



NLTNet

**Nordic Technology Network
in
Applied Language Technologies**

FINAL

Pre-Project

Nordic Innovation Center

Project number: 04222

Final Report on Copenhagen Workshop

June 2005

Responsible: Center for Sprogteknologi& ITEK

Authors: Hanne Fersøe assisted by the consortium.
Special thanks to Jan Zneider and Gisle Andersson

Document History

Version	Date	Status	Notes
01	May 4	First Draft	Creation of document template Report outline (Content)
02	May 12	First Draft	Content added
03	May 24	First Draft	Content added
04	May 25	First Draft	Content added, First draft sent to partners
05	June 1	Second Draft	Content added, partner comments added
06	June 7	Second draft	Content added, partner comments added
07	June 10	Second draft	Content added, logos added
08	June 13	Third draft	Corrections etc. from partners to second draft
09	July 8th	Final Version	Content added, minor corrections and additions

Abstract for Dissemination

This report describes the discussions and decisions of the workshop *Language Technology and Business Innovation in the Nordic Countries*, which took place in Copenhagen on March 14th of 2005.

The workshop was supported financially by a grant from Nordic Innovation Centre and NordForsk.

A consortium of 5 Nordic universities and 5 Nordic business associations organised the workshop in order to discuss the idea of formally creating a Nordic Network of Language Technology. The mission of such a Nordic NLTNetwork would be to promote research results in the area of language technology with the purpose of promoting market take-up and commercial exploitation in the Nordic area. The Nordic businesses experience parallel problems that can be solved by applying similar methods and tools, and by taking advantage of the synergy that can be achieved at technology level among the Scandinavian languages.

The main conclusion from the workshop was the decision by the members to join forces and continue towards the creation of a network. It was obvious that there was a good synergy and that the combination of the universities with the business associations would result in the strongest and best imaginable platform for the promotion of language technologies to the Nordic business communities.

The mission of the network will be achieved through practical activities, and it is stipulated that the researcher/business collaboration is innovative and will lead to new insights and understanding, such as e.g.:

- give the companies new knowledge and insight into the state-of-the-art of language technology
- give the companies an understanding of the business potential in these technologies
- give the researchers a clearer understanding of what the businesses know about language technologies and about ways in which businesses might exploit the technologies
- enable the researchers to focus their dissemination of research results
- ease the way to establishing strategic alliances between researchers and businesses
- improve the market take-up of language technology in the Nordic area for the benefit of Nordic synergy
- make Nordic languages more visible in Europe and more resistant towards “erosion”

The activities include:

- an electronic information portal
- events/meetings/seminars/conferences
- testimonials/case studies/reports and white papers
- courses
- demonstrators on the web
- exchange of employees in order to facilitate cross fertilization between the communities
- establishing of “expert” groups on a Nordic level

Contents

1. SUMMARY	4
1.1 BACKGROUND.....	4
1.2 THE CONSORTIUM.....	4
1.3 THE WORKSHOP IDEA AND THE EXPECTED RESULTS.....	5
1.4 MISSION AND ACTIVITIES OF A NORDIC NLTNETWORK	6
2. THE BUSINESS CASE FOR LANGUAGE TECHNOLOGY (LT)	7
2.1 THE SPOKEN LANGUAGE	7
2.2 THE WRITTEN LANGUAGE	8
3. INFORMATION AND AWARENESS.....	9
3.1 GENERAL AWARENESS ACTIVITIES.....	9
3.1.1 <i>Electronic Information Portal</i>	9
3.1.2 <i>Partner Roles in the Electronic Information Portal</i>	9
3.1.3 <i>Events</i>	9
3.1.4 <i>Partner Roles in Events</i>	10
3.2 SPECIFIC AWARENESS ACTIVITIES	10
3.2.1 <i>Meetings</i>	10
3.2.2 <i>Testimonials</i>	11
3.2.3 <i>Reports</i>	11
3.2.4 <i>Partner Roles in Specific Awareness Activities</i>	11
4. KNOWLEDGE AND TECHNOLOGY TRANSFER.....	12
4.1 OVERALL CONSIDERATIONS	12
4.2 EXAMPLES OF ACTIVITIES.....	12
4.2.1 <i>Tutorials, Courses</i>	12
4.2.2 <i>Development of Demos and Demonstrators</i>	12
4.2.3 <i>Temporary Exchange of Employees</i>	13
4.2.4 <i>Testing a Technology Transfer Model</i>	13
4.3 PARTNER ROLES IN KNOWLEDGE AND TECHNOLOGY TRANSFER.....	13
5. SUMMARY OF PARTNER ROLES	14
6. FINANCIAL OUTLINE FOR THE NETWORK ACTIVITIES	16
6.1 TYPES OF COSTS.....	16
6.2 ESTIMATES	16
7. CONCLUSIONS	17
7.1 FORMING A CONSORTIUM.....	17
7.2 SEEKING FUNDING.....	17
7.3 DEFINING MISSION AND ACTIVITIES.....	17
REFERENCES	18
APPENDIX A. WORKSHOP AGENDA	19
APPENDIX B. THE CONSORTIUM.....	20
APPENDIX C: ACREO'S BUSINESS SERVICES MODEL	22

1. Summary

This report describes the discussions and decisions of the workshop *Language Technology and Business Innovation in the Nordic Countries*, which took place in Copenhagen on March 14th of 2005. Agenda in Appendix A.

The workshop concluded that the participants agree to continue their collaboration and pursue funding to support core activities of a Nordic Language Technology Network with a perspective of three years. The mission and the activities of such a network were discussed during the workshop, and they will be described in this report.

