Catia Cucchiarini / Linde van den Bosch

Medium-sized languages and the technology challenge: the Dutch language experience in a European perspective

1. Introduction

During the last decade the language policy of the Dutch Language Union (DLU, in Dutch Nederlandse Taalunie) has focused, among other things, on strengthening the position of Dutch in language and speech technology. Our approach is not technology or research driven, but user driven, i.e. oriented towards the language community as a whole and its communicative needs. The general aim is to guarantee the full integration of Dutch in modern ICT applications, at a level of excellence comparable with that of the big languages which surround our language area (English, French and German). From a language policy perspective this HLT policy is considered as a crucial contribution to the general aim of keeping Dutch a fully fledged language that can be used in all occasions and environments.

In this paper we present our policy and discuss some of the more important initiatives in the light of what is considered to be relevant for EFNIL and its member organisations. In particular, we address the following questions:

– what can be learned from the Dutch language experience?
– how could national language resources be integrated into a multilingual technical language infrastructure?

The remainder of this paper is organized as follows. In Section 2 we briefly introduce the Dutch Language Union. We describe its geographical and topical areas of activity, the nature of this organization, its policy and some of its most important achievements in the 30 years of its existence. In Section 3 we explain the rationale behind the DLU's involvement in HLT and the approach it has chosen. In Section 4 we describe a selected number of relevant HLT initiatives that were set up by the DLU. Section 5 is devoted to the lessons we learned through more than ten years of HLT policy and to our perspectives for the future. We end with some concluding remarks in Section 6.

2. The Dutch Language Union (DLU)

The Dutch Language Union is a joint effort of the Netherlands, Belgium and Surinam to promote the Dutch language, Dutch language teaching, the literature in the language and to support the Dutch language as such (infrastructure). As a living language, Dutch is constantly evolving so as to remain suited to the demands of our times. Although this happens largely by itself, it occasionally needs a little "push" by the parties working together to keep the Dutch language a vital, modern language. To achieve this for the entire Dutch-speaking world, in 1980 the Netherlands and Belgium signed the Treaty concerning the Dutch Language Union, in which the two countries agreed to pursue a common policy on the Dutch language. Owing to the Belgian state reform (federalisation), Flanders became the official partner of the treaty. This cross-border language area treaty is the only one of its kind in the world.

2.1 The Dutch language area

The Dutch language area is principally comprised of the Netherlands, Flanders and Surinam.

- The Netherlands is virtually 100% Dutch-speaking. Its capital city is Amsterdam, and its seat of government is The Hague, where the Dutch Language Union also has its official place of establishment.
- Belgium is a multilingual country, with Dutch being spoken in the northern region (Flanders), French spoken in the south (Wallonia) and a small German-speaking area in the east. Brussels, the capital of Belgium and of Europe, is officially both French and Dutch-speaking. The Dutch Language Union also has a small office in Brussels, responsible for the activities for Dutch as a foreign language.
- Surinam is a country in northern South America. It is a former colony of the Netherlands, and uses Dutch as its language of government and education.

2.2 The Dutch Language Union: a governmental, intergovernmental and international organisation

The policy of the Dutch Language Union is established by the Committee of Ministers (Comité van Ministers), a commission comprising the Dutch and Flemish ministers for education and culture and a representative from Surinam.

The Interparliamentary Commission (Interparlementaire Commissie), comprising Dutch and Flemish representatives, oversees the policy.

The Council for Dutch Language and Literature (Raad voor de Nederlandse Taal en Letteren), comprising experts and prominent language users, advises the policymakers.

The General Secretariat (algemeen secretariaat), which prepares and implements policy, works closely with individuals and organisations from within the language region and beyond.

The Dutch Language Union is also an intergovernmental organisation: it was founded in 1980 by the Dutch and Belgian governments. Surinam joined as an associate member in 2004.

The union also cooperates with the Caribbean islands that have Dutch as an official language

2.3 2010: 30th anniversary of the Dutch Language Union

As the Dutch Language Union Treaty was signed by the Netherlands and Belgium in 1980, 2010 marks the 30th anniversary of our organization. The Dutch Language Union's motto is 'Dutch without barriers'. It signifies the Dutch Language Union's desire to help all Dutch speakers continue to be able to use their language for every purpose that a language can serve. The major areas in which the Dutch Language Union has devoted its efforts in the past 30 years are the language itself, Dutch in electronic applications, Dutch

language teaching (both teaching in Dutch and the teaching of Dutch as a second language), literature, promoting the position of Dutch in Europe and around the world and last, but not least, providing a single, uniform, official spelling for the Dutch language.

