Application-oriented research (15.4%)
4. End-user targeted demonstration projects (10.2%)

(See Fact file, *The STEVIN Priorities*, p. 13-14 for details)

Though no strict geographically determined division of the available financial resources between Flanders and the Netherlands was required, it was still the aim to involve researchers and developers from both the Netherlands and Flanders, in a 2:1 proportion (section 7, p. 34). The actual proportion between the Netherlands and Flanders is 63%:37% for R&D projects, 66%:34% for demonstration projects, and 64%-36% overall (see Fact file, *Summary statistics STEVIN funding R&D and demonstration projects (in k€)*, p. 7 for details). The proportion for educational projects is 68%:32% and for master classes 100%:0%. (NB. only 3 educational and 2 master class projects were funded).

It was the goal to have the largest part of the programme carried out by knowledge institutes (section 5.2) but it was also a clear aim to involve SMEs in STEVIN, and to ensure that knowledge institutes and SMEs would participate in common projects, in order to create new networks and stimulate knowledge transfer (sections 5.2, 5.8, 6.1, 6.3). The actual proportion between knowledge institutes and SMEs is 83%:17 %, completely in line with the original target (see *ibidem* for details). SMEs participated in 13 of the 19 funded R&D projects. In total, SMES participated 16 times in the R&D projects, whereas KIs participated 65 times (i.e. a proportion of approximately 20%:80%).

Naturally, SMEs dominated in the demonstration projects, which were targeted at SMEs. SMES participated in all demonstration projects, 40 times in total, while KIs participate only 7 times in total (i.e. a proportion of approximately 85%:15%).

Master classes are organised completely by companies (100%), while in education activities knowledge institutes dominated (approximately 83% v. 17 %).

Both language and speech technology are represented in STEVIN. Some projects focus purely on speech technology, some on language technology, and others combine the two. In terms of funding the proportion is 53.1% for language technology aspects vs. 46.9% for speech technology aspects. So these are more or less equally represented, in accordance with the original aims. (See Fact file, *The STEVIN Priorities*, p. 14 for details)

As is clear from the STEVIN Fact File, many more project proposals were submitted than could be funded. Most of the unsuccessful proposals were also of high quality and would have been funded if we had had more financial resources. For R&D, in total 52 proposals were submitted (68 if we also count the Second Open Call pre-proposals), though only 19 could be funded (36.5% success rate).

Within the R&D projects, the calls for tender have relatively few submitted proposals. The proposal for the written corpus, however, proposed a project consortium which covered all players in the selected annotation domains in the Netherlands and Flanders (resulting in the project SONAR). The same holds for the proposal for a speech recognition toolkit (resulting in the project SPRAAK). For the semantic lexicon two competing proposals were submitted, one of which was selected.

For demonstration projects 41 proposals were submitted of which 14 were selected to be funded by a committee of civil servants (30% success rate).

For educational and master class projects, the success rate is higher: for educational projects, 5 proposals were submitted and 3 could be funded (60% success rate); for master classes, 3 proposals were submitted and two could be funded (67% success rate). Because of the lower number of submitted proposals for these types of projects, and because some of the submitted educational proposals were judged not to be appropriate as an educational project (but rather a demonstration project), the PC decided not to issue calls for educational and master class projects in 2010 but to spend the money in other ways.

In general, the resources provided for each project were sufficient. However, especially the SONAR and SPRAAK tender projects consortia- targeting high priority large complex resources - carried out work for which not all costs were covered (participants invested themselves as well).

Especially the SONAR project suffered from some additional problems. First, the funding bodies insisted on testing the fairness and openness of the tender procedure prior to opening the tender which required that the preparatory project D-Coi had to be completed and its results checked by external experts. The tender was opened 5 months after the D-Coi project formally ended, causing problems for maintaining the employees and the expertise they had been built up. Secondly, when in August 2007 the SONAR proposal after some revising was proposed for funding, the board decided to first only fund a start-up phase because not all requested funding could be committed as some informal Dutch financial commitments to the STEVIN budget had not been formally substantiated yet. Though in the end, all the requested money did become available, it had a negative effect on the efficiency of the SONAR project execution and led to quite some uncertainty and frustration among the involved researchers. On the other hand the start up phase focusing on acquiring data did result in an improved procedure and set-up for obtaining the required written resources and the licences to allow for optimal reuse of the data.

Finally, it can be stated that the financial resources have been used in a correct and adequate manner since the evaluation and monitoring of projects was transparent, impartial and objective. This will be elaborated on in more detail in the answer to Question 8.

Question 2: Effect

- **Heeft STEVIN de doelstellingen zoals neergeschreven in het STEVIN-meerjarenplan gehaald?**
- **Welke vooropgestelde activiteiten of doelstellingen werden niet gerealiseerd, en waarom niet?**
- **In welke mate hebben de projecten/flankerende activiteiten die STEVIN heeft gerealiseerd in de beschouwde periode bijgedragen tot de realisatie van de vooropgestelde beleidsdoelstellingen zoals bondig omschreven in bijlage 1 punt 3?**
- **In welke mate zijn de projecten/flankerende activiteiten die STEVIN heeft gefinancierd/georganiseerd in de beschouwde periode een concrete en onderbouwde vertaling van de doelstellingen (in het STEVIN-meerjarenplan)?**
- **Welke belangrijke doelgroepen heeft STEVIN hierbij geïdentificeerd en op welke manier werden deze benaderd?**

In English, these questions can be formulated as follows:

- **Did STEVIN achieve the goals as written down in the STEVIN work programme?**
- **Which planned activities or goals were not realized, and why not?**
- **To what degree have the projects and accompanying activities that STEVIN realized contributed to the realization of the targeted goals?**
- **To what extent are the projects and accompanying activities that STEVIN financed/organized in the relevant period a concrete and justified translation of the objectives (in the STEVIN work programme)?**
- **Which important target groups has STEVIN identified in this process and in which ways have they been approached?**

The main objectives of the STEVIN programme are:

1. *The realisation of an effective digital language infrastructure for Dutch, based on the BaTaVo priorities*
2. *The execution of strategic research in the field of language and speech technology*
3. *The creation of networks and the consolidation of language and speech technology activities, educate new experts, and promote discussion and transfer of knowledge*

We dedicate a separate subsection to each of these objectives.

STEVIN Objective 1:
The realisation of an effective digital language infrastructure for Dutch, based on the BaTaVo priorities

A 'digital infrastructure for Dutch Language Resources' involves at least three aspects:

1. Relevant language resources must exist or be created and guidelines, conventions, and best practices for creating these and other language resources must be made.
2. The relevant language resources must be stored and maintained, and these resources with their documentation, as well as guidelines, conventions and best practices must be made easily available to all researchers and developers in a non-discriminative way
3. The intellectual property rights (IPR) surrounding the language resources must have been adequately dealt with, so that the resources can actually be used by researchers and developers.

We discuss each of these aspects in turn.

Language Resources
In the STEVIN programme three open calls for research and development (R&D) projects have been launched. Because the availability of adequate resources is often a necessity for basic as well as for more applied research, several of the projects granted in the first call for proposals were focused on the collection and annotation of resources (both corpora and tools), and guidelines for creating them. Examples are D-COI and Jasmin-CGN. D-COI was meant as a pilot study for one of the main goals of the STEVIN programme: the creation of a huge (>500M words) Corpus of Written Dutch. Also several projects from other calls were focused on creating new or extending and enhancing existing language

resources (e.g. DPC, LASSY, STEVIN can PRAAT). Projects (from all calls) were always stimulated to yield language resources as a side benefit next to their their main research objectives, and many awarded projects have done so (examples are Autonomata, IRME, COREA, Daeso, DAISY, DUOMAN, PACO-MT).
In order to ensure the creation of a number of high priority language resources, three specific tenders have been issued:

1. one for the creation of a semantic lexicon, resulting in the project Cornetto
2. one for the creation of a state-of-the-art Dutch Speech Recognizer and speech recognizer toolkit, resulting in the project SPRAAK
3. one for a richly annotated corpus of written Dutch containing more than 500M words from a wide variety of sources, resulting in the project proposal SONAR

The SPRAAK and Cornetto projects have been carried out, and the SONAR project, which addresses one of the top priorities from the BaTaVo report, is well underway and has published its first release.

The original priorities for STEVIN Objective 1in the domains of language technology were more specifically:

1. A large corpus of written Dutch
2. An electronic lexicon
3. Parallel corpora

Projects which contribute(d) to the construction of these resources are:

- D-COI: preparation for the construction of a 500 million word corpus of written Dutch.
- LASSY: a 1 million-word tree bank of written Dutch, based on D-COI.
- CORNETTO: a lexical semantic database for Dutch, covering 40 thousand entries.
- DPC: creation of a 10 million word sentence-aligned parallel corpus for Dutch-English and Dutch-French.

Besides, there are projects which have the creation of annotated corpora or lexical resources as one of their secondary objectives:

- COREA: a limited amount of data annotated with co-referential information
- IRME: a lexical database of Dutch multi-word expressions, covering 5 thousand entries
- DAESO: an annotated monolingual Dutch parallel corpus of 1 million words
- DAISY
- DUOMAN
- PACO-MT

The original priorities of STEVIN Objective 1 in the field of speech technology were:

1. speech and multimodal corpora for:
   - applications such as CALL (Computer Assisted Language Learning);
   - applications in which names and addresses play an important role;
   - CCQA applications (questions and answers in call centres), educational applications;
2. multimodal corpora for applications of broadcast news transcription or person identification;
3. text corpora for the development of stochastic language models;
4. tools and data for the development of:
   - robust speech recognition;
   - automatic annotation of corpora;
   - speech synthesis;

Projects which address these objectives are:

- AUTONOMATA: a corpus of spoken person and address name utterances. Since many non-natives and young people are recorded, this corpus may also be relevant for CALL and educational applications for children.
- JASMIN-CGN: an extension of the Spoken Dutch Corpus with speech of children, non-natives, elderly people and human-machine interaction. It can therefore be seen as support for applications such as CALL and educational applications for children.

- D-COI and SONAR: for the construction of a large Dutch written corpus, which can, inter alia, be used for stochastic language modeling.
- SPRAAK: a speech recognizer toolkit and speech recognizer for Dutch
- STEVIN can PRAAT: extension of the functionalities of the PRAAT program

All bullets of the first objective for speech technologies can be said to have been satisfactorily addressed, with the exception of the second bullet: multimedia corpora have not been produced in the STEVIN programme. Fortunately, some (Dutch) organizations (TNO, Twente) were involved in the European FP6 projects AMI and AMIDA[1], and many were involved in the Dutch national programme IMIX[2], which covers some aspects of multimedia resources. There also were no real projects in the area of speech synthesis, in part because commercial speech synthesis systems are available for Dutch and of high quality. However, one can reasonably state that Autonomata can make a small contribution to improvements of the synthesis of names.
Overall it can be stated that the first objective has been addressed almost completely.

Availability, Accessibility, and Maintenance
In order to guarantee the availability of the language resources, both resources that existed prior to the STEVIN programme as well as resources that have been or are being created in the STEVIN programme are being made available via the Dutch HLT Agency (TST-Centrale). The Dutch HLT Agency has been started up in 2003, to serve as a central portal and service centre for Dutch language and speech technology resources.

The main tasks of the Dutch HLT Agency are acquisition and IPR, management and maintenance, and distribution and service. A pricing committee has been formed to support the HLT Agency in pricing matters. The current pricing policy distinguishes (non-commercial) research and commercial use. Commercial use prices are in accordance with the market, for (non-commercial) research resources are freely available.