Section 1, Summary, describes the background for the workshop, the rationale behind the composition of the consortium, and states the mission and the core activities for a future Nordic Language Technology Network. The consortium intends to seek financial support for this.

Section 2, The Business Case for Language Technology, gives a brief overview of the language technologies for written and spoken language, respectively, and some of their possible applications and their business potential.

Sections 3, Information and Awareness, and 4, Knowledge and Technology Transfer, describe the proposed and discussed core activities to be undertaken by the consortium members together in order to succeed with the overall mission.

1.1 Background

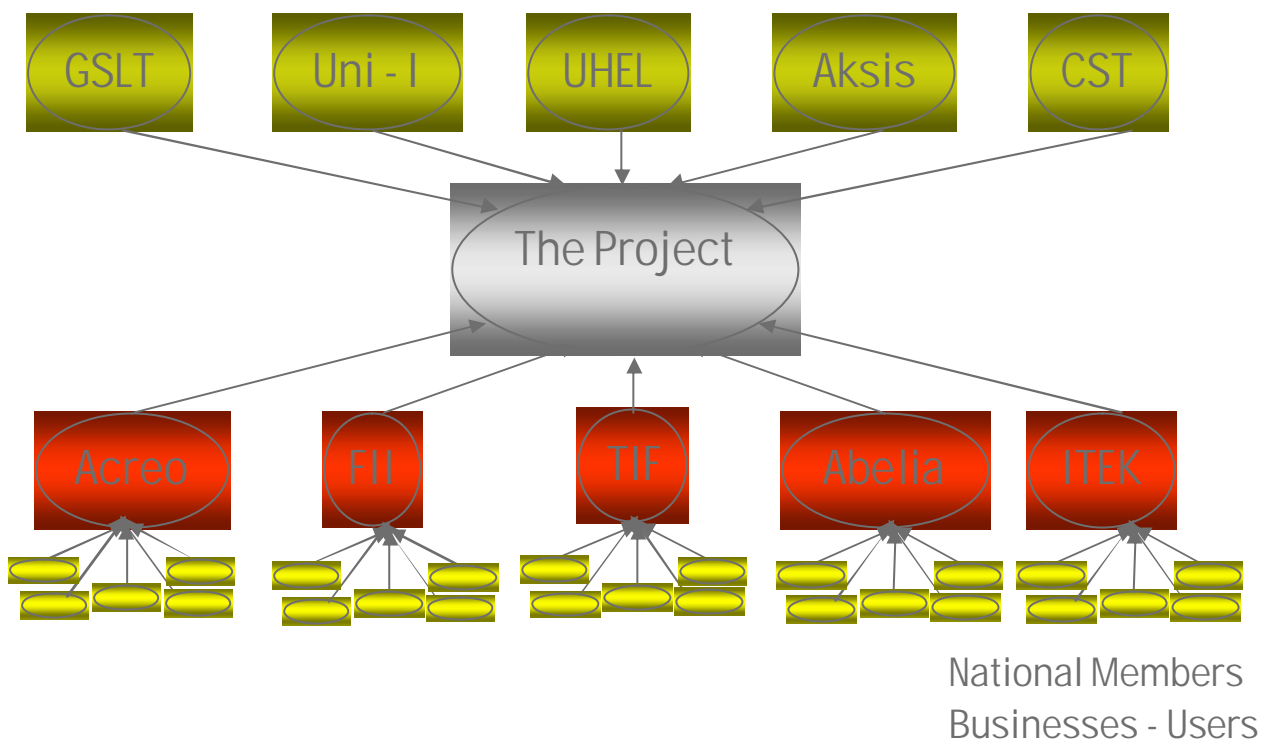
Five Nordic universities, in Denmark, Finland, Iceland, Norway and Sweden, respectively, have worked together in a network for four years with the purpose of creating a common Nordic Portal for Language Technology Information, primarily for students and academia in the field. This network was supported by NorFA. The network is called *NorDokNet*, and the portal can be found at www.nordoknet.org.

During this period of time, national business associations from the same five Nordic countries and the three Baltic countries have worked together in the *Provisions of System-on-Chip Technology for Small and Medium-sized Enterprises* project called SoC-SME. This project was supported by The Nordic Industrial Fund (now Nordic Innovation Centre). For more information see the project web page at www.soc-sme.net.

To stay competitive, businesses have to invest in human talent capable of creating innovative information technology based products. Language Technology components constitute key building blocks in future intelligent IT-systems.

1.2 The Consortium

The consortium covers all the five Nordic countries, it represents the Nordic businesses and industries through national associations, and it represents the universities where the research takes place, where prototypes are built and where the students are educated. The universities in the consortium have good networking contact with the language technology communities in the other universities in their own countries. The idea is that the associations will act as the mediators between the research communities and the businesses, driven by the common goal of innovation through technology transfer. The consortium can be illustrated like this



The Project, in the middle, represents the common goals and the agreed core activities in a future Nordic Language technology network.

On the one side, the universities provide the content to the project. Content in this context means knowledge about technologies, prototypes, projects, players, etc.

On the other side, the associations are responsible for liaison with the businesses, they are the mediators who are used to contacting the businesses and interacting with them, and as such their role is very essential.

Neither the researchers nor the individual SMEs have the knowledge and the resources available to provide direct one-to-one information about the latest developments and needs in the language technology area. In this project model, the associations enable one-to-many information dissemination because they have primary access to information and experts on the one hand, and they have, on the other hand, the tools and methods to reach their members, many at a time .