As the Dutch Language Union is a relatively small organisation, these activities have been carried out in close co-operation with other professional organisations and associations both within and outside of the Dutch language area. This is a key characteristic of the way we work. It is thanks to such cooperation that we are today able to look back on the past 30 years, and conclude that a great deal has been achieved, since the DLU was founded in 1980:

- In total there are more than 30,000 students of Dutch as a foreign language across the world, and Dutch is taught at 180 universities in 40 different countries;
- Books by Dutch and Flemish authors have been translated into 100 languages;
- Close cooperation links have been established with Surinam, Curacao, Sint-Maarten, Aruba, South Africa and Indonesia;
- Advice on a range of Dutch language and linguistic issues is freely available to the public. In 2009, 5.5 million items were consulted, and 6,300 new questions were submitted and answered;
- The DLU website, "Taalunieversum", receives over 17 million visitors a year;
- Numerous digital Dutch language resources have been made available to researchers and to the general public, as will be explained in the remainder of this paper. These resources are now being managed and maintained for future use.

3.   The DLU and Human Language Technologies (HLT)

As explained in the previous section, the initiatives by the Dutch Language Union cover all aspects of language policy. Each one is aimed at creating the right conditions to make it easier for Dutch speakers to use their language in as many different situations as possible. It is ultimately not the governments of the Netherlands, Flanders and Surinam, who are the DLU's most important 'clients', but the people who use Dutch to communicate.

In the digital world communication largely takes place through and with computers or other electronic devices. The DLU acknowledged the growing importance of Human Language Technologies (HLT), which make it possible to use natural language in information and communication technology applications, and realised it had an important role to play in guaranteeing that such technologies would become available for Dutch.

3.1   Rationale

According to the DLU embracing HLT would be the way to ensure that Dutch speakers keep using their mother tongue in all daily life situations. Taking this commitment to HLT seriously involves the development of HLT applications for a specific language, and requires the availability of a digital language infrastructure (comprising basic software tools, language and speech data, corpora and lexicons) for that language.

During the last decade the language policy of the Nederlandse Taalunie (NTU, Union for the Dutch Language) has focused on strengthening the position of Dutch in language and speech technology. Our approach is not technology or research driven, but user driven, i.e. oriented towards the language community as a whole and its communicative needs.

The general aim is to guarantee the full integration of Dutch in modern ICT applications, at a level of excellence comparable with that of the big languages which surround our language area (English, French and German). From a language policy perspective this HLT policy is considered as a crucial contribution to the general aim of keeping Dutch a fully fledged language that can be used in all occasions and environments.

As Dutch is a so-called medium-sized language and companies are not always willing or able to invest in developing HLT for a language with a relatively small market, government support was needed. On the other hand, the development of HLT is considered essential, if a language is to survive in the information society. It was against this background that the DLU set up a number of initiatives aimed at strengthening the position of Dutch in human language technologies.

3.2   Approach

The approach to stimulating language and speech technology that has been adopted for the Dutch language is comprehensive in many respects. First of all, because it is based on co-operation between government, academia and industry both in Belgium and in the Netherlands. Co-operating saves money and effort, boosts the status of the language and means not having to reinvent the wheel over and over again. Second, because it encompasses the whole range from basic resources to applications for language users. Third, because it concerns the whole cycle from resource development to resource distribution and (re)use.

4.   HLT initiatives by the Dutch Language Union

DLU activities in the field of HLT date back to 1998, when a first explorative survey was carried out. Subsequently, different initiatives were set up in cooperation with various partners in the Netherlands and Flanders. Co-operation is specially required in HLT, because the market for Dutch HLT products is relatively small so it is necessary to share the high investments, this also increases efficiency and it makes it possible to establish a common agenda. Our partners in HLT initiatives are:

– Dutch Ministry of Culture, Education and Science,
– Netherlands Organization for Scientific Research,
– Flemish Fund for Scientific Research,
– Flemish Department of Economy, Science and Innovation,
– Dutch Ministry of Economic Affairs.