IPR
To allow for the most optimal reuse of resources, the Board had stipulated that: a) all IPR to STEVIN results had to be transferred to the Dutch Language Union (the original owners retained a free licence to use the results themselves) and b) that project consortia had to ensure that licences to background knowledge and to incorporated data and tools were also obtained and transferred to the Dutch Language Union as well. Ensuring the correct treatment of intellectual property rights (IPR) is a non-trivial issue, especially if it has to be done by researchers and developers who are not particularly knowledgeable in this area. The newly created Dutch HLT Agency attempts to support the researchers, but also for them most of the matters relating to this aspect were new. This caused problems and delays for several of the projects, especially (but not exclusively) from the first call. For this reason, the programme committee proposed to set up an IPR committee, with members with some experience in this field from the programme committee (including the Dutch HLT Agency representative) and HLT Board, supported by a lawyer with expertise in this area. This committee has given advice to several projects (for example D-COI, DPC, SPRAAK) and produced several other results, in particular a leaflet (c.f. STEVIN Fact File) to help convince external data providers (e.g. publishers) to make available their data and give them the confidence that doing this is not harmful to them or hurting their normal business interests, a schematic overview of the relations between actors involved (c.f. STEVIN Fact File) and a number of standard contract templates[3] that can be used by project coordinators as a basis for dealing with IPR issues in an adequate manner. In addition, it has been decided that an Open Source strategy towards the IPR issues is a possibility within STEVIN, and some projects have delivered or will deliver their results in this way (e.g. SPRAAK). These results have been very useful, but they do not imply that all IPR issues are now solved, and it is certainly true that IPR issues have led to delays in delivering some of the project results.
Guidelines, conventions and best practices have been created in several projects, and are also available via the HLT Agency. They have of course also been made available in the form of scientific publications, e.g. in LREC conferences and proceedings.

[1] http://www.amiproject.org/
[2] http://www.nwo.nl/imix
[3] http://www.inl.nl/index.php?option=com_content&task=blogcategory&id=103&Itemid=5

STEVIN Objective 2:
The execution of strategic research in the field of language and speech technology

The first call already resulted in some projects carrying out strategic research, including COREA, IRME, and Autonomata. The second call for proposals was mainly focused on strategic scientific research as can be seen in the overview of the granted projects in the STEVIN Fact File. The third and last call for R&D proposals was focused on strategic applied research, using where possible the results of the previous projects. All five granted projects involve applied research.

The priorities of STEVIN objective 2 in the domain of language technology were:

1. semantic analysis (tagging, integration with syntax and morphology)
2. text pre-processing (tokenization, spelling correction, named entity recognition, ...)
3. morphological analysis (compounding and derivation)
4. syntactic analysis: a robust parser for Dutch

Projects which primarily address these priorities are:

- COREA: development of a robust system for automatic co-reference resolution.
- IRME: methods and tools for the automatic identification and representation of multi-word expressions.
- LASSY: further development and refinement of a robust parser for Dutch
- DAESO: tools for automatic alignment, classification of semantic relations and text-to-text generation
- DAISY: aims to develop and evaluate essential technology for automatic summarization of Dutch informative texts.
- DUOMAN: aims to transform the volumes of online information that threaten to leave media analysts information-bound into aggregates of attitudes organized by topic by employing classification, information extraction, and cross-document linking.
- PACO-MT: aims at building a hybrid machine translation system (for Dutch-English and Dutch-French) combining the positive features of corpus based and rule based systems

Most of these projects cover aspects of semantic analysis. Parts of D-COI cover other aspects of semantic analysis (thematic roles and spatio-temporal relations), as well as various aspects of text pre-processing. LASSY combines the construction of a treebank with the further development and refinement of a robust parser for Dutch. A major lacuna is the absence of a project on morphological analysis for derivation and compounding.

The original priorities for STEVIN Objective 2 in the field of speech technology were:

1. robustness of speech recognition;
2. output treatment (inverse text normalization);
3. confidence measures;
4. adaptation;
5. lattices.

They are addressed by:

- AUTONOMATA: grapheme to phoneme conversion for Dutch and Flemish names.
- MIDAS: noise reduction techniques and confidence measures, to be integrated in SPRAAK
- N-BEST: benchmarks for the evaluation of speech recognizers for northern and southern Dutch
- SPRAAK: a speech recognizer toolkit and speech recognizer for Dutch
- DISCO: development of a prototype of an ASR-based CALL application for training oral proficiency for Dutch as a second language
- AUTONOMATA TOO: to build a demonstrator version of a Dutch/Flemish Points of Interest (POI) information providing business service, and to investigate new pronunciation modeling technologies that can help to bring the spoken name recognition component of such a service to the required level of accuracy

So, also for the second objective it can be stated that it has almost completely been addressed. Three out of four projects have already finished, so also the realization of this objective is well

underway. This does of course not mean that e.g. the problem of robust speech recognition has now been completely solved. There are many different situations in which adverse conditions make speech recognition a challenge, and within STEVIN only a few of such adverse conditions could be addressed. Similarly, JASMIN-CGN created a speech database for elderly people and children, neatly complementing the CGN (Spoken Dutch Corpus)[4], but this does not exhaust all possible speaker groups that might require special resources or research in speech recognition.

The BaTaVo report also listed the creation of bench marks as an important priority. The STEVIN work programme does not, though it does state that the creation of test suites, test data and benchmarks can be part of individual projects and often must be to allow for systematic evaluation. In actual practice, only for speech a separate project dedicated to creating a benchmark was proposed (and awarded): N-BEST. But in most other projects, data sets have been created that were used internal to the project for testing purposes and can serve as a reference for other researchers. In the area of data, most projects creating data have had the resources they created validated by external parties, in accordance with partially well-established and partially newly developed validation guidelines.

Applications
The STEVIN work programme lists a number of potential (classes of) TST applications as illustrations. STEVIN projects can contribute to the realization of such applications in many ways, sometimes very indirectly, e.g. by creating resources required for the development of the technology underlying the application or by doing fundamental research on underlying technologies; they may contribute more directly by carrying out application oriented research; and they may even be involved in the creation of the application itself, e.g. in demonstration projects. We briefly list the example TST applications from the STEVIN work programme and describe how various projects can contribute (or have contributed) to such applications:

- **Information extraction from Speech**. Projects that contribute to this class of applications include the R&D projects AUTONOMATA, JASMIN-CGN, SPRAAK, STEVINcanPRAAT, N-BEST, AUTONOMATA TOO and MIDAS. The demonstration projects Rechtspraakherkenning, NEON, and SNRT address this type of application directly
- **Detection of accent and identity of speakers**. Projects that contribute to this class of applications include JASMIN-CGN, SPRAAK, and DISCO. The demonstrator project Rechtspraakherkenning includes this aspect in its application. Finally the educational project Diademo illustrates the problem and a possible solution in a demonstrator at the Flemish documentation centre for science and technology Technopolis[5] for educational purposes.
- **Extraction of information from (monolingual or multilingual) text**. Projects that contribute to this class of applications include COREA, IRME, D-COI, SONAR, DPC, LASSY, CORNETTO, and PACO-MT. It is the focus of application-oriented research in DAESO and DUOMAN, and the demonstration projects Gemeenteconnect and YourNews.
- **Semantic web**: Projects that contribute to the semantic web include D-COI and SONAR, but especially CORNETTO.
- **Dialogue systems and Q&A solutions** Projects that contribute include especially DAISY and DUOMAN, as well as the demonstration projects Gemeenteconnect and Web Assess.
- **Automatic summarization and text generation** Projects that contribute to this call of applications include D-COI and SONAR, but especially DAESO and the demonstration project Web Assess.
- **Automatic Translation** projects that contribute to this class of applications include D-COI, SONAR, LASSY, IRME, COREA, CORNETTO and more directly DPC and PACO-MT.
- **Educational systems** projects that contribute include very indirectly all resource creation projects, but more specifically DISCO. Many demonstration projects make applications that fall in this class, including SpelSpiek, Primus, HATCI, WooDy and AAP.

[4] http://www.inl.nl/en/corpora/spoken-dutch-corpus-cgn
[5] http://www.technopolis.be/nl/index.php

STEVIN Objective 3:
The creation of networks and the consolidation of language and speech technology activities, educate new experts, and promote discussion and transfer of knowledge

In order to achieve the third goal, a series of supporting activities, using ±10% of the total STEVIN budget, have been planned and have largely already been carried out. These activities include calls for demonstration projects, calls for educational projects, calls for master classes, funding of networking activities (such as conferences, workshops, etc.), the organisation of brokerage events related to STEVIN calls, and the organization of internal STEVIN Programme Meetings in which STEVIN-funded researchers and developers can discuss their projects and results.
STEVIN has played an important role in encouraging network relations between the complete spectrum of players involved in developing, implementing and embedding LST resources and technologies. This has been achieved in a number of highly visible ways:

- By the nature of the calls for proposals, which have asked for joint projects involving industry and universities.

- Through a number of major events to which representatives from across the spectrum of researchers, developers and users have been invited. Examples of such events are the brokerage events organized in connection with the calls for proposals, "Taal in Bedrijf" ('language@work'), STEVIN programme meetings, etc. These events were generally very successful and had high attendance numbers (see the STEVIN Fact File).

- By supporting – both financially and in person – various networking events organized by research groups, companies and others. Various conferences[6] (from the scientific oriented CLIN[7] and InterSpeech2007[8] towards more applied TaalInBedrijf[9] and ICT-Delta[10]) were sponsored and/or used as a platform for networking and promotional activities. More details about the organized or sponsored networking events until 2009 (29 in total) can be found in the STEVIN Fact File.

Just from the collaboration in the projects over 330 binary cooperation link occurrences have arisen, which is a clear indication of the high degree of collaboration.
Some twelve publications on the STEVIN programme have appeared in several journals, in particular journals published by the programme bureau organisations and the journal of the NOTaS organization (DIXIT). See the STEVIN Fact File for more details.
The STEVIN programme also opened demonstration calls for proposals for companies to develop LST-related demonstration applications (with or without universities). Companies could present interesting proposals for software applications where state-of-the-art language and speech technology were used in a new situation. If selected, the Nederlandse Taalunie would commission the consortium to develop the application. The application had to be realised within 15 months. Bureaucratic overhead was reduced to a minimum, the consortium would be paid a fair hourly wage for the number of hours estimated to build the application up to a maximum of € 100.000 for the completed demonstrator.
It resulted in 14 demonstrator projects (see STEVIN Fact File). These applications cover a wide variety of LST for different groups in the Dutch and Flemish society (schools, counties, police, courtroom, children, etc.), but two themes are dominant: Not surprisingly (STEVIN is all about language), many demonstration projects developed applications or services to facilitate the creation of or access to language content: Spelspiek allows one to check the correct spelling of a word, Primus provides a spelling and grammar checker and WooDy a word predictor for dyslectic speakers; Rechtspraakherkenning makes access to court proceedings easier, Audiokrant makes written newspapers audible, NEON adds textual content to audio-visual signals (subtitling) to make them more accessible to people with bad or no sight, YourNews organizes the massive amount of news text available in accordance with the user's own profile, and HATCI helps people with cochlear implants improve their auditive perception. Second, many demonstration projects deal with applications or