1.3 The workshop Idea and the Expected Results

The idea of the workshop was to bring these two already existing and well functioning consortia together in order to discuss the idea of formally creating a network project with a three year duration, and to discuss the core activities to be undertaken and the methods to be used in such a network collaboration.

The effort is driven by the individual goals of the two types of partners and by the fact, that collaboration will be beneficial for both groups of partners, because they can reach their own goals through it.

Universities in general are interested in raising awareness about their activities. Currently they may be even more interested in this and in illustrating their usefulness, because criteria for funding are changing. So the universities are interested in demonstrating that research results and methods exist, which represent a potential for innovation if taken up and further developed and industrialised by the businesses.

The business associations are formed in order to take care of the interests of their members. In order to succeed in this, they may work along several lines, e.g. political and technical. Currently globalisation and outsourcing are big issues in the public debate, and the challenge for businesses is to stay competitive through technological

innovation. In this light the associations are interested in supplying information to their members, which will enable these to develop their businesses through innovative new technologies, such as language technologies.

1.4 Mission and Activities of a Nordic NLTNetwork

The mission of a Nordic NLTNetwork is to promote research results in the area of language technology with the purpose of promoting market take-up and commercial exploitation in the Nordic area. It is important to build such a network at the Nordic level because the Nordic businesses experience parallel problems that can be solved by the same methods and tools, and because there is great synergy to be achieved at technology level at least among the Scandinavian languages. So this mission carries in it a strong Nordic added value both for business and for research, and at the same time it gives the Nordic area a stronger place in Europe.

The instruments to be used to succeed with such a mission were discussed at great length under two main headlines: 'Information and Awareness Activities' and 'Knowledge and Technology Transfer Activities'.

The sections 3 and 4 of this report describe in detail the proposed instruments and their expected effect.

2. The Business Case for Language Technology (LT)

Human Language Technology (HLT) is the application of knowledge of written and spoken language to the development of information, transaction and communication systems, so that they can recognise, understand, interpret, and generate human language.

Increased globalisation and recent technical and scientific advancements have increased the need for precise, cross-language communication, and the overflow of information in our society requires efficient means for man-machine interaction, storing and retrieval of relevant information from large and many different sources via devices that require interfaces with different modalities such as keyboard, voice, movements. On top of these, globalisation also increases the need for effective handling of multiple languages at the same time.

Language technologies for the spoken and for the written language play an increasingly significant role in enabling businesses to meet these needs. At the same time the researchers often experience that many people and businesses still have relatively vague, if any, ideas of what language technology is and what these different types of technologies (for written and spoken language, respectively) may do for a company, a product or a process, whether the company is a technology developer itself or whether it is a user of IT technologies.

A special challenge for the Nordic countries is the limited market size of HLT-related products for the respective languages. On the other hand, the fact that Nordic languages are closely related means that many technical challenges are similar in nature and have related solutions; this makes it natural to seek cooperation at a Nordic level. For businesses aiming to make products available in the Nordic languages, it is a strategic question whether to develop products from scratch or to adapt existing technology for larger languages.

2.1 The Spoken Language

The application potential of technologies for the *spoken language* is the easiest to immediately understand.

Speech technology enables users to communicate with a computer without the use of a keyboard or a mouse, and it is used in desktop computers, telephony services, PDAs, mobile telephones etc. Its main components are speech recognition and speech synthesis. Speech recognition converts spoken language to text, while speech synthesis takes text as input to generate speech artificially.

Enabling computers to understand and recognise naturally spoken language is a challenge and requires a complicated set of techniques. Speech recognition is based on large databases of recorded speech, representing all dialects and varieties of a particular language, and representing all possible sounds and combinations of sounds that can occur in that language. The recorded speech must be annotated with information about where a particular sound (phoneme) begins and where it ends, and which word the phoneme is a part of. On the basis of a large set of annotated recordings from a great number of speakers, a computer can be trained to understand the acoustic properties of each sound, specifically the frequency values of the sound segments in the recorded wave forms, and how the sound varies according to the environment in which it occurs. The result of the training process is an acoustic model of a language. On the basis of such a model, the speech recognition system compares incoming acoustic data with the existing model, and thereby generates a suggestion as to what words the spoken sounds represent.

Speech recognition systems are used in a number of voice-controlled services and applications, for example telephone directory services, traffic information systems, financial account information, medical dictation and dictation systems for the desktop computer.

Speech synthesis represents the other half of man-machine interaction, and is a technology which converts written language to machine-generated speech. A variety of competing technologies exist, some of which are entirely synthetic in that artificial sounds are generated by the computer, and others which are acoustic-based in that the sound is generated from an acoustic database. In the latter case, product developers record a speaker who reads a manuscript which is constructed so that it includes all the sounds in the language in different contexts. On the basis of a sophisticated system of rules and calculations, the speech synthesis system selects segments from the database, glues them together (concatenative synthesis), and modifies the speech signal and plays the generated sentence.

Speech synthesis requires a set of language-processing rules for converting the text from ordinary text to a phonetic representation. This is a challenging task due to the fact that the relationship between letters and sounds is often irregular, and that languages contain many words which look the same but sound differently. The conversion therefore relies on lexical, morphological and syntactic processing rules which convert a sentence into a representation that is annotated with information about not only which speech sounds should be generated, but also which intonation the sentence should have, where pauses should occur, the duration and pitch of the individual phonemes, and so on.

The technology is used for a variety of purposes, such as voice-controlled directory services, screen readers, talking SMS, and aids for sight-impaired, so the potential is huge. There is a considerable market take-up already in Europe for the large languages such as English, German, French, Italian and Spanish.