In this section we do not present all HLT initiatives that have bee launched over the years. We limit ourselves to discussing three of them: the **HLT Platform Project** (1999-2003) which paved the way for two other important HLT initiatives, the **STEVIN programme** (2004-today) and the **HLT Agency** (2004-today).

4.1 The HLT Platform Project

In 1999 a decision was taken by the Dutch and Flemish governments to work closely together in matters concerning HLT for Dutch. The Nederlandse Taalunie took the initiative to install an HLT Platform, bringing together all Flemish and Dutch government bodies involved. Within the framework of the HLT Platform project a number of action lines were carried out which culminated in concrete achievements in 2003.

4.1.1 Raising awareness

A first action line was aimed at stimulating cooperation between all parties involved (Dutch and Flemish industry, academia and policy institutions), at raising awareness and at disseminating the results of HLT research so as to stimulate market take-up. Part of the activities were carried out within the framework of the European Euromap project. In 2003 it was clear that this action line had helped create transparency and structure in the HLT field in the Netherlands and Flanders, and had clearly improved communication between interested partners. A co-operative framework was now available that provided a forum for discussing, exchanging and sharing experiences, best practices, information, data and tools. This success was partly due to the participation of DLU as the National Focal Point (NFP) in the Euromap project.

4.1.2 HLT infrastructure

A second action line aimed at defining the BLARK (Basic LAnguage Resources Kit), the basic elements required to create a suitable HLT infrastructure for the Dutch language. In addition, it aimed at determining which of these essential elements were already available for Dutch and which ones were missing (Cucchiarini et al. 2001). This resulted in priority lists for language and speech technology specifying which parts of the BLARK had to be developed (Binnenpoorte et al. 2002).

A complementary study commissioned by the Dutch Ministry of Economic Affairs investigated the functioning of the HLT innovation system and its contribution to sustainable growth in the Netherlands and Flanders, to identify the optimal form of financial support for the HLT sector. The results indicated that the HLT sector had economic potential in the Dutch language area and that the optimal form of government intervention should envisage three lines of activities: realizing the prioritized HLT resources; strengthening innovation oriented strategic research in academia in response to industry needs and stimulating the demand of HLT products.

These preparatory activities laid the basis for the STEVIN programme, which was eventually launched in 2004 under the auspices of the DLU.

4.1.3 Management, maintenance and distribution of HLT resources

The action line aimed at developing a blueprint for the management, maintenance and distribution of language resources was intended to identify the necessary requirements for the re-use of digital language resources developed with government money. A Blueprint for management, maintenance and distribution of digital materials developed with

public funds was prepared by a team of language technology experts from the Institute for Dutch Lexicology (INL) and speech technology experts from other institutes, under supervision of the DLU (Binnenpoorte et al. 2002; Beeken/Van der Kamp 2004). This document also made recommendations for organising a structural form of co-operation in this respect, which eventually materialized in the HLT Agency, a central repository for HLT resources which was set up and financed by the DLU and hosted by the Institute for Dutch Lexicology in Leyden, the Netherlands (with an auxiliary branch in Antwerp, Belgium).
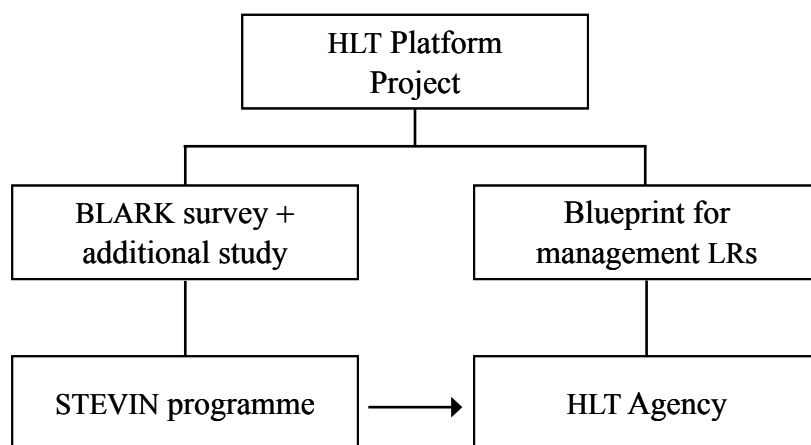
Figure 1: The relationship between the three HLT initiatives