[6] For a list of events supported by STEVIN networking grants see: http://taalunieversum.org/taal/technologie/stevin/netwerking/
[7] CLIN: http://lands.let.ru.nl/~clin2007/sponsors.html
[8] INTERSPEECH2007: http://www.interspeech2007.org/
[9] Taal in Bedrijf: http://taalunieversum.org/taal/technologie/taalinbedrijf/
[10] ICT-Delta: http://www.ictdelta.nu/

services in the area of education: these include Spelspiek, Primus, HATCI and AAP, while also the demonstrator Diademo made in an educational project can be counted among these.
The major goal with these demonstration projects was to increase the visibility of language and speech technology by creating attractive demonstrators and show these demonstrators to governmental organizations, companies, educational organizations, researchers, the general public, etc., thus trying to increase demand to language and speech technology.
The results of these projects have been demonstrated at various events, including InterSpeech2007 in Antwerp, TaalinBedrijf, ICT-Delta, STEVIN programme meetings and at various other events. Illustrations of the results of these projects can also be found at the STEVIN website.[11]
From the beginning it was clear that a strong effort should be exerted in the dissemination of Language and Speech Technology for school children (age 17+) and bachelor students. The last years the number of language and speech technology-students has decreased considerably (for several reasons) and it becomes more and more difficult to find excellent (Dutch speaking) Master and PhD-students. Spreading the word for schoolchildren and bachelor students may help to increase the number of students in language and speech technology related areas and, as a side effect, make youngsters aware of something like LST. Three calls for "educational" projects have been launched (in 2007, 2008 and 2009). In each year one project was awarded. In the 2007 and 2008 calls only one proposal was submitted (and awarded). In 2009 3 project proposals were submitted. However, two of them clearly did not qualify as an "educational" project, and the third one was a continuation of the project awarded in 2007. Therefore the working group for supporting activities (WGFA) and the PC have decided not to organize a new call in 2010.
"Master classes" are aimed at increasing general awareness of LST research and applications with companies and government organizations. They are short (1-3 days) but intense courses in the area of LST and its applications targeting the people who are in a position to make decisions on the use of language and speech technology in companies and in government organizations. STEVIN has launched two calls for master classes, one in 2008 and in 2009. Two proposals were submitted in the first call, one in the second call, and each calls yielded one awarded project. Both awarded master classes still have to be held, so it is not possible to evaluate these at this point in time.
In combination with (at least some) of the projects granted in the third open call, we believe all the targets of the third goal have been addressed and have to a large extent also been reached in an excellent way. Their realization has therefore contributed significantly to achieving the overall policy goals.

Target Groups
STEVIN from the start intended to cover all four layers in the layered LST innovation model: basic resources, strategic research, application-oriented research and bringing applications and services to the end-user. In addition, a good balance of funding knowledge institutes and companies was aimed at, as well as a balanced distribution of funds between the Netherlands and Flanders. The first two open R&D calls and the tenders targeted the creation of basic resources and strategic research. These calls required that the lead partner of a project consortium was a knowledge institute, but stimulated and rewarded collaboration with companies and transnational Dutch-Flemish cooperation. Application-oriented research was targeted for in R&D Call 3. Again proposals required a knowledge institute as main applicant, but an additional requirement was that a well-identified user group was defined and involved in testing the application. Bringing applications and web services to the end-users has been implemented mainly by the calls for demonstration projects, for which SMEs were required as main applicant.
Calls were distributed via the usual Programme Bureau (NWO and SenterNovem) and Dutch Language Union channels, via channels available at the funding agencies (e.g. EWI, EZ, OC&W), via good contacts with individual LST companies, the professional association NOTaS and research networks such as CLIF, by balanced participation of researchers and developers both from knowledge institutes and from companies in STEVIN groups such as the Programme Committee and the WGFA, and by announcement via other relevant channels (mailing lists, website, at relevant events, etc.). In future programmes some more effort might be put into trying to also actively involve researchers of the Dutch language.

[11] More specifically http://taalunieversum.org/taal/technologie/stevin/etalage/ and http://taalunieversum.org/taal/technologie/stevin/pers/#demo

Summary
It has been clearly demonstrated in the preceding sections that STEVIN has largely achieved the goals from the STEVIN working programme. Targets not covered have been identified: these involve mainly morphology (derivation and compounding) and multimedia resources. However, even though STEVIN has dealt with most of its target areas, this does not imply that these areas are now fully covered. Due to budgetary and timing constraints many of the areas could only be covered in part. The large number of good (i.e. in principle fundable) submitted proposals testifies to that as well. It has been clearly described how the several activities and projects contribute to the realization of the STEVIN objectives, and we submit that the projects carried out and accompanying activities taken are an excellent concrete and justified translation of the STEVIN objectives. Finally we claim to have been successful in reaching out to all target groups, both from knowledge institutes and from companies, as well as from the Netherlands as from Flanders.

Question 3: Added Value and Relevance

- **In welke mate zijn de projectoproepen/flankerende activiteiten die STEVIN heeft gerealiseerd in de beschouwde periode doorgedrongen tot het TST-veld (onderzoekers en TST-ontwikkelaars) én (specifieke) gebruikersdoelgroepen in het bijzonder. In welke mate leidden de projecten tot bruikbaar materiaal voor het veld en (maatschappelijke) gebruikersdoelgroepen?**
- **In welke mate is de technologische en wetenschappelijke stand van zaken in het TST-domein geëvolueerd door de projecten/flankerende activiteiten die STEVIN heeft gefinancierd/georganiseerd?**
- **Welke meerwaarde heeft STEVIN al dan niet voor de wetenschappelijke wereld, de industrie en de maatschappij opgeleverd? Welke zijn de aanwijsbare verbeteringen in de TST-infrastructuur als gevolg van het STEVIN-programma?**

In English, these questions can be formulated as follows:

- **To what extent have the project calls/accompanying activities that STEVIN has realized, reached the LST-field (researchers and LST developers) and to (specific) targeted user groups in particular. To what extent have the project led to usable material for the field and targeted user groups in society?**
- **To what extent has the technological and scientific state of the art in the LST-domain evolved by projects/accompanying activities that STEVIN financed/organized?**
- **What is the added value of STEVIN for the scientific world, industry and society?**
- **Which are the concretely specifiable improvements in the LST infrastructure that can be ascribed to the STEVIN programme?**

Added value and relevance of Research & Development

Impact in Academia

The Open Calls for project proposals, including the announcements of Information Events aimed to explain the purpose of the calls and the conditions for potentially successful proposals, have been published and publicized widely in the institutes for research and higher education in Flanders and the Netherlands. Especially for the groups that are aware of what is going on in the language and speech technology field all potentially interested researchers have been reached. This is testified by the coverage of the proposals that have been received by the Programme Committee. LST developers have been reached sufficiently as well as witnessed by their presence in all types of projects.
STEVIN did not elicit successful proposals from research groups in the Humanities in general that base their research on written (and spoken) documents in the Dutch language. This was not among the goals of the STEVIN programme, but one might expect reaching this group because of the focus of STEVIN on the Dutch language. But the CATCH programme and CLARIN-NL are more focused on these target groups.

Given the focus on research groups that are active in the language and speech technology field it should not come as a surprise that the resources and tools that have become available through STEVIN have so far mainly been used for research and development in the LST-field. We believe that most, if not all of these assets will also have a substantial impact in other humanities disciplines, once the awareness of the advantages offered through these resources starts growing.

Impact on technological and scientific development

STEVIN has made very substantial contributions to strengthening collaborations between Flanders and the Netherlands. Moreover, STEVIN has helped to raise awareness of the importance of LST-approaches in neighbouring disciplines, even if that awareness is still not at the level that one would want to see.
Relative to the total investment in language and speech technology both in Europe and globally, the joint Dutch-Flemish investment – though considerable – is relatively small. Consequently, the STEVIN programme only has a restricted impact on the evolution of language and speech technology in general. Especially so since STEVIN specifically targets the (often non-trivial) adaptation and

recreation of international state of the art technology and resources for Dutch.
Having said this, it must be concluded that the Dutch-Flemish HLT community has been able to retain their top position in the international HLT community which prepared them for a leading position in the European CLARIN endeavour.

Improvements of the infrastructure
Thanks to STEVIN we now have essential resources such as the Cornetto Lexical Database, the D-coi (and soon the SoNaR) written language corpus, the SPRAAK platform for developing ASR applications and many other resources specifically geared to the Dutch language available for research and development. Given the preference for strategic and fundamental research in the Dutch and Flemish Research Organizations it is evident that the same level of infrastructure could never have been built without a dedicated programme such as STEVIN. The existence of these essential resources already has and will increasingly have impact both on LST commercial development and on the presence of Dutch and Flemish research at the international top.

Another aspect where STEVIN has made a significant impact is in educating junior and senior staff members. Without STEVIN the number of researchers in the Flemish and Dutch universities who have active knowledge of and hands-on experience with language and speech technology tools and resources would most probably be smaller than it is now.

The BaTaVo Report recommended to create a Dutch Language HLT Agency, with the tasks to collect, document, maintain and redistribute LST-tools and resources. In response to this recommendation the TST-Centrale was created, as a project of the Dutch Institute for Lexicology. So far, the TST-Centrale has played an important role in collecting and redistributing resources and tools. It remains to be seen whether it is at all realistic to also charge an organization such as the TST-Centrale with tasks related to documentation and maintenance, especially with regards to software tools.

Added Value and Relevance of Accompanying Activities

To what extent have the project calls/accompanying activities that STEVIN has realized, reached the LST-field (researchers and LST developers) and to (specific) targeted user groups in particular.

Demonstration projects
In general one may say that (nearly) all of the Duct and Flemish LST-(related) companies were aware of the existing of the STEVIN-programme. This becomes clear in the high number of different companies participating in the requests for demonstration projects. In some demonstration projects there was a strong collaboration between LST-companies and universities, resulting in knowledge transfer, lowering the boundaries between companies and universities and intentions for future collaboration. Another positive effect of the demonstration projects was the increased awareness of the STEVIN-programme by the end-users. Independently of whether a proposal was granted or not, end-users were told about STEVIN, LST, and its promising possibilities for their companies.
The 14 granted demonstration projects will exist at least for some time on the websites of the NTU and the companies involved. Furthermore, especially NWO, SenterNovem and EWI have written about several of these projects in their widely distributed publications. Moreover, the demonstrations are being used as show case by the companies to convince potential customers to use LST.

Master classes
Master classes, organised by companies with support of universities clearly show the potential of LST. Demonstration of working applications that are used by real companies and/or organisations is a strong plus point. They are aimed at decision makers: people in companies or in government that can (co-)determine the policy on the use of LST technologies, so they may have very beneficial effects for increasing awareness and stimulating demand. However, the real impact of the master classes cannot be measured yet since they are still to be held.

Educational projects
Educational projects aim at raising the interest in LST of high school pupils and young students. Two of the three awarded projects produce content for Kennislink.nl, targeting high school pupils. The project itself investigated in how far it reached its targeted audience, with positive results. The third project has created a demonstrator that is available at the Flemish science museum Technopolis. Clearly, Technopolis will attract many high school pupils, but it is too early to assess the success of the Diademo demonstrator in this respect.

Networking events
The networking events organised by STEVIN generally had high attendance numbers (see STEVIN Fact File) and indeed reached the targeted public. The fact that STEVIN supported events relevant to Dutch LST was well announced and well known in the community, and STEVIN received many requests for sponsorship, which could only partially be awarded.

To what extent have the project led to usable material for the field and targeted user groups in society?
Within the accompanying activities, first of all the demonstration projects led to usable material for the field and targeted user groups. They showed the various (potential) user groups the possibilities of LST. Not only what should be possible but also what is possible! One of the biggest problems with new technology is that potential end-users nearly always ask for an already existing and working prototype they can visit. For small companies however, it is risky to make such working prototypes. The help of STEVIN broke this Catch-22 situation and resulted in 14 demonstrations of LST-applications.
The educational projects have led to usable material as well: all the articles on Kennislink on LST will remain there for quite some time and may be used by many high school pupils and students in the years to come. Furthermore, the Diademo demonstrator presence at Technopolis will keep attracting attention of Technopolis visitors and may contribute to an increased interest in LST in general but especially among high school pupils and young students..