Despite the general turbulence in the HLT sector, acoustic models and speech recognition and synthesis systems exist for all the Nordic languages, but the Nordic countries are lagging behind in using them [1]. The actors in this field have focused on different applications with varying degree of success. A special challenge for the Nordic countries is the limited market size of HLT-related products for the respective languages. On the other hand, the fact that Nordic languages are closely related means that many technical challenges are similar and makes it natural to cooperate at the Nordic level.

2.2 The Written Language

The application potential of technologies that handle the *written language*, i.e. text and document based language technology, is not so intuitively easy to understand, on the one hand because the diversity and application areas of the technologies cover a large number of areas, and on the other hand because these language technologies are usually embedded into applications where they enable the IT-systems to manipulate language better or even intelligently. In many cases, embedding of these technologies into existing IT-applications will enable the systems, and their users, to deliver better results to well known tasks.

The basic language technologies involved in handling and manipulating text comprise but are not limited to such components as document filters, language recognisers, normalisation routines (of numbers, spelling, acronyms, signs, etc.), lemmatisers, morphological analysers, part-of-speech taggers, stop word lists, name recognisers, NP-recognisers, key word identifiers, repetitiveness checkers, language analysers (parsers), language descriptions (grammars), language generators, mono- and multilingual dictionaries, taxonomies, thesauruses, semantic networks, ontologies, text corpora, etc. Components such as these may be integrated into more complex functionalities such as spell checking, automatic classification, automatic translation, etc.

Many language technology components may either be rule based or statistically based. Rule based systems use a linguistic description (formal or algorithmic) of the problem they are intended to solve, e.g. a spell checker, or a morphological analyser, or a bi-lingual dictionary. Statistically based systems use a set of statistical probabilities for a certain solution to a certain problem. Some translation systems are statistically based and they determine the most likely translation to a certain sequence from one language to another by having processed huge amounts of parallel texts from which they have 'learned' the best solution.

In short, the application areas for language technologies for the written language are within all aspects of mono- and multilingual document handling, including information retrieval.

In the area of storing, managing and retrieving documents and knowledge the challenge is to be able to store the data in such a way that the information can be retrieved and re-used in all thinkable ways and contexts. This requires advanced technical computational solutions, and these may be enhanced by the use of language technology components, which can operate on the content of the text and not merely on the form. Most of the routines mentioned above are relevant in this context.

In the area of mono- and multilingual production of text, including translation and other globalisation processes, the challenge is extremely complex. The processes involve both keeping track of the source text and the style in which it should be written and the terminology which should be used, and subsequently or simultaneously translating it into one or more other languages in a way consistent with a multitude of requirements. This requires advanced work flow models and these may be enhanced by the use of computer systems with embedded language technology functionality such as name recognisers, dictionaries, terminology databases, repetitiveness checkers, translation tools, etc.

3. Information and Awareness

One very important instrument in the promotion of research results in the area of language technology will be a series of well designed and executed information and awareness activities. The main tool for the general awareness will be an electronic information portal offering all the relevant information supported and supplemented by national and cross Nordic events. More specific awareness may be pursued through activities such as meetings for focus or user groups, provision of video clips and testimonials, reports and white papers summing up research results for a non scientific audience.

3.1 General Awareness Activities

3.1.1 Electronic Information Portal

An Internet Information Portal will be created which collects, presents and makes accessible the knowledge and experience available about HLT research results, resources, prototypes, applications, key actors, projects, publications, etc. and which on an ongoing basis collects and presents new knowledge in the area. The Information Portal is offered to the public free of charge.

The creation of the portal will take as its starting point the existing portals for the Nordic and international levels [2]. This means that resources need not be spent on collecting information which has already been collected and which will be constantly verified until June 15th 2005. The innovative aspect of the NLTNet portal compared to the existing ones is that it will focus on also presenting the information in ways which meet the needs of businesses. The portal will offer information in all the Nordic languages and in English, and it will offer the users the facility of a cross-lingual search function such that information will be retrieved and offered not only in the search language but also in the other languages covered by the portal.

This in itself is innovative and will at the same time serve as an example of how research results can be applied on an internet portal for the benefit of the users – or customers.

In developing the portal and organising the information, the network will seek collaboration with parallel portals and networks in Europe. It is a well known fact that the major European languages are important business languages in the Nordic area, and that Nordic businesses have to face the reality of handling multiple languages in their day to day business. For this reason European collaboration in the organisation and classification of information is needed.

3.1.2 Partner Roles in the Electronic Information Portal

The universities will be responsible for collecting, organising and presenting the information and for implementing research prototypes on the portal, such as cross lingual search. The universities already have experience with presenting the information for academia and students, and in the future they will collaborate with the associations about including business relevance. They will also be responsible for collaborating with other research groups in Europe who are engaged in creating similar portals. Among the universities there will be one task leader.

The associations will be responsible for marketing the portal to their members, collecting feed-back regarding content and organisation, and providing guidelines and input to the universities regarding the specific needs of the companies to assure that the portal is seen as relevant for them.

3.1.3 Events

Events will be organised to raise the general level of awareness, and at the same time such events will also be an excellent opportunity for the players from research, provider companies and user companies to meet and exchange views. Participating with presentations in other related events, ‘piggy backing’, is also a relevant option.

Events will typically feature presentations from the research as well as from the business perspective, and previous Nordic experience from e.g. Denmark, Sprogteknologisk Forum, and European experience from the series of

LangTech conferences [3] shows that a parallel exhibition with demos of applied language technologies and products combined with user cases attracts a varied and interested audience. It is of significant interest for the audience if there is also an invited speaker of a reasonably high political or strategic level to present national or regional priorities, goals and action plans.