4.2   The STEVIN programme

The respective Dutch and Flemish policy institutions acknowledged the recommendations that resulted from the various action lines, and budgets were assigned for a comprehensive HLT Programme, STEVIN. STEVIN is a Dutch acronym that stands for 'Essential Speech and Language Technology Resources'. Simon Stevin was a 16th century applied scientist who worked both in Flanders and the Netherlands and who, among other things, introduced Dutch terms for mathematical and physical concepts. In line with the priorities identified in the preparatory phase, the STEVIN programme aimed at realizing the prioritized HLT resources, strengthening innovation oriented strategic research in academia in response to industry needs and stimulating the demand of HLT products. In addition, it will strengthen the economic and cultural position of the Dutch language in the modern ICT-based society

4.2.1   STEVIN programme management structure

The STEVIN programme is jointly financed by the Flemish (Department of Economy, Science and Innovation) and Dutch governments (Ministry of Education, Culture and Science, Ministry of Economic Affairs and the Netherlands Organisation for Scientific Research). STEVIN runs until 2011 with a total budget of 11.4 million euros. STEVIN is coordinated by the Dutch Language Union and supervised by a board of representatives of the funding bodies (STEVIN Board). A Programme Committee, including both academic and industrial representatives, is responsible for scientific and content related issues (D'Halleweyn et al. 2006). An International Advisory Panel of highly-respected

HLT-experts evaluates the submitted R&D proposals. A Programme Office, a joint collaboration of the Netherlands Organisation for Scientific Research and the Dutch innovation agency AgentschapNL, takes care of the operational matters (see Figure 2).

Academic institutions and companies submit proposals that are first assessed and ranked independently by the International Advisory Panel and then by the Programme Committee. Evaluation criteria are quality, innovative features and economic aspects of the project proposal, contribution to the STEVIN Programme, proper treatment of IPR, use of standards, prevention of duplication. Based on the Programme Committee's recommendations, the STEVIN Board finally formulates a binding advice to the Dutch Language Union as to which projects should be funded.
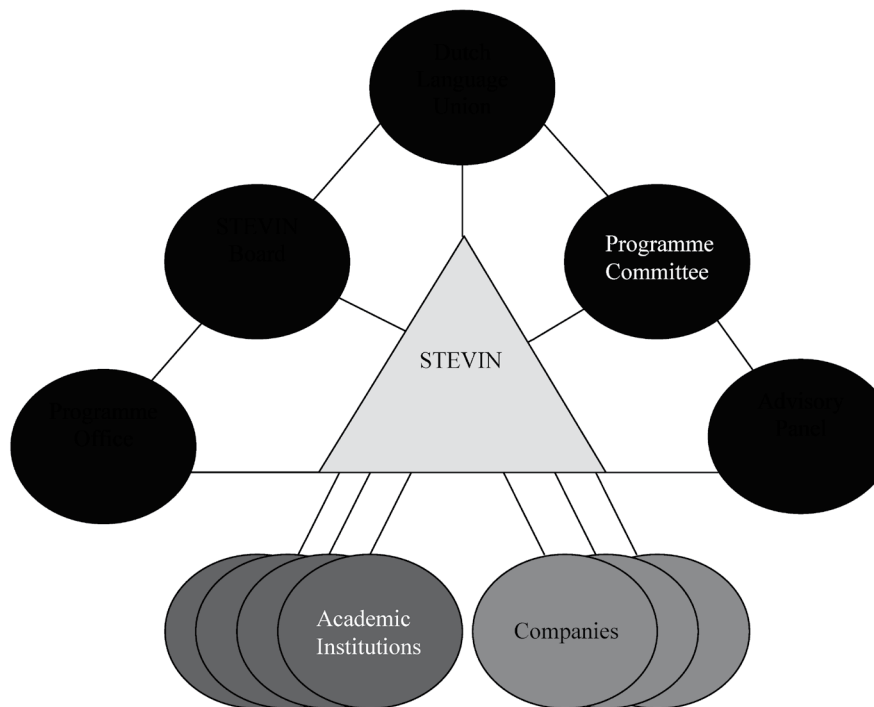
Figure 2: Management structure of the STEVIN Programme

4.2.2 Strategic research, resource development, and application development

Various calls for strategic research, resource development, and application development were launched in STEVIN. Cross-border consortiums were stimulated by increasing the standard bench fee by 50%. In total, 19 projects have been funded addressing basic HLT resources development, strategic research and application-oriented research:

1. Automata for deriving phonemic transcriptions of Dutch and Flemish names (AUTONOMATA);
2. Coreference Resolution for Extracting Answers (COREA);
3. Dutch Language Corpus Initiative (D-Coi);
4. Identification and Representation of Multi-word Expressions (IRME);
5. Extension of CGN with speech of children, non-natives, elderly and human-machine interaction (JASMIN-CGN);
6. Detecting and Exploiting Semantic Overlap (DaESO);

7. Dutch Parallel Corpus (DPC);
8. Large Scale Syntactic Annotation of written Dutch (Lassy);
9. Missing Data Solutions (Midas);
10. Northern and Southern Dutch Benchmark Evaluation of Speech recognition Technology (NBest);
11. STEVIN can PRAAT;
12. Speech Processing, Recognition & Automatic Annotation Kit (SPRAAK);
13. Combinatorial and Relational Network as Toolkit for Dutch Language Technology (Cornetto), a lexical resource for the semantic processing of Dutch;
14. Autonomata, Transfer of Output (Autonomata TOO);
15. Dutch lAnguage Investigation of Summarization technologY (Daisy);
16. Development and Integration of Speech technology into COurseware for language learning (DISCO);
17. Dutch Online Media Analysis (DuOMAn);
18. Parse and Corpus based Machine Translation (PaCo-MT);
19. Stevin Nederlandstalig Referentiecorpus (SoNaR), an annotated written Dutch corpus.

4.2.3   Raising awareness and stimulating the demand of HLT products

An important priority of STEVIN is to improve the visibility of the HLT sector, promoting cooperation, information exchange, and dissemination of research results. The former partners of the HLT Platform now meet in the board of the STEVIN programme. Within STEVIN, a substantial budget is allocated for 'accompanying measures', e.g. conferences, demonstration projects and network subsidies. The already existing cooperation is maintained and intensified through instruments that have proven to be successful, such as the STEVIN website, newsletters, conferences, workshops and seminars.

To narrow the gap between technology and the market and to address the end user conferences have been organized – a Dutch-Flemish counterpart of LangTech (2002 in Berlin, 2003 in Paris) – in the Taal in Bedrijf (intentionally ambiguous between "Language in Business" and "Language in Action") series, to bring together HLT and related-fields companies, as well as current and potential users of speech and language technologies. To attract as many different potential professional users of HLT as possible, business cases from sectors such as media, education, health care, transportation and logistics, tourism and recreation, public administration, telecom, and finance were presented.

To set an example of successful HLT applications and thus stimulate the demand of Dutch HLT applications, so-called **demonstration projects** were funded within STEVIN.

Essential characteristics of such projects are that they make use of "proven technology", that they try to access new markets for already established products or that they port established technologies to new domains. In total, 14 demonstration projects have been funded, which vary from a spoken dialogue system to optimize information provision to citizens and a speech-driven license plate retrieval tool for the police, to an auditory training system for children wearing cochlear implants.

To enhance the visibility of HLT among students and to attract them to HLT, three **educational projects** were funded aimed at making students between the age of 15 and 20 aware of the possibilities of language and speech technologies. In addition, **masterclasses** were organised on the following two topics: a) HLT and Dislexia and b) HLT for government bodies and public services.

4.2.4 Evaluations of the STEVIN programme

To be able to measure the impact of the STEVIN programme on the HLT sector, a benchmark study was carried out at the onset of the programme to establish a baseline for evaluation. In 2008, a scientific midterm programme review was carried out by the STEVIN International Assessment Panel (Spyns/D'Halleweyn 2010).

In 2010 the STEVIN Programme Committee conducted an internal evaluation of the programme, while a final evaluation of STEVIN was carried out by an external agency, the Technopolis Group (for further details, see www.stevin-tst.org/programma/#evaluaties).

The evaluations indicated that STEVIN largely achieved its goals, but this does not imply that all target areas are now fully covered. Owing to budgetary and timing constraints not all areas could be addressed, but the various projects and activities are an excellent concrete and justified translation of the STEVIN objectives. The programme has also managed to address the various target groups: companies and knowledge institutes in the Netherlands and Flanders. Other findings will be discussed in the section on Lessons Learned further on in this paper.