To what extent has the technological and scientific state of the art in the LST-domain evolved by projects/accompanying activities that STEVIN financed/organized
The strong collaboration between companies and universities increased the LST-knowledge of the participating companies. Moreover, some projects were using existing technology in a new environment, thereby increasing the number of potential end-users of LST.

What is the added value of STEVIN for the scientific world, industry and society?
Still focusing on accompanying activities, the biggest added value is the increased collaboration of the various players. STEVIN resulted in a serious lowering of the walls between scientific world, industry and society. Although it may seem a one-way direction from scientific world → industry → society, this isn't true. Increased trust and awareness of LST may lead in an increased accessibility for the scientific world of the industrial world. An example is the access to data: universities often have difficulties to get real world data, data that is available at the end-users. Lowering the walls between the parties may result in an increased access to this real world material.

Which are the concretely specifiable improvements in the LST infrastructure that can be ascribed to the STEVIN PROGRAMME?
Not relevant for accompanying activities.

Question 4: Additionality

Additionaliteit (zowel input-, output-, als gedragsadditionaliteit)
- **Welke voor de start van STEVIN gesignaleerde knelpunten in het TST-domein werden dankzij het STEVIN-programma aangepakt, en met welke mate van succes? Welke niet?**

In English, this question can be formulated as follows:

Additionality (input, output and behavioural additionality)
- **Which bottle necks in the LST domain that were observed before the start of STEVIN were addressed thanks to the STEVIN programme, and what degree of success did they have? Which ones were not addressed?**

The question about input additionality is relatively easy to answer for the academic sector. If the STEVIN programme would not have been funded and carried out, we may assume there were no or only very limited other funding opportunities for projects targeting the creation of a digital infrastructure for the Dutch language in the Netherlands and Flanders. Also in Europe, funding opportunities for LST were very limited in the past years, and working specifically on the Dutch language would be impossible due to the subsidiarity principle. We can thus safely say that the resources, tools and applications developed in the STEVIN programme would not have existed, that the intensification of the network between the LST players in knowledge institutes and industry in the Netherlands and Flanders would not have taken place, and that certain other more recent developments (e.g. cooperation between the Netherlands and Flanders in CLARIN, inter alia on data and tools originally developed in STEVIN) would not have occurred.

The question about input additionality for the industrial sector is however more difficult to answer. Could a company (e.g. Tele Atlas in the projects Autonomata and AutonomataToo?) have financed the development of the resources entirely on its own, or would it have developed them if funding covered only half in stead of the full cost? In order to answer this question, one has to carry out a cost-benefit or ROI analysis of the investment. This involves estimating the total additional market created by the application of the resource in particular products, customers' willingness to pay for these enhanced products, the time horizon over which the investment can be monetized and the reaction of competitors. Each one of these estimations carries a high degree of uncertainty, the bundled outcome even more. These very difficult questions are sometimes answered by calling upon "the wisdom of crowds" (James Surowiecki), just like a company is valuated by the wisdom of the investors' "crowd" trading its shares on the stock market, or large public infrastructure projects are decided by the "crowd" of the people's representatives in parliament, or the value of any product is determined by the market crowd. This is the approach chosen by STEVIN, by considering the resources as stand alone items with a value, acting as a proxy for their real commercial value (that cannot be calculated), and determined by the (arguably too small) crowd of the Pricing Commission. But again this is not the full answer. Does the price lead to optimum sales of the resource? How large will the market be for each resource, at its current price? How long will the resource be valuable?

In any case, making reliable estimates about the economic viability of specific products is difficult to predict for any individual company as well. In addition, it must be able to provide the necessary investments. We may safely assume that the opportunities offered by STEVIN for industry to participate in R&D projects and in demonstration projects surely have taken away obstacles for industry to invest in new or improved data and technologies and in applications or services leading to new products that can be offered. It is very likely that, at least in some cases, these new products would not have existed were it not for STEVIN.

It is also clear from the number of good (i.e. in principle fundable) proposals submitted both in R&D calls and in calls for demonstration projects that still a lot of resources, research and demonstration products could not be realized in the STEVIN programme due to budgetary limitations.

Behavioural additionality is clearly present. Researchers have met and become familiar with colleagues in other institutes and the industry on both sides of the Dutch-Belgian border. Many LST

companies in the Low Countries did not originate as spin-offs of universities, so contacts between the industry and the academic world were rare before STEVIN. Some companies have been able to recruit hard to find valuable people for their STEVIN contribution. In short, network creation and knowledge transfer have been very successful in STEVIN, as can be seen from the large number of cooperation links that have been established in or as a consequence of the STEVIN programme.

Output additionality As for output additionality, several remarks can be made. First, many projects that concentrated on creating data (annotated text corpora, lexicons, speech data) additionally yielded a variety of best practices, procedures and tools for the collection, creation, annotation, inspection and validation of these data. These best practices, procedures and tools can be used for other data as well, thus contributing to a potential of increasing speed of data availability for the Dutch language. Several of these tools will actually be used in the cooperation project between the Netherlands and Flanders in CLARIN so that they also become available to humanities and social science researchers and at the same time they will become even easier to use for LST researchers and developers.

Second, many projects that focused on strategic or application-oriented research yielded new data as a secondary result. Due to the requirements in the STEVIN programme, all these data are or will become available via the Dutch HLT Agency and thus to the whole LST community.

Third, several of the demonstration projects did not only yield a demonstration but also led to actual products in the form of applications or services that are currently being employed by the companies that developed them.

Question 5: Efficiency

- **Heeft STEVIN de doelstellingen zoals neergeschreven in de overeenkomst EWI-NTU en STEVIN-meerjarenplan gehaald op een aanvaardbare/efficiënte manier?**

In English, this question can be formulated as follows:

- **Did the STEVIN programme attain the objectives laid down in the EWI-NTU agreement and in the STEVIN multi-year plan in an acceptable/efficient way?**

As already mentioned in the response to Question 2, the main objectives of the STEVIN programme were the following:

1. *The realisation of an effective digital language infrastructure for Dutch, based on the BaTaVo priorities (=BAsis TAal&spraakVOorzieningen),*
2. *The execution of strategic research in the field of language and speech technology*
3. *The creation of networks and the consolidation of language and speech technology activities, the training of new experts, and the stimulation of discussion and transfer of knowledge*

Adopting the layered model for LST innovation that was introduced in the programme proposal, the Programme Committee (PC) aimed for a gradual shift from resource development over strategic research to applied research, culminating in the demonstration of existing and new technology in novel applications. To make this happen the PC originally envisaged a bottom-up approach with open calls for proposals, with peer review by international experts and with a scientific monitoring of the selected projects by these experts. However, it soon turned out that this was not the optimal strategy for a programme like STEVIN. Therefore, the PC promptly developed a more diversified strategy, comprising:

(1) the publication of three open calls for 'scientific' proposals with a thorough evaluation of the proposals by an International Assessment panel (IAP),
(2) the publication of tenders for soliciting proposals for the realization of particular resources that were considered top priorities of the programme, to be evaluated by members of the International Assessment Panel and independent PC-members.,
(3) the publication of open calls for small-scale demonstration project proposals that would be subject to an evaluation by people with experience in stimulating innovative product development,
(4) the appointment of two portfolio holders (PHs) per project to follow-up the progress in the course of the project and to make a general assessment of the quality of the deliverables at the end of the project, taking into account circumstantial evidence,
(5) the establishment of an IPR working group including a representative of the Dutch HLT Agency (TST-Centrale), a lawyer and scientists with experience in IPR to assist in matters of IPR.

The PC believes that its continuous efforts to improve its instruments for steering and monitoring the programme were necessary in order to ensure that the main objectives of the STEVIN were realized in an effective way, and in particular, to ensure that the LST community got the chance to demonstrate in a convincing way, and within the time scope of the programme, the effectiveness of the resources and methodologies that were developed in the early stages of the programme. The PC is convinced that the present strategy can be taken as a good example for steering a follow-up programme with similar objectives. If the strategy would have been in place at the start of the STEVIN programme, some of the problems that were encountered in the initial phases of the programme (see below) would have been avoided.

By making it explicit in the calls for proposals that the PC and the IAP would favour mixed consortia involving academia and industry and implying cross-border collaboration, the STEVIN programme succeeded in creating a lot of new collaborations, in facilitating a transfer of skills, etc.

The WGFA has turned out to be a very enthusiastic body for assisting in the development and evaluation of initiatives contributing to the full realization of the third main target of the STEVIN-programme.
In sum, the PC believes that in spite of some initial problems, the STEVIN programme has achieved its objectives in an efficient way. The following sections provide more details about the encountered problems and the proposed solutions.

STEVIN objective 1:
Realisation of an effective digital language infrastructure for Dutch

Realizing the top priorities
As for the realization of language resources the STEVIN programme explicitly listed three top priorities:

1. A large corpus of written Dutch that can serve as a major resource for the development of a large number of LST tools.
2. A multi-functional electronic lexicon that covers all the words appearing in the written Dutch corpus, and that incorporates a rich collection of annotations for each entry.
3. A modular toolkit for speech recognition to support speech technologists in their search for improved speech recognition approaches.

Since the top priorities were well known in the field at the start of the programme, the PC took it for granted that an open call for proposals with a clearly stated emphasis on resource development would attract enough good proposals to launch a pilot project to prepare the realization of the first priority and to realize the two other top priorities. However, the PC had insufficiently anticipated that the maximum budget for a proposal was set too low to solicit good proposals that could fully cover the latter two priorities. As a consequence, the projects that were granted as a result of the first call were (1) a pilot project to prepare the written Dutch corpus collection (D-Coi), and (2) a number of projects which were no more than modest steps towards the realization of the envisaged multi-functional lexicon. A proposal for the speech recognition toolkit was turned down because, driven by budgetary constraints, it had settled for the delivery of a sub-optimal solution.

Based on this first experience, the PC decided to issue three tenders with specific topics, providing detailed requirements for the intended resources and setting aside a realistic budget for realizing them. This strategy turned out to be more successful: the tenders for a multi-functional lexicon, including semantics, and for a powerful speech recognition toolkit were issued in 2005 and resulted in the projects CORNETTO and SPRAAK. The tender for the corpus of written Dutch was released much later, in December 2006. This delay was partly due to problems encountered in the pilot project with respect to IPR issues and partly to the fact that the external validation of the corpus took more time than foreseen. The tender led to the selection of the SONAR project, but financial problems hampered the full-scale launching of this project and led to a split in two stages, SONAR-1 and SONAR-2, each with their own budget.

As a result of these actions, the three top priorities are bound to be realized by the end of the programme, but the original aim of making them available already during the programme has only partly been realized. The delayed delivery of the speech recognition toolkit, for instance, caused some inconveniences and delays for a number of demonstration projects and applied research projects (see below) which planned to use a toolkit that was not yet available at the start of the project. Adopting the tender strategy from the very beginning of the project and setting aside a sufficiently large budget would most probably have avoided these problems.

Realizing the other priorities
Apart from the three resources mentioned above, the STEVIN programme had listed in a fairly open way the type of resources (data and tools) that were highly ranked in the BaTaVo list and that would be considered for development with STEVIN support. The PC has used this list as a guideline for defining the calls for proposals and for assessing the received proposals. Before issuing a new call, the PC updated the list by removing resources that were already covered by formerly granted projects.