One model for the organisation of events over a three year period is to start with relatively small focused national events or maybe just meetings the first year, followed by larger national events the second year leading up to a large final Nordic event the third year. The supposed advantage of this model is that the associations will have the opportunity of approaching their members and marketing the idea of language technology in a step wise manner leading up to the large final event, and that project results (new applications of language technology, new companies using language technology) can get maximum exposure at such a final event. The drawback is that by starting low key it may be difficult to generate enough awareness.

So consequently, another model is to start with a 'big bang event' so to speak, in order to create as much attention as possible and launch the project. Such a large event will then be followed up by smaller and more focused follow-up events. The drawback of this approach is that since the businesses do not know the field beforehand it may be difficult to attract enough participants to such a large event.

There was no conclusive solution to this matter. Both choices have advantages and drawbacks. Basically the choice of model will depend on an assessment of the national situation of each country, and the consortium should be prepared to have member countries using the model which suits every country best.

3.1.4 Partner Roles in Events

The associations will play a key role in organising and marketing events to their members and in providing liaison between businesses and researchers.

The universities will be responsible for organising the language technology content of events based on proposals from all the partners.

3.2 Specific Awareness Activities

Apart from the broad and general awareness activities, ideas for more specific types of activities were brought forward and briefly discussed. The purpose of such activities will be to supplement the general awareness by e.g. preparing for an event, addressing special needs for information or customising information for specific user profiles (e.g. health services, government institutions, etc.). Collectively the consortium members have positive experience with various types of specifically targeted awareness.

3.2.1 Meetings

Meetings will be organised for focus or user groups about specific topics of interest for them. Another method of addressing these groups is to offer ourselves as speakers on their events. Below is a non-exhaustive list of examples of such topics and their possible target groups:

- automatic classification and indexing - for providers of e-government tools and for public administrations,
- mono- and multilingual document production - for producers and developers of industrial products,
- automatic speech recognition in e.g. telephony applications – for providers and users of customer support tools,
- automatic speech synthesis in different types of applications – for producers and users of tools for handicapped people,
- text mining (to extract knowledge from text)– for producers and users of document management tools,
- the use of linguistic knowledge in the semantic web - for producers of search engines and content providers for information portals,
- standardisation of linguistic resources, encoding and formatting – for technology developers and providers who seek to integrate linguistic resource into their application software.

3.2.2 Testimonials

From previous projects [4], both the universities and the associations have good experience with the awareness effect of different forms of testimonials, such as:

- video clips,
- articles,
- brochures,
- fact sheets

In different ways these all documents how a certain company or type of company or administration has benefited from applying a certain technology in their business area.

3.2.3 Reports

Research papers and scientific publications are not immediately accessible to the commercial world because they are written for a different type of audience. This alone makes it cumbersome for businesses to keep themselves updated on the latest trends in technology development or to acquire a full overview of the pros and cons of existing techniques and methods, etc. If there is a documented need for broad and general overview descriptions of certain phenomena, technologies, potentials etc. the network should make sure that these are provided and published on the portal in the form of reports or white papers. See an example of a white paper on the use of Taxonomies written by the Danish Company Navigo Systems A/S [5].

3.2.4 Partner Roles in Specific Awareness Activities

The associations will play a key role in proposing concrete activities based on contact with their members, and in subsequently organising and marketing the activities. See Appendix C 'ACREO's Business Service Model' for an example of a model for such a marketing effort. Collaboration with the universities is of key importance here, because they possess the knowledge about the technologies, i.e. the content.

The universities will be responsible for proposing activities and providing content for them in collaboration with the associations.

4. Knowledge and Technology Transfer

Other important instruments in the promotion of research results in the area of language technology will be a series of pragmatic and action oriented tasks that have the common goal of transferring knowledge and maybe even technology from the researchers to the businesses. While the awareness activities tell the story about language technology, the transfer activities entail doing something concrete together with language technology.

4.1 Overall considerations

The debate at the workshop was quite lively and there were many viewpoints and opinions.

One general conclusion to be drawn from the discussion is that there is a strong link between the awareness and the transfer activities. Companies learn about the existence of language technologies through the awareness activities and thus become motivated to get more in depth knowledge of the field, for instance by trying out technologies and methods in their own environment or by joining a course. The results of such concrete actions, in turn, will be promoted on the information portal and thus contribute to motivating other companies to go into transfer activities.

Another general conclusion is that those partners who already have some experience with awareness activities in the language technology area are highly motivated to initiate some more concrete transfer activities, while the newcomers to the area have reservations as to whether it is possible to motivate their members into participating in such activities.

4.2 Examples of activities

Below find the catalogue of ideas and suggestions which were discussed at the workshop.

4.2.1 Tutorials, Courses

Short courses or tutorials covering new technologies, new methods, standards, etc. will be developed and offered to the public. The courses will be taught by the experts in the universities or even in the businesses, and by invited lecturers. The possibilities of integrating the courses into the existing range of external courses offered by the universities should play an important role in the planning, and as far as possible, courses should also be offered via the Internet as distance learning.

4.2.2 Development of Demos and Demonstrators

Easy access to simple tools for standard manipulation of language based data would constitute a valuable support for business development in the area of knowledge management, and even more so if these tools were standardised so that integration and exchange of components would be simpler. In addition, it would be valuable to be able to study concrete examples of what the implementation of embedded language technology functionalities means for a certain system or type of systems. Development of such demonstrators would contribute to standardisation in the area of 'system integration with language technology', and the experience gained from developing them would contribute to the general knowledge and technology transfer through the information portal.