Quality of the developed resources
With respect to the effectiveness of the digital infrastructure for Dutch Language Resources, the PC has always valued the opinion of the IAP on the scientific quality of the proposals. In fact the PC always accepted the IAP selection, though sometimes made some relative changes and amended the assessment reports by adding some extra funding conditions. These changes were motivated by arguments that usually referred to the following requirements:

1. the developed language resources must be relevant and created according to good guidelines, conventions, and best practices which are also well documented,
2. the developed language resources must be maintained and made easily available to all researchers and developers in a non-discriminatory way,
3. the intellectual property rights (IPR) concerning the language resources must be adequately dealt with, so that no problems arise for researchers and developers who want to use them.

In order to guarantee the relevance of the developed resources, the project proposals had to discuss the potential of the envisaged resources in view of future research and/or application developments. In order to guarantee that the delivered corpora would be of good quality, the projects had to reserve a budget for the formal external validation of the core data, the protocols and the documentation. It was also made clear that the project deliverables would only be accepted if this validation is positive, or in the opposite case, if sufficient measures are taken to accommodate the problems that are raised in the validation report. The external validation of the quality of software resources is of course more problematic. It was therefore decided to settle for an examination of the circumstantial evidence in support of that quality by the portfolio holders. E.g., are there any publications or reports demonstrating state-of-the-art performances obtained with the resource on the task it was designed for, are there already representative users (outside the group who designed the tool) that successfully applied the software in their research or development, etc.? The PC is convinced that the majority of delivered resources can indeed be considered relevant and created according to good guidelines, conventions, and best practices. The apparent success of the actually running applied research projects, as emerging from the reports received from the portfolio holders, is often a direct consequence of the availability of the high-quality tools being developed by STEVIN projects granted in the first and second call.

Making the resources available in a non-discriminatory way
Both the language resources that existed prior to the STEVIN programme and the resources that are created in the STEVIN programme should be made available to the LST community through the Dutch HLT Agency. This is, in fact, the mission of the Dutch HLT Agency, a central portal and service centre for Dutch language and speech technology resources. Its main tasks are acquisition and IPR, management, maintenance, distribution and service. In the course of the STEVIN programme, a pricing committee has been formed to support the HLT Agency in developing a pricing policy which distinguishes between commercial and non-commercial (research) use. For commercial use prices are in accordance with the market; for non-commercial use, the resources are made freely available at handling costs. By communicating its decisions to the developers of the resources, and by awaiting feedback from these developers, the pricing committee can ensure that the final prices for commercial use are in accordance with the market.

In order to ensure the availability of software products, the original model of depositing the software at the HLT Agency was soon experienced as sub-optimal. By the time that the SPRAAK toolkit was delivered and that the applied research proposals were ready for evaluation, it was decided that an Open Source model should be preferred as a valid software dissemination model.

Taking care of the IPR
Ensuring the correct treatment of intellectual property rights (IPR) turned out to be a non-trivial issue, especially for researchers and developers who are not particularly knowledgeable in this area. The Dutch HLT Agency tried to provide support, but also for them most of the matters relating to this aspect were new. Therefore the PC decided to set up an IPR committee, with members of the HLT Board and the PC having experience in this field (including the Dutch HLT Agency representative in

the PC), and with support of a lawyer with expertise in this area. This committee has given advice to several projects (for example D-COI, DPC, SPRAAK) and it produced several other results, in particular a leaflet to convince publishers that donating text material is not hurting their normal business interests, a schematic overview of the relations between actors involved, a number of standard contract templates that project responsibles could take as a basis for dealing with IPR issues in an adequate manner. Thanks to the efforts of the IPR committee, the Dutch HLT Agency, the PC and the PB most of the IPR problems are being solved in an adequate manner, and it is fair to state that all delivered resources will become available to developers and researchers. Nevertheless, it cannot be denied that some of the projects granted in the first call experienced a lot of inconveniences (and loss of time) due to the unsatisfactory models that were handled at the beginning of the programme.

STEVIN objective 2:
The execution of research in the field of language and speech technology

The STEVIN action plan explicitly stated that **strategic research** in the field of language and speech technology would be supported. The plan also provided a number of suggestions about domains that were considered relevant.

It was mainly in the second open call (March 2005) that proposals for strategic research were solicited, and that call did indeed elicit a number of good proposals in the domains that had been identified as relevant in the action plan. The granted projects included (1) research on semantic analysis (DAESO: Detecting and Exploiting Semantic Overlap), listed as domain 1 under language technology in the action plan, (2) research on robust syntactic parsing (LASSY: large scale syntactic annotation), listed as domain 4 under language technology in the action plan, and (3) research on noise robust speech recognition (MIDAS: Missing Data Solutions), listed as domain 1 under speech technology. Although these projects have only recently been completed or are still running, there is already good evidence that they will deliver interesting scientific results, contributing significantly to the improvement of Dutch language and speech technology. By publishing the results in important journals they have also raised the visibility of the STEVIN programme on the international scene.

In line with the stratified chain model of innovation in LST, the STEVIN action plan finally emphasized the importance of applying the existing and newly acquired knowledge and tools in novel applications. For the application of proven technology, several calls for demonstration project proposals were issued (see next section). For the development of novel applications which also involve additional research, the PC issued a third open call (April 2007) with a clear emphasis on **applied research**. This call was very successful and was able to attract 15 proposals, 5 of which were rated as 'of high quality' by the IAP, and 5 of which were finally granted. Although the projects are still ongoing, the echoes coming from the portfolio holders who visited the projects are extremely positive. Some already produced one or more papers in refereed journals.

STEVIN objective 3:
The creation of networks and the consolidation of language and speech technology activities, educate new experts, and promote discussion and transfer of knowledge

In order to realize this objective ±10% of the total STEVIN budget was reserved for supporting activities, e.g. demonstration projects, educational projects, networking activities (such as conferences, workshops, etc.), brokerage events to attract industry, internal STEVIN Programme Meetings, etc.
Thanks to the STEVIN programme a number of new local and cross-border collaborations between universities and industries, between industries and between universities have been initiated or strongly reinforced. The following instruments have proven successful to accomplish this.

1. The open calls for scientific proposals gave preference to joint projects involving industry and universities, and cross-border collaboration.
2. The calls for demonstration projects stimulated companies to develop new applications with proven LST.
3. The calls for educational projects supported activities to interest more young people in studying LST and to convince decision makers of the potential of LST in their business.

4. Events were organized with the intention to attract representatives from across the spectrum of researchers, developers and users. Examples of such the brokerage events organized in connection with the calls for proposals, "Taal in Bedrijf" ('language@work'), STEVIN programme meetings, etc. These events were generally very successful and had high attendance numbers.
5. Various networking events organized by research groups, companies and others were supported – both financially and in person. Various conferences (from the scientifically oriented CLIN[12] and InterSpeech2007[13] to the more applied ICT-Delta[14]) were sponsored and/or used as a platform for networking and promotional activities. A full listing of supported events is presented in the STEVIN Fact File and on the STEVIN website

The calls for demonstration projects were successful, inter alia because the application procedure was kept lean and simple: the budget only had to specify a fair hourly wage for the number of hours estimated to build the application, up to a maximum of € 100.000 for the completed demonstrator. The PC is very satisfied with the results of the demonstration projects, not only did they prove interesting in terms of the applications that were developed, but they also generated a lot of interest for LST in the media. This was good value for money.

In its midterm evaluation report that was in general very positive, the IAP pointed out that the visibility of the STEVIN programme on the international scene should be improved by the publication of good papers in high-level peer reviewed journals. It emerges from the list of publications mentioned in Question 1 that the quantity of publications has increased significantly in comparison to the mid term evaluation, as well as the number of publications in journals and in high impact conference proceedings.

[12] CLIN: http://lands.let.ru.nl/~clin2007/sponsors.html
[13] INTERSPEECH2007: http://www.interspeech2007.org/
[14] ICT-Delta: http://www.ictdelta.nu/

Question 6: Positioning

- **In welke mate is er eventueel overlap met andere activiteiten/projecten van andere actoren?**
- **Welke rol speelt STEVIN binnen het domein van de taal- en spraaktechnologieonderzoek in het Nederlandse taalgebied en internationaal?**
- **In welke mate werd taal- en spraaktechnologie op de beleidsagenda van relevante overheden gezet.**

In English, these questions can be formulated as follows:

- **To what extent is there overlap with other activities/projects of other actors?**
- **Which role does STEVIN play within the domain of language and speech technology research in the Dutch language area and internationally?**
- **To what extent was language and speech technology put on the policy agendas of relevant governments?**

The stratified setup, the focus on LST for Dutch, and the inclusion of both strategic and application oriented research make the STEVIN programme unique. During the STEVIN programme, there have been several other projects or programmes in which LST research is addressed. First of all, the IMIX programme (*2003-2008*) funded by NWO. This programme was well underway when STEVIN started, but the two programmes ran in parallel for a couple of years. They share their focus on the Dutch language and the explicit inclusion of partners from the industry. However, the scope of the IMIX-programme was much smaller, and the focus more targeted: the goal of this programme was the development of a multimodal dialogue system. The CATCH programme (from 2005) focuses on funding information science research contributing to the accessibility of cultural heritage and on collaboration between researchers and organizations with cultural heritage collections. The programme has funded a number of projects addressing (the application of) Language and Speech processing research, but the focus is always on cultural heritage, not on LST. In 2008 the European CLARIN preparatory project started. This project aims at preparing a European framework for creating a digital infrastructure in which languages resources (tools and data) are made visible and accessible for scientific research in the humanities and social sciences. The Dutch project CLARIN-NL is part of this European framework The scope of CLARIN is on the use of language data and tools by humanities researchers, but this requires adaptations of existing LST tools to make them more user-friendly, and the focus is on a broad range of languages, not just Dutch or just European languages. Finally, there are a number of European funding programmes that allow for project proposals with a strong LST component, such as the DG INFSO, in addition to Dutch (NWO and ICT Regie) and Flemish (IWT-SBO) programmes. However, there is no programme other than STEVIN that facilitates strategic LST research focusing on the Dutch language.

The role of STEVIN in the Dutch and Flemish LST community must be considered very large. STEVIN also has a significant impact internationally. The programme has already delivered a number of resources that are publicly available and that are already in use in new projects or in applications, e.g. CORNETTO, and STEVIN projects are well-represented at international conferences such as LREC and Interspeech, which was organized in Antwerp in 2007. In addition to the very tangible results in the form of project deliverables, presentations and publications, some less tangible, but equally important, results have been achieved: through STEVIN, the community accumulated knowledge and experience in areas such as pricing, IPR, and licensing, which it will no doubt need in future projects and programmes. The IAP concluded in its mid term review that STEVIN sets an example for the research communities in other countries.

The focus on collaboration, both between Flanders and the Netherlands and between academia and the industry, has created new networks and strengthened existing relations. The programme events have contributed to this building and strengthening of networks as well, as they served as a platform for researchers in the LST domain to meet each other, in a smaller and more focused group than the yearly CLIN conference allows, in this way neatly complementing it.

For a number of years, it was difficult for companies to find suitable, well trained researchers who were willing to make the step from academia to the industry. STEVIN has not also employed and educated a large number of researchers in Language and Speech Technology but also strengthened contacts between academia and industry by funding collaborative projects and networking events.

One of the goals of the programme was to make available all results of STEVIN projects. Each project had a delivery commitment. Furthermore, the programme included a number of demonstrator projects that incorporated proven (research) technology into practical (demonstrator) applications to show the strengths and pros of LST to a wider audience. This result-oriented approach appears to have made its way into recently started research programmes: in most new (plans for) funding research programmes (in Flanders and the Netherlands) there is a focus on valorisation and use (application) of results.

This focus on applications of results was also expressed by 3 of the 4 STEVIN funders when visited by the Dutch Language Union (NTU):

- the Dutch Ministry of Economic Affairs (EZ): referred to SenterNovem, not EZ, as the executing agency for new research programmes.
- the Dutch Ministry of Education, Culture and Science (OCW): responded in a positive manner. They stressed the importance of forward looks that sketch or exemplify how LST can/would be applied in specific contexts (e.g. education).
- the office of the Flemish minister of Economy, Science and Innovation: the adjunct head of office stressed the importance of valorisation, usage and maintenance. He was pleased with the positive outcomes of the programme.
- the Dutch science agency (NWO): stressed the importance of a close link between research and applications. This will become a major point of attention for NWO in the (near) future. In so far as this link is given sufficient attention in a follow-up initiative of STEVIN, NWO will constructively collaborate in organising a follow-up initiative.

The NTU internally has given high priority to a follow-up of STEVIN, bringing the STEVIN results to three specific domains or contexts of use, namely education, health, and governmental organisations. These efforts are grouped under the label of "LST info desk": networking activities promoting the application of the STEVIN results by others.

Also in its relationship to its official institutions (the NTU Committee of Ministers, the Council for Dutch Language and Literature, and the Inter-Parliamentary Commission), the NTU puts a focus on LST for Dutch, which has globally resulted in a broad interest in LST for Dutch and the joint Flemish-Dutch collaboration in this matter.

The general secretary of the NTU has presented the STEVIN programme to the Dutch-Flemish steering committee on Economy, Science and Innovation, in which the STEVIN funding partners participate. This steering group has included LST in its series of "innovation lunch discussions" organised by the Dutch and Flemish diplomatic representations, aiming at informal discussion between civil servants of both territories on potential avenues of cooperation in the realm of economy, science and innovation. Although not leading to concrete results, it has shed light on the positions of the various administrations and hopefully injected a new impetus in the overall process.

Additionally, a report was created that lists a range of successful LST applications, inter alia to serve as inspirational source for ideas on a follow-up of STEVIN.[15] The NOTaS group has held some meetings of a "Resonansgroep", i.e. a group with players from the LST field to obtain input on priorities for LST in the Netherlands. The PC will continue to prepare for a basis for a STEVIN follow-up.

The Flemish department of Economy, Science and Innovation (EWI) has provided the Flemish researchers with 792K euro for the implementation of three demonstrator projects and a cooperation project with the Dutch researchers under the CLARIN umbrella. From the Dutch side, the CLARIN-NL

[15] A. van den Bosch & A. van Hessen(2008), 'Toepassingen van taal- en spraaktechnologie in Nederland', internal report University of Tilburg & University of Twente

consortium has reserved 1,5M euros for the cooperation project and 19 Dutch demonstrator projects. Although this CLARIN-VL-NL initiative is not an immediate follow-up on STEVIN[16], it does result in STEVIN resources and software being adapted to the CLARIN requirements. In essence, the CLARIN goal here is to offer and use STEVIN results as web services to researchers in the humanities and social sciences (not necessarily LST specialists). EWI has also contacted the Flemish IBBT (Institute for Broadband Technology), the Dutch Technologiestichting STW, and the Flemish Hercules Foundations to discuss the possibilities of funding for LST. Already in 2008, EWI has organised a LST forecast in order to have an idea of potential application contexts.

The NTU, as an intergovernmental organisation and STEVIN coordinator, will in the near future try to consolidate the outcomes of these steps already taken. Together with the PC chair it will draft a framework for a coherent and synergetic deployment of the follow-up initiatives. This will serve as the reference document for contacting policy setting individuals and organisations (including embassies) and funding organisations. In addition, this framework will be instrumental for steering the operational activities of the Dutch HLT Agency, which takes care of the distribution and maintenance of the STEVIN (and other LST for Dutch) materials.

[16] But rather of the renewed official declaration of the intention of Flanders and the Netherlands to cooperate in the domain of economy, science and innovation.

Question 7: SWOT analysis

- **Wat zijn de sterke/zwakke punten van STEVIN? Welke uitdagingen en bedreigingen dienen zich aan in de nabije toekomst, rekening houdend met het ruime beleidskader (zowel in Vlaanderen en Nederland als daarbuiten)?**
- **Wat is de evolutie t.o.v. de SWOT-analyse uit de M&I Partners-voorstudie (zie bijlage 1 punt 3)?**

In English, these questions can be formulated as follows:

- **What are STEVIN's strong/weak points? Which opportunities and threats are expected to occur in the near future, taking into account the broad policy framework (both in the Netherlands and Flanders and internationally)**
- **What is the evolution with regard to the SWOT analysis from the M&I-Partners report?**

We list strong points and weak points of the LST field in the Netherlands and Flanders, indicate how the STEVIN programme has contributed to these points, and make a comparison with the strong and weak points identified in the M&I Report before the start of the STEVIN programme.[17] We also list some opportunities and some threats for LST.

Strong Points

Scientific quality: The scientific quality of LST researchers in the Netherlands and Flanders was uncontested before the start of the STEVIN programme (M&I Report, p. 60). The STEVIN programme has contributed to consolidating this position. Though some of the STEVIN projects focused on the creation of language resources and may not have been scientifically cutting-edge projects, this has created new opportunities for increasing the scientific excellence in the near future since many new language and speech resources have been created in the STEVIN programme that allow existing research questions to be addressed in a better (and more data-intensive) way, and allows the formulation of completely new research questions.

Excellent international visibility: International visibility was excellent before the start of the STEVIN programme (M&I Report, p. 60). This international visibility has been consolidated by the STEVIN programme, witness the number of publications in journals and conference proceedings, and the many presentations of researchers from the Netherlands and Flanders in various conferences, workshops and other meetings (see STEVIN Fact File). The IAP observed a lack of visibility in their mid term evaluation report. We believe that this was due to two factors: (1) the projects from the first Calls focused on the creation of language resources for the Dutch language and did often not involve cutting-edge LST-research; (2) publications usually appear long (1-2 years) after the research has been carried out. This analysis appears to be confirmed by looking at the list of STEVIN publications in the STEVIN Fact File, and it gives reason to expect that many more publications can be expected in the near future. It was also suggested that STEVIN should increase its international visibility as a programme, since it could serve as an excellent model for other countries or regions as well. Work on this is ongoing but as yet insufficiently realized to conclude that this aspect has already been fully covered.

Intense cooperation between companies and knowledge institutes: Cooperation between companies and knowledge institutes was ongoing also before the STEVIN programme and listed as one of the strong points of the field in the M&I report (p. 60). STEVIN has strongly intensified this cooperation. Most of the R&D projects involve both knowledge institutes and companies. Companies have been strongly involved in the STEVIN programme, especially in demonstration projects, but also via events such as TaalinBedrijf. STEVIN also has made the Flanders company landscape more visible to Netherlands researchers and vice versa.

[17] We will use "M&I Report: to refer to: Jaap Akkermans et al. (2004), "Technologieverkenning Nederlandstalige Taal- en Spraaktechnologie", ministerie van EZ.

Large amount of basic language resources: STEVIN has clearly achieved one of its major aims: creating a large amount of high quality basic language resources and making them easily accessible to all LST researchers and developers, both in knowledge institutes and in companies. Most of the resources that were lacking before the STEVIN programme and that were listed in the priorities in the STEVIN work programme have been realized. STEVIN has thus been able to turn a weak point in the M&I Report ("insufficient basic language resources') into a strong point of the LST field in the Netherlands and Flanders. Nevertheless, some basic resources that were listed in the priorities are still lacking, among them morphological lexicon and analysis systems for compounding and derivation and multimedia resources. In addition, some of the resources that were created are limited in size and/or in scope. For example, the amount of (manually verified) semantically annotated text is small; annotation of discourse and rhetorical relations is not absent but has not been addressed systematically and at large scale; speech resources have been created but they cover only a part of the target user groups and use cases that may pose adverse conditions for speech recognition, etc.

Language resources are easily accessible: The relevant resources have not only been created but are (or will soon be) easily accessible via the Dutch HLT Agency. The set up of the Dutch HLT Agency, and the requirements imposed on each project to make available the data and tools produced via the Dutch HLT Agency have resulted in a situation where the results of the STEVIN programme (but also of other, earlier programmes and projects) have become systematically available to all researchers and developers. Accessibility here includes two aspects: (1) visibility of the resources (via the Dutch HLT Agency catalogue), so that researchers can easily find out what exists, and (2) Easy obtainability of the resources via downloads, on-line browsing and searching, or physical transports of data and tools.

Systematic approach to IPR issues: STEVIN has made an attempt to arrange for IPR-issues around language resources in a systematic manner. It has (via the IPR working group) created a limited number of licensing template agreements, complemented with some templates obtained from actual practice. It has created awareness for the importance of IPR issues among researchers, and it has actually arranged many license agreements with third parties (mostly via corpus creation projects such as D-COI, SoNaR and DPC). It continues to raise the interest of potential data providers in order to obtain more resources for LST research and development. Not all problems have been solved, the process of arranging IPR sometimes did not run smoothly, and the proper arrangement of IPR issues will remain a continuing effort. But it is safe say that the STEVIN programme has dramatically improved the IPR situation, which was problematic before the programme started ("The field is not able to make proper arrangements for IPR of language resources", M&I report, p. 62).

De facto **standards for Dutch-language resources:** Though before the start of the STEVIN programme some *de facto* standards for Dutch-language resources had emerged, the intense cooperation in STEVIN has significantly contributed to the establishment of additional *de facto* standards for Dutch-language data and tools. In fact, even the *de facto* standards that existed before STEVIN were largely due to an earlier programme of cooperation between multiple research and development groups in the Netherlands and Flanders. Intense cooperation between different groups, where one group takes the output of another group as a basis for its work creates a good environment for the emergence of *de facto* standards that originate from the research and development groups themselves, and have been tested in practice (rather than being imposed from above and not being in touch with reality) .

Actual use of LST for new domains: A large part of the data and tools developed in STEVIN are soon going to be used in a new domain, in particular they will become part of a research infrastructure being developed in the CLARIN project that aims to facilitate scientific research for the humanities and social sciences. Even though most of the tools require adaptations, especially to make them more user-friendly for non-technical users and to make them compliant with CLARIN requirements of metadata, semantic interoperability, web services and work flows, the core functionality of these tools and the data formats that operate on have been developed in the STEVIN programme.

Organisation and focus of the LST field: The M&I Report stated the lack of organisation of the LST field and the insufficient focus of the field on defending its interests as a weakness. The STEVIN programme itself has acted to unify the field so that it can operate as a single community and thus better represent its interests to other fields and potential funding candidates. We therefore state this

as a strong point of the STEVIN programme. However, it is not obvious that this will remain so after the STEVIN programme. The coordination by the Dutch Language Union provides some positive prospects that the community can continue to act as a single community after the STEVIN programme, but there are no guarantees for this and it will require continuous attention. The Netherlands still does not have an organisation such as CLIF in Flanders, and though many LST-players are member of NOTaS, it does not represent all players. Ensuring the operation of the LST community as a single community also after the STEVIN programme is a challenge to be addressed, but an excellent opportunity for this has been created by the STEVIN programme.

Relations between knowledge institutes and companies: The M&I report observed (and lists as a weakness) the fact that discussions between knowledge institutes and companies, though intense, were insufficiently to the point before the start of the STEVIN programme. Indeed, there were many, often very emotional discussions between these organisations on the roles of the different players, their mutual relations; there were fears of market disturbance; it was felt that what knowledge institute were investigating had no relevance to industry, and conversely, that no interesting research could be carried out on topics that industry was working on. Though the latter issues still play and have not gone away, it can safely be said that the STEVIN programme had a very beneficial effect on relations between knowledge institutes and companies. By the fact that they were actually cooperating in joint projects, and by the fact that there were also opportunities to work (largely) separately (e.g. knowledge institutes focusing on research in R&D projects; companies in demonstration projects), a better understanding of each other's problems has arisen, and most of the emotional and sometimes uninformed judgments have been taken away.