The following is a list of examples of demos and demonstrators which may serve the double purpose of providing a hands-on integration experience for the specific company involved in creating them, and at the same time demonstrating a specific application of one or more language technologies to others:

- Cross-lingual search functionality in a standard search engine – relevant for providers and users of search engines, content management systems and information portals,
- Improved monolingual search results with integrated linguistic resources – relevant for providers and users of search engines and document management systems,
- Improved document storing and retrieval with automatic classification and indexing – relevant for providers and users of e-government solutions and document management systems,

- Knowledge organisation with term banks and ontologies – relevant for document production, storing and retrieval
- Text to speech – having information read aloud by a synthetic voice – relevant for handicapped people, for people who need to have their hands free, in the entertainment business and in a multitude of other scenarios,
- Automatic speech recognition – relevant as an alternative to keyboard interfacing, in games, and in a multitude of other scenarios.

4.2.3 Temporary Exchange of Employees

The idea is that it should be possible for the businesses to temporarily station employees in the research environments in connection with certain projects or with the purpose of acquiring a specific knowledge. In a similar way it should also be possible for a researcher to be temporarily stationed in a business in connection with e.g. the application oriented parts of a research project. This would be the case when e.g. company specific data or processes play a central role in testing a hypothesis or a method, etc.

This particular means of technology and knowledge transfer was not viewed as having much of a chance, mostly because of the structural difficulties that can be foreseen in connection with it. It was not deleted from the list of possible activities, though, in part because it corresponds in content and intention to an initiative which has recently been launched by the Danish government called 'Regionale Videnspiloter' [6], and in part because it should be available as an option in case a company and a research group finds a way of using it.

4.2.4 Testing a Technology Transfer Model

The idea is to test an established method (ACREO's Business Service Model, see Appendix C) on a new subject area (Language Technology), with the purpose of successfully matching the researchers' competences to the businesses' needs and thus have made technology transfer happen. Another result will be the experience gained by all the parties involved from including the researchers into the process.

The first step in the model is the Need Analysis where the optimized business impact in a single company or company group will be mapped. The second step, the Match Making, regards matching the right competence provider to the identified needs. The competence provision is the third step, and it may either take place as Consultancy or as the development of a joint R&D project. A successful step three constitutes a successful technology transfer step.

4.3 Partner roles in Knowledge and Technology Transfer

The universities will be responsible for carrying out activities in close collaboration with both the associations and the companies.

The associations will be instrumental in establishing the necessary contacts between selected companies among their members and the research groups. Much of this contact making may also take place during events and other awareness activities.

5. Summary of Partner Roles

In the table below **U** partner or **A** partner stands for Universities or Associations, respectively. The workshop discussed the roles of the partners in general terms: the universities will provide content, the associations will provide contact; allocation of specific responsibilities for specific work packages was not decided, except for the general project management role.

<u>Work Package:</u> Number and Name	<u>Task:</u> Number and Name	<u>Role:</u> Responsible/Participant	<u>Short Description:</u> The Work that has to be done in the Task
WP1 Project Co-ordination and Project Management	T1.1 Project Co-ordination	R: CST and ITEK P: All other U and A partners	R: Overall project management and administration, including reporting to funding office, financial administration, organisation of project meetings, and quality control. P: Input to reports, meetings etc.
	T1.2 Day to day project management	R: Each partner	R: Local country and task specific project management including input for financial and general reporting
WP2 General Awareness Activities	T2.1 Information Portal	R: One U P: All other U and A partners	R: Technical solution for the portal and general management of its organisation and content. Provide input to reporting. P: Provide local content. The Universities must keep the LT content updated. The Associations must take contact to their businesses and map their information needs.
	T2.2 Events	R: One U with one A P: All other U and A partners	R: General planning and idea generation, identification of suitable candidates as guest speakers etc. Assess quality of event plans and provide input for reporting P: Mapping of information needs, local planning and organisation.
WP3 Specific Awareness Activities	T3.1 Meetings	R: One U with one A P: All other U and A partners	R: General planning and design, propose subjects, collect input from the other partners, assess quality of plans, and provide input for reporting. P: Need mapping, organise and hold the meetings locally.
	T3.2 Testimonials	R: One U with one A P: All other U and A partners	R: Research and propose content and production plan for 2-6 specific testimonials (video clips, case stories). Collect input from the other partners, assess quality of plans, and provide input for reporting. P: Need mapping, produce local testimonials.
	T3.3 Reports	R: One U with one A P: All other U and A partners	R: Research and propose content and production plan for 2-6 reports, white papers, fact sheets etc. Collect input from

<u>Work Package:</u> Number and Name	<u>Task:</u> Number and Name	<u>Role:</u> Responsible/Participant	<u>Short Description:</u> The Work that has to be done in the Task
			the other partners, assess quality of plans, and provide input for reporting. P: Need mapping, produce local reports.
WP4 Knowledge and Technology Transfer	T4.1 Tutorials, Courses	R: One U P: All other U and A partners	R: Research and propose content and planning for 2-4 tutorials/courses based on input from the consortium. Development of courses may be delegated to other partners. Assess the quality of courses and provide input for reporting. P: Need mapping, input provision, marketing
	T4.2 Demos and Demonstrators	R: One U P: All other U and A partners	R: Research and propose 3-4 demos based on input from the consortium. Development of demos may be delegated to other partners. Assess the quality of demos and provide input for reporting. P: Need mapping, input provision, participate in building local demos/demonstrators.
	T4.3 Exchange of Employees	R: One U with one A P: All other U and A partners	R: Research and propose 1-2 exchange scenarios and describe them using input from the consortium. Assess the quality of plans and provide input for reporting. P: Need mapping, input provision, marketing.
	T4.4 Testing a Technology Transfer Model	R: One A P: All other U and A partners	R: Research, propose and complete 1-6 test cases in collaboration with the consortium. Provide input for reporting. P: Input and content provision, make the test locally.