Demand Side: The M&I report observed that the demand side was insufficiently developed. Many companies could introduce LST and would benefit from it, but lack of knowledge and avoiding risks prevents this from happening. The STEVIN programme has made an attempt to remedy this situation, in particular by the demonstration projects which took away impediments for (small) investments in new applications and services, and by the master classes that aim to increase the knowledge and awareness of LST and its potential among decision makers in companies and governmental organisations. It is too early to conclude that this is a success (that will have to become clear in the years to come), but at least STEVIN has already been able to involve various publishers and other content providers (broadcasting companies), and various city councils.

Weak Points

The M&I report mentions several weak points of the LST field in the Netherlands and Flanders some of which have not really been changed:

Large fluctuations in available (financial) resources for LST: Though with the STEVIN programme a reasonable amount of financial resources were spent on LST, it is absolutely unclear whether there will be financial resources for LST after STEVIN. It is certainly no automatism that a successful execution of the STEVIN programme (if confirmed by positive external evaluation reports) will lead to additional funding for follow-up projects. The matter is complicated by regular changes in the organization of financial funding, and by the differences in organization and timing of funding programmes between the Netherlands and Flanders. The fact that research funding in the Netherlands is currently targeting either talents or grand challenges makes it even more difficult to get funding for LST. This issue is currently being investigated.

Standardisation: The M&I report observed, before the start of STEVIN programme, that public knowledge institutes participate in standardisation efforts, but insufficiently, and with insufficient feedback to other knowledge institutes and LST companies. The PC believes that this is still the case and has not really been changed by the STEVIN programme, though there surely are opportunities to ameliorate the situation soon (e.g. in the context of CLARIN, FLaReNet and META-NET). See under opportunities.

Start-ups: The M&I report observed that the LST sector is ideal for new start-ups. Before the STEVIN programme, new start-ups have been created, but less than one could expect. Though the PC does

not have a complete and systematic picture of the current situation, it believes that the situation has not really changed. In the period from 2004-2009, as far as is known to the PC, relatively little new start-ups have been created.

Integrators: The M&I report observed that companies that integrate LST in applications and services and that are not operating in a specific (language or speech-oriented) niche but deal with LST as just one of the many components for their applications/services are not easy to mobilize for LST. Though the PC did not analyze the situation in any detail, it has the impression that this has not really changed despite the STEVIN programme. Most integrators and application developers that participated in STEVIN are indeed focusing on a language- or speech-specific niche.

Threats

Financial Crisis: The financial crisis may have negative effects for the LST field, since the willingness and ability to make new investments will be smaller.

Netherlands v. Flanders: The opportunities for obtaining funding for cooperation between the Netherlands and Flanders in the area of LST is complicated by the fact that the two countries have different governments, are organized differently, and by the fact that funding programmes do not run synchronously. In fact, last year some funding was available in Flanders, but not in the Netherlands, which made it impossible to spend additional funding on pure LST research projects.

Opportunities

Organisation and focus of the LST field: The STEVIN programme has acted to unify the field so that it can operate as a single community and thus better represent its interests to other fields and potential funding candidates. Ensuring the operation of the LST community as a single community also after the STEVIN programme is a challenge to be addressed, especially since there is no guarantee and there are no immediate prospects for a follow-up project of STEVIN. It will have to be investigated whether this requires setting up new organizations in the field (e.g. a counterpart of CLIF in the Netherlands, and/or a counterpart of NOTaS in Flanders), or that the current situation can be retained. Can the Dutch Language Union play a role here, or will it be too focused on goals that fit in its mission? In any case, the situation of a unified field operating a single community created by the STEVIN programme creates opportunities to make a flying start in consolidating this situation even after the STEVIN programme and independently of new programmes (if any).

Promotion of *de facto* standards for the Dutch language: The *de facto* standards developed in the STEVIN programme for resources for the Dutch language should be promoted more internationally. This of course first of all involves aspects of standards that are not specifically tied to the Dutch language. However, even the purely Dutch-specific aspects are relevant, since these often make use of or are selected from larger language-independent sets. Opportunities to promote these standards offer themselves: first of all, CLARIN, with its aim of semantic interoperability of tools and data as one ingredient to offer adequate functionality for non-technical humanities and social sciences researchers. Second, there are opportunities to promote these and similar standards by providing input the to FLaReNet[18] project that aims to formulate recommendations for EU LST policy for the coming years. Third, opportunities are offered at the European level by a facility (called META) that is being developed by META-NET (in the T4ME project), which will include the META-SHARE common data and services repository and exchange.[19]

Usage of LST in the humanities and social sciences: Concomitant with the preceding point, LST for the Dutch language will be made available for use by humanities and social science researchers in the CLARIN project (both in the Netherlands and in Flanders)

[18] http://www.flarenet.eu
[19] http://www.meta-net.eu/

Excellent Scientific LST research: The STEVIN programme has yielded a wealth of data and tools. These were necessary for excellent research in the area of LST. With most of the relevant data and tools available already now and more to come in the very near future, there should be excellent opportunities for carrying out cutting-edge scientific research. Obviously, there will have to be funding to carry out such research and make use of the wealth of data and tools STEVIN has created.

New Applications and Services: The STEVIN programme has yielded a wealth of data and tools. These were necessary for excellent research in the area of LST. With most of the relevant data and tools available already now and more to come in the very near future, there should be excellent opportunities for creating new exciting applications and services in which LST is optimally deployed to serve customers, not only in companies and governmental organisations but also individual end-users. Obviously, this will require entrepreneurs who are willing and able to invest in such new applications and services based on a solid business plan.

Financial Crisis: The financial crisis may create new opportunities for LST. Though the willingness and ability to make new investments might be smaller, there will also be an urgency to reduce costs and increase efficiency. Language and speech technology, when properly deployed and when its limitations are taken into account properly, should allow usage which will bring a fast return on a relatively small investment.

Netherlands v. Flanders: The governments of the Netherlands and Flanders have reiterated the ministerial statement of intention on further strengthening the collaboration between the Netherlands and Flanders in economics, science, and innovation (April 17, 2008). Though the opportunities for obtaining funding for cooperation between the Netherlands and Flanders in the area of LST is complicated by the fact that the two countries have different governments, are organized differently, and by the fact that the timing is not running synchronously, this can sometimes be used to one's advantage. In particular, if funding is available in one country but not (yet) in the other, one can sometimes use the availability in one country as a lever to increase the chances for a positive decision and/or speeding up decisions in the other country.

Question 8: Monitoring

- **Op welke manier werd/wordt de performantie (zoals efficiënte en effectieve uitvoering) van STEVIN gemonitord?**
- **Was deze adequaat?**

In English, these questions can be formulated as follows:

- **In which manner was/is the performance (such as efficient and effective execution) of STEVIN monitored**
- **Was this monitoring adequate?**

Transparent, impartial and objective evaluation and monitoring of R&D projects
All calls for proposals have been widely announced via a variety of media (newsletters, NWO website, etc.). For the first and the second call, brokerage events have been organized to explain the nature of the calls and to give potential applicants the opportunity to meet with possible co-applicants. These brokerage events have been a great success, as can be seen from the number of participants in each brokerage event (160 and 165 participants respectively), and from the large number of proposals submitted in the three open STEVIN calls van R&D projects (52 full proposals, 68 if we include pre-proposals). In this way it was ensured that everyone with an interest in participating to the programme was aware of the opportunities STEVIN offers. This was the best guarantee for obtaining the best proposals and the best coverage of topics.

To make the evaluation impartial and objective, the assessment of the scientific quality of the project proposals was done by the IAP, a committee of independent and internationally highly respected speech and language technologists from all over Europe. The PC adhered to that assessment by not selecting any submitted project that was not highly ranked by the IAP. If the final PC ranking slightly differed from that of the IAP, this usually had to do with additional elements such as the responses of applicants to the assessment of the IAP, the balance between speech and language and between Dutch and Flemish involvement, the amount of collaboration implied, etc. The PC always justified the deviations of its ranking from that made by the IAP.

In analogy with the code of conduct used by the European Commission for its Framework Programme, a STEVIN Code of Conduct was formulated for the IAP and PC. All members were required to sign a declaration of conflict of interest and confidentiality and to formally indicate in which – if any – of the submitted projects they had any involvement. In doing so, the members committed themselves to strict confidentiality and impartiality concerning their tasks. If a member of the PC had a direct or indirect link with the project(s), or any other vested interest, or is in some way connected with the project(s), or has any other allegiance which impairs or threatens to impair his/her impartiality with respect to the project(s), the STEVIN Programme Bureau has ensured that this member did not participate in the review and ranking of the project(s) concerned.

To make the evaluation of project proposals transparent, the applicants were always informed about the assessments made by the IAP, and from the second call on, the applicants were given the opportunity to respond to these assessments. These responses were then taken into account by the PC in making the final ranking. Since some involvement of PC members in the ranked project proposals is unavoidable, the final decision on the funding of projects is made by the Board, which mainly consists of representatives of the Dutch and Flemish funding bodies and some LST experts.

For the monitoring of projects, the PC asked the projects for half-yearly reports and for each project, two of its members (the so-called portfolio holders, abbreviated as PHs) are asked to perform a project progress control. A full check of the scientific quality and validity would be impossible to realize, because the PHs are often insufficiently knowledgeable in the very specialized topics addressed in the projects they have to monitor. Different projects are monitored by different PH teams, and the PH's are not directly related to the projects they have to monitor.

The monitoring of progress by the PC initially did not run optimally, because of several reasons. First, the exact nature of the task of the PHs was not clearly defined in the beginning, meaning that

monitoring was done in different ways by different PHs. Second, the submissions of reports were not synchronized, but happened following a schedule that was related to the project start date. This made it difficult to centralise the monitoring of projects in particular PC meetings. So there was clearly room for some improvement in this respect.

The PC has therefore taken the following actions.
1. The PC has more clearly defined the task of the PHs, and developed a form to assist the PHs in their work.
2. The projects are required to submit their half-yearly reports on fixed dates.
3. A PC meeting with project monitoring as an agenda item will be planned approximately a month after such a report submission date, so that the PHs have enough time to carry out their task in preparation of the PC meeting.
4. Finally the PB organised site visits where PHs could discuss the progress with the STEVIN consortia members and try and find solutions to problems that had arisen.

The monitoring of IPR related issues is left to the representative of the Dutch HLT Agency in the PC. These measures clearly have led to a better and more systematic follow-up of the individual projects.

Another instrument that was set up for the monitoring of projects is the organisation of so-called STEVIN Programme Meetings. On such a workshop-like event, all projects present their activities, progress and results to representatives of the other running projects, to PC members, Board members and members of any other group within STEVIN programme organisation. So far, four of these meetings have been organized (2006, 2007, 2008, 2009), and one is planned for September 2010.

Transparent, impartial and objective evaluation and monitoring of other projects
Apart from scientific projects, the STEVIN programme also funds demonstration projects, educational projects, and master class projects.

For monitoring the demonstration projects, the educational projects and the master class projects, PHs from the WGFA have been appointed, and though it took some time before the monitoring and final evaluation of the projects took off, the procedure for this is now in place and the relevant information to be able to carry out a proper project monitoring is becoming available.