6. Financial Outline for the network activities

The workshop agenda comprised a brief discussion of the budget items involved to support the network and its activities. Comprehensive analysis and detailed cost estimation will be made in connection with a specific application for funding.

6.1 Types of costs

The overall picture comprises

- Network costs – network meetings and events and direct costs related to these
- Project costs – awareness and technology transfer activities and direct costs related to these

Both of these include labour costs, i.e. the manpower spent on the activities, and material costs, i.e. travelling and meeting costs, events, the purchase of external services e.g. software licenses for the information portal and video production assistance.

6.2 Estimates

The estimates in the table below are based on experience from previous Nordic Networks and EU-projects with similar activities. Ultimately the cost of an activity is determined in the lower end by the minimum of effort needed to even embark on it, and in the higher end on the quality requirements and the level of ambition defined for it. Here, creating a network, an information portal and organising a series of events has been estimated according to a reasonably high level of ambition, because these are seen as core activities. The special awareness and transfer activities are considered more complicated to embark on, all partners may not see the same need in their countries, and they are not core activities in the same sense, so the level of ambition has been set lower.

Activity	Estimate, NOK, 3 years
The Network	1.700.000
The Information Portal	3.275.000
Events	2.495.000
Special Awareness Activities	950.000
Knowledge and Technology Transfer Activities	1.520.000
External Services	50.000
Total estimated cost for 10 partners for 3 years	9.990.000

This cost corresponds to a cost per partner of approximately 999.000 NOK, or 333.000 NOK per year.

7. Conclusions

7.1 Forming a Consortium

The main conclusion from the workshop was the decision by the members of the two former consortia to join forces and continue the work as one entity. It was obvious that there was a good synergy and that therefore the combination of the two groups would result in a strong consortium. Such a consortium will be the best imaginable platform for the promotion of language technologies to the Nordic business communities.

7.2 Seeking Funding

It was also decided that this new consortium would seek funding from both Nordic Innovation Centre and from NordForsk, since the proposed activities would not likely be undertaken in this particular way without external support. Furthermore, the overall objective of the NLTNetwork supports the general Nordic strategy regarding the languages of the Nordic region, language technology holds a clear innovative potential for the businesses, and both the research communities and the associations will attract attention and be strengthened by such a collaboration.

7.3 Defining Mission and Activities

The mission of a Nordic NLTNetwork will be the promotion and dissemination of research results in the area of language technology with the purpose of encouraging market take-up and commercial exploitation in the Nordic area. This shall be achieved by carrying out a number of core activities, as described in this report. A common feature for the activities is the fact that they are well known types of activities (conferences, course etc.), which are carried out in an innovative way (collaboration between researchers and business associations), which might otherwise not have been tested, and it is stipulated that this will lead to new insights and understanding for the benefit of the Nordic Region, the Nordic businesses and the consortium members.

Seen from the perspective of the participating Associations – which are seen as the necessary link to the businesses in the individual countries – the project will:

- give the companies new knowledge and insight into the state-of-the-art of language technology
- give the companies an understanding of the business potential in these technologies
- give the researchers a clearer understanding of what the businesses know about language technologies and about ways in which businesses might exploit the technologies
- enable the researchers to focus their dissemination of research results
- ease the way to establishing strategic alliances between researchers and businesses
- improve the market take-up of language technology in the Nordic area for the benefit of Nordic synergy
- make Nordic languages more visible in Europe and more resistant towards “erosion”

The core activities decided upon to achieve these goals are described in the report. They include:

- an electronic information portal
- events
- meetings/seminars/conferences
- testimonials/case studies
- reports and white papers
- courses
- demonstrators on the web
- exchange of employees in order to facilitate cross fertilization between the communities
- establishing of “expert” groups on a Nordic level

Seen from the perspective of future development in the domain of “Embedded systems” the project should facilitate this development since the trend towards “Pervasive Computing” and miniaturization will necessitate new methods of dialogue with such systems – keyboards will simply be too large for many future systems.

References

- [1] Lockwood, Rose & A. Joscelyne (2003). *Benchmarking HLT Progress in Europe. The Euromap Study*. EU-project HOPE (IST 1999-12595), Final Report available at <http://www.enabler-network.org/reports.htm>
- [2] www.nordoknet.org and www.ltworld.org
- [3] Sprogteknologisk Forum <http://cst.dk/sprogforum2002/> and LangTech conferences <http://www.lang-tech.org/>
- [4] Euromap <http://www.hltcentral.org/htmlengine.shtml?id=56> and www.euromap.dk .
SocSME www.soc-sme.net
- [5] http://www.navigosystems.com/dansk/ontolica/pdf/WhitePaper_taksonomi.pdf
- [6] http://www.videnskabsministeriet.dk/cgi-bin/doc-show.cgi?doc_id=232250&doc_type=35&markwords=Regionale+Videnpiloter

Appendix A. Workshop Agenda

Language Technology and Business Innovation in the Nordic Countries

Workshop 14. March 2005

Venue: ITEK – Dansk Industri, H.C. Andersens Boulevard 18, 1787 København C
Meeting Room 106, 1st floor