Demonstration projects are intended to stimulate the creation of new services and applications with existing speech and language technologies. Evaluation criteria focus on the potential contribution of a project to the visibility of services enabled by Language and Speech Technology. Another important evaluation criterion is the involvement of SMEs in these projects.

The proposals for demonstration projects have been evaluated and ranked by a small committee of mainly civil servants. The reasons for working with a civil servant committee were twofold: (1) avoiding conflict of interest; (2) the civil servant committee members have experience with these kinds of projects from other programmes and are especially experts in assessing the economic benefits of such projects.

The Programme Committee has had virtually no influence on the selection of application projects. There is only a so-called `sanity check` for the proposals that have been proposed for funding by the civil servant committee, but there usually was insufficient time to carry this out in a proper manner. Even though the reasons for working with a civil servant committee have been discussed with and accepted by the PC, at least some PC members have the opinion that the PC involvement (or, alternatively, involvement by the committee for accompanying actions) should have been more intense than it currently is.

With respect to the monitoring of demonstration projects, most demonstrations can be accessed via the STEVIN result display page of the STEVIN website maintained by the Dutch Language Union. Monitoring the progress of demonstrations has only recently been properly implemented by assigning this task to specified members of the WGFA.

Since `educating new experts' is one of the goals of the STEVIN accompanying activities, three calls for **education projects** have been issued with the aim of soliciting educational projects aiming at high school children (17+) and bachelor students. The formulation and evaluation of these calls are managed by the WGFA, supported by a small committee of educational experts. After the first call for educational projects, some adaptations of the evaluation procedure have been made. As stated before, three projects have been awarded (one in each call).

One project has finished and has been evaluated positively by its PH. The WGFA adopted the PH's recommendation for this project. The second project is about to finish, and the third one has just started, and they will be monitored by the WGFA PHs.

Calls for **master class projects** have also been organized by the WGFA, and the submitted proposals have been reviewed by this same group. Since the master classes still have to be held, monitoring of these projects is ongoing.

In 2008 the programme committee organised that the progress of the STEVIN programme was assessed by the International Advisory Panel that had been involved in selecting the STEVIN R&D proposals.

Management Summary of the STEVIN Midterm Review International Assessment Panel

The IAP congratulates the Dutch and Flemish HLT community with the accomplishments STEVIN has reached so far. STEVIN's well structured and stratified approach, with open calls and calls for tender each focusing on different aspects and levels of HLT needs, is well chosen and can serve as an example for other language communities. The set up as a joint Flemish-Dutch programme allows the direct co-operation of Dutch and Flemish research groups with excellent international track records. The co-operation between language and speech projects has also been growing, but is not optimal yet. Surprisingly good dynamics were observed in the collaboration between industry and academia, especially in small industrial enterprises, certainly due to the effective supporting activities initiated by STEVIN.

The IAP emphasizes the importance of the availability and maintenance of the developed resources and tools. The use of one Dutch HLT portal via the Dutch HLT Agency (TST-centrale) is highly appreciated and strongly encouraged.

The IPR issues, causing some delays in the starting phase of STEVIN, are now well served by the installed IPR Committee. However, IPR issues still remain a complex matter and they deserve continuous attention in the next phase of the programme. In this midterm review report the committee formulates some specific advices on dealing with IPR for the different products and results of STEVIN.

All scientific priorities and objectives are very well-covered by the STEVIN projects, apart from multimodal / multimedia projects and some semantic aspects, which are scarcely represented. The IAP acknowledges that it is not possible anymore to address these themes to their full extent within the current STEVIN Programme, but strongly suggests considering these as important items on the research agenda for the near future.

It is considered as very important that the STEVIN programme is known outside the Dutch language community. The first STEVIN results have become visible, also internationally at conferences like LREC and InterSpeech. Now that, in addition to building resources, also research results are becoming available, publications in top level conferences and journals should be feasible. The high quality of recent results, also presented at the Midterm Review STEVIN conference in June 2008, certainly justifies an ambitious publication target, and the IAP sees this as a central aim for STEVIN in the next few years.

The IAP compliments the Programme Committee for their considerate and effective operation concerning some complicated procedural issues that arose during the first phase of STEVIN. Serious effort has been exerted to deal with possible conflicts of interests which exist in the well networked Dutch and Flemish language and speech technology community. Where present, they were handled

appropriately and with integrity. Some good attempts are made for monitoring and evaluating scientific quality of projects inside the programme, although this remains a difficult subject. Participation in (international) evaluation exercises is encouraged in this respect but also publications in highly rated peer-reviewed journals.

In a final conclusion, the IAP assesses the first phase of the STEVIN programme as very successful, with excellent work done so far, which looks very promising for the results in the second phase. Moreover, STEVIN clearly provides an outstanding service to the Dutch language. Although still a lot of work needs to be done to allow the continued use of the Dutch language in ICT, funders and participants should be proud of their accomplishments so far.

Summary

We can conclude that the installed procedures for granting, launching and monitoring the projects and for transferring and approving the project results are adequate. However these have had to large extent to be developed and refined during the programme. The major reason for this is that the whole set-up and constellation of the STEVIN programme was new – result-oriented in stead of effort-oriented - and, especially in the early phase, there was too little time to prepare procedures, forms etc. in advance. It is clear that this has led to delays in the start-up and finalization of certain projects, especially the first round projects. However, most procedures have been elaborated and experiences gained have been incorporated, so that similar problems have been avoided for most of the projects from the other rounds.

Question 9: Governance

- **Op welke manier werd/wordt het STEVIN-programma georganiseerd?**
- **Was deze manier adequaat, efficiënt, effectief?**

In English, these questions can be formulated as follows:

- **What is the organisational structure of the STEVIN programme?**
- **Was this approach appropriate, efficient, effective?**

Organisational structure of STEVIN

The organisational structure of the programme was specifically designed to divide the programme management into complementary task packages that require specific and partly complementary expertise, to make different committees responsible for these tasks, and to setup a general procedure for the making of decisions.

The STEVIN Programme Committee (PC) was composed of Dutch and Flemish representatives of academia and industry together representing expertise in scientific research, application development and application introduction in the market. The PC was responsible for (1) developing the programme and formulating the calls for proposals that had to realise the aims of the programme, (2) proposing a selection of awardable applications to the Board of the STEVIN programme, taking into account the IAP assessment, and (3) monitoring the progress made by the awarded projects.

The International Assessment Panel (IAP) was composed of eight internationally respected language and speech technologists and it was responsible for (1) providing an assessment of the scientific quality of the submitted proposals (including their anticipated contribution to the STEVIN programme objectives) to the PC, and (2) participating in the midterm assessment of the STEVIN programme.

The Board of the STEVIN programme (Board) consisted of representatives of the funding organisations, and a few scientific persons with experience in running large research and development programmes. It was responsible for supervising the progress of the programme towards the programme goals and for ensuring the transparency and fairness of the assessment and ranking of project proposals by the PC. The Board was also responsible for supervising two other activities that are related to the STEVIN programme, i.e. the Dutch HLT Agency and the 'Makel en Schakel' activities carried out by the Dutch Language Union. To guarantee a good communication between the Board and the PC, the chair person of the PC was also invited to the Board meetings.

The Programme Bureau (PB) consisted of persons from NWO and SenterNovem, two organisations with a long experience in the practical implementation of large research projects. The PB offered logistic assistance to the IAP, the PC and the Board, and in this way it acquired the survey that was needed to coordinate all the STEVIN activities. Obviously, the PB was also an excellent common source of information for the different committees.

In addition, two advisory groups were set up:

- The WGFA coordinated the LST supporting activities in the Netherlands and Flanders. This group included representatives from non-STEVIN LST programmes and projects, alongside a subgroup of the PC.
- The IPR Working Group coordinated and optimized STEVIN IPR practices. This group was led by the Dutch Language Union, including legal experts and academic and industrial LST experts on IPR.

Since the PC always valued the project proposal assessments carried out by the IAP, since the Board almost always followed the proposals made by the PC, and since the PB was always able to carry out the recommendations made by the other committees, there have been no reasons for questioning the validity of the devised organisation. During the programme there was one case in which the Board seriously deviated from the PC recommendation, but this difference of opinion originated from the

fact that the two committees had a different opinion on how to put pressure on the Dutch sponsors of the programme to commit an amount of 628.000 Euro that had to complete the funding of the STEVIN programme (see below).

Though the work by the WGFA was very successful, this was the case despite the fact that this working group was not functioning optimally. There were several causes for this. First, the chair of the working group – a SME representative - could not spend sufficient time in actively preparing meetings. Second, the members of the working group were generally inexperienced with committee work and were not assertive enough (in part also due to lack of time). But most importantly, the support by the programme bureau was suboptimal. One reason for this was that project officers had to deal with too many projects in parallel, so that their attention span for STEVIN was limited (as was observed with all of them so far). The frequent changes of project officers (4 different project officers in 6 years) did not really contribute either (to say it mildly), especially since the transition period from one officer to the next was each time too short or even almost nonexistent so that transfer of knowledge and expertise was not or insufficiently possible.

Strength-weakness analysis

The STEVIN governance reflects the programme's goal to cover a stratified innovation system. The various committees/working groups acting within STEVIN consist of diverse representatives (academics in the field of both language and speech technology, industrial partners, IPR experts, education experts, legal advisors, etc. from Flanders and the Netherlands), addressing (a) particular actor(s) within the system. Below the main strengths and weaknesses of this approach are listed.

Strengths

- STEVIN's multilevel governance enabled greater scope and expertise within the LST field across Flanders and the Netherlands.
- STEVIN was a true Flemish-Dutch joint venture in the sense that all finances were pooled. As a result, the realisation and implementation of joint policies and procedures were facilitated considerably. This, in turn, facilitated scientific and technical collaborations within the projects.
- Impartial and objective evaluation of the projects was maximised by the clearly defined, separate tasks assigned to the committees/board, where the IAP was responsible for the scientific evaluation, the PC for monitoring the STEVIN priorities, and the board for assessing the selection procedure and making the final decision.
- The collaboration between NWO and SenterNovem, together constituting the programme office and responsible for operational matters, proved to be successful and fruitful. Both organisations have a long experience in the practical implementation of large research projects. Noteworthy is SenterNovem's contribution regarding the demonstration projects. The concept of this successful funding instrument mainly resulted from their input.
- The NTU made an important contribution to the success of the programme, especially because of its role in ensuring the availability of the STEVIN resources (via the TST-Centrale), its ties with the Flemish LST academic and industrial community and its outreach activities via the STEVIN website.

Weaknesses

- In retrospect, the role and responsibilities of the WGFA were insufficiently defined. This led to frustration of the WGFA members and overlapping efforts with, for example, the PR working group initiated by the NTU.
- The collaboration between NWO and SenterNovem ran smoothly in most respects. The division of labour was clear. Some delays were caused, however, due to the fact that the coordinating position at SenterNovem shifted to a different person a number of times. This meant that extra time was necessary in order to hand over the work, which was also done insufficiently.
- The support provided by the programme office (SenterNovem) to the WGFA was not always optimal, partly due to changes in the coordinating position mentioned above. This was mainly apparent in that WGFA members were not always sufficiently kept up to speed.
- Looking back, the role and responsibilities of the NTU should have been defined better. Over the course of the programme, the NTU put in a significant amount of time and effort into

coordinating the STEVIN programme – much more than was anticipated when the programme was set up. As a result, the STEVIN budget allocated to the NTU was insufficient.

- Although the collaboration between Flanders and the Netherlands was predominantly successful, it cannot be denied that culture differences at times complicated certain processes. These countries differ, for example, in how financial and administrative matters are dealt with, sometimes resulting in confusion.