Moderator: Jan Zneider, ITEK

- 09.00 – 09.50 **1. Opening**
Presentations – 2-3 minutes per partner
The workshop idea and the expected results – Hanne Fersøe, Center for Sprogteknologi and Jan Zneider, ITEK – 15-20 min.
- 09.50 – 11.00 **2. The Business Case for Language Technology**
The spoken language – Jens Otto Kjærsum, Prolog Development Center – 25-30 min.
The written language – Gisle Andersen, Aksis – 25-30 min.
The Nordic countries and Europe – Lise D. Jørgensen, Center for Sprogteknologi – 5-10 min.
- 11.00 – 11.15 **Break – signing of Samarbeidsavtale**
- 11.15 – 12.00 **3. Information and Awareness Activities**
Short introduction – Koenraad de Smedt, Aksis – 10-15 min.
See Annex 1 for preparation
Discussion - everybody
- 12.00 – 12.45 **4. Knowledge and Technology Transfer Activities**
Short introduction – Hanne Fersøe, CST – 10-15 min.
See Annex 1 for preparation
Discussion - everybody
- 12.45 – 13.45 **Lunch Break**
- 13.45 – 14.30 **5. Follow-up discussion on full scale project**
- 14.30 – 15.15 **6. Partner Roles and Budget Items in a full scale project**
Short introduction – Hanne Fersøe, CST, 15-20 minutes
Discussion - everybody
- 15.15 – 15.40 **7. Conclusions and action plan**
Jan Zneider, ITEK with input from everybody
- 15.40 – 16.00 **8. Budget practicalities for this workshop**
-

Appendix B. The Consortium

NLTNet Consortium Contact Data

February 10, 2005

Nordic Universities	Nordic Industry Associations
<p>Denmark</p> <p>Center for Sprogteknologi University of Copenhagen Njalsgade 80 DK - 2300 København S Contact: Hanne FERSØE</p> <p>Tel: +45 3532 9090 (Direct: +45 3432 9073) Fax: +45 3532 9089 E-mail: hanne@cst.dk Web: www.cst.dk</p>	<p>Denmark</p> <p>ITEK – Dansk Industri H.C. Andersens Boulevard 18 DK – 1787 København C Contact: Jan ZNEIDER</p> <p>Tel: +45 3377 3344 (Direct: +45 3377 3341) Fax: 45 3377 3920 E-mail: jzn@di.dk Web: http://www.di.dk/itek</p>
<p>Finland</p> <p>Department of General Linguistics University of Helsinki -UHEL P.O.Box 9 (Siltavuorenpenger 20 A, room 328) FIN-00014 Helsinki Contact: Prof. Kimmo KOSKENNIEMI</p> <p>Tel: +358 9 191-29300 Fax: +358 9 191-29307 E-mail: kimmo.koskenniemi@helsinki.fi Web: http://www.ling.helsinki.fi</p>	<p>Finland</p> <p>Technology Industries of Finland Eteläranta 10 P.O. Box 10 FIN-00131 Helsinki Contact: Marja HAMILO/Juha Ylä-Jääski</p> <p>Tel: + 358 9 192 3378 Fax: +358 9 635 855 E-mail: Marja.Hamilo@techind.fi juha.yla-jaaski@techind.fi Web: www.techind.fi</p>
<p>Iceland</p> <p>Leksikografisk Institut Íslands Háskóla Neshaga 16 IS - 107 Reykjavík Contact: Prof Eiríkur RÖGNVALDSSON</p> <p>Tel: +354 525 4403 Fax: +354 525 4242 E-mail: eirikur@hi.is Web: www.lexis.hi.is</p>	<p>Iceland</p> <p>FII – Information Technology Federation of Icelandic Industries Borgartún 35, Box 1450 IS - 121 Reykjavík Contact: Ingi Bogi BOGASON</p> <p>Tel: +354 591 0126 Fax: +354 591 0101 E-mail: ingi.bogi@si.is Web: www.si.is</p>
<p>Norway</p> <p>Aksis Universitetet i Bergen Allégaten 27 5007 Bergen Contact: Prof. Koenraad de SMEDT Section of linguistic studies</p> <p>Tel: +47 55 58 20 52 Fax: +47 55 58 94 70 E-mail: desmedt@uib.no Web: http://spraktek.aksis.uib.no/</p>	<p>Norway</p> <p>Abelia – Foreningen for IKT og Kunnskabsbedrifter I NHO Middelthunsgt 27 Postbox 5490 Majorstua N – 0305 Oslo Contact: Gro-Anett Sundstrøm OLSEN</p> <p>Tel: +47 23088077 (mob. +4799229345) Fax: +47 23088071 E-mail: gro-anett.olsen@abelia.no Web: www.abelia.no www.abelia-innovasjon.no</p>
<p>Sweden</p> <p>Institutionen för lingvistik</p>	<p>Sweden</p> <p>ACREO AB</p>

<p>Göteborgs Universitet P.O. Box 200 SE - 405 30 Göteborg Visiting address: Renströmsgatan 6, 412 55 Göteborg Contact: Prof. Robin COOPER</p> <p>Tel: +46 773 2536/5916 Fax: +46 773 4853 E-mail: cooper@gslt.hum.gu.se Web: www.cling.gu.se/</p>	<p>Electrum 236P.O. Box 5510 S-164-40 Kista Visiting address: Isafjordsgatan 22 Contact: Stellan GRANSTRÔM</p> <p>Tel: +46 8 632 77 00 Fax: +46 8 750 54 30 E-mail: stellan.granstrom@acreo.se Web: www.acreo.se</p>
---	--

Appendix C: ACREO's Business Services Model