4.3 The HLT Agency

4.3.1 A central repository for digital Dutch language resources

The HLT Agency is a central repository for digital language resources based within an existing language planning institute of major importance, i.e. the Instituut voor Nederlandse Lexicologie (INL, Institute for Dutch Lexicology).

The resources that are developed within the STEVIN programme are subsequently handed over to the HLT Agency which takes care of their future lifecycle. This is a completely different situation from the one existing before the HLT Agency was established. At that time it was not uncommon that official bodies such as ministries and research organisations financed the development of LRs and no longer felt responsible for what should happen to those materials once the projects were completed. However, materials that are not maintained quickly lose value. Moreover, unclear intellectual property right (IPR) arrangements can create difficulties for exploitation.

To ensure that HLT resources developed with public funding become available for interested users (academia and companies) the Nederlandse Taalunie, as the owner of a number of these resources, took the initiative to set up the HLT Agency. The aim was to combine the infrastructures required for different projects, thus reducing the costs for equipment, data, software, licences, experts, and personnel, and at the same time to ensure optimal visibility and accessibility by offering resources through a one-stop-shop supplier.

In addition, to prevent HLT resources developed with public funding from lying unused on the shelf, it is necessary to make sure that they stay usable, which may entail debugging or updating to new platforms. All these activities concerning management, maintenance and distribution are carried out by the HLT Agency, which is hosted by the Institute for Dutch Lexicology. Further information on the activities of the HLT Agency, has been provided in previous publications (Beeken/van der Kamp 2004; Boekestein et. al 2006; Van Veenendaal et al. 2010).

4.3.2 IPR Issues

To enable the use and re-use of the results produced by STEVIN projects, a particular IPR-arrangement has been set up. The materials (software, data etc.) must be handed over to the Dutch Language Union so they can be made available to third parties through the Dutch HLT Agency ('TST Centrale', www.tst.inl.nl). The Dutch HLT Agency helps resolve IPR issues, is responsible for the management, maintenance and distribution of materials, and also acts as a servicedesk. For this reason, the HLT Agency is involved in the evaluation and negotiation procedures concerning the STEVIN projects. Through these IPR arrangements it can be ensured that all developed resources will become available for the whole language community in the Netherlands and Flanders.

4.3.3 Evaluations of the HLT Agency

Although it is clear that the existence of the HTL Agency has considerable advantages, we are interested to know whether and how the services offered by the HTL Agency could be improved. To this end evaluations are regularly carried out.

In 2007 a three-fold evaluation was carried out consisting of a self evaluation by the HLT Agency, a digital user inquiry by the Dutch Language Union and interviews with a selected group of users, project partners and suppliers held by an external evaluation committee. The main results of the evaluation were incorporated by the HLT Agency in a plan for improvement. The main focus was on increasing the visibility of the Agency in the field and improving collaboration and communication with suppliers and project partners.

In 2010 a similar evaluation was carried out which again indicated that a central repository for the maintenance and distribution of available Dutch language resources is highly appreciated, that IPR issues deserve continuous attention in the future, and that marketing strategies should be developed to stimulate the re-use of available resources.

These evaluations are also important to keep partners and users involved and to stay informed about the needs of the field. Other findings will be discussed in the section on Lessons Learned further on in this paper.

4.4 The European Dimension

While it should have become clear by now that the DLU stands for the Dutch language, it is also important to point out that it does not operate in a vacuum, but rather in a complex, multilingual context. As a consequence, the DLU is also concerned with the relationship between the Dutch language and other languages and/or language varieties. This

is particularly important in the field of HLT and the DLU has been involved in European HLT initiatives from the early stages of its commitment to HLT. For this reason, the DLU has participated in projects such as Euromap and ENABLER and is now involved in FLaReNeT, CLARIN-EU, the CLARIN-ERIC and META-NET. In addition, since 2000 the DLU has been presenting its HLT policies and results at the various LREC conferences (Cucchiarini et al. 2000; Cucchiarini/D'Halleweyn 2002, 2004; D'Halleweyn et al. 2006; Spyns et al. 2008; Spyns/D'Halleweyn 2010; Van Veenendaal et al. 2010).

5.   DLU HLT Policy: Lessons Learned and Future Perspectives

In the Introduction we mentioned two important questions we would address in this paper:

(a) What can be learned from the Dutch language experience?

(b) How could national language resources be integrated into a multilingual technical language infrastructure?

In this section we attempt to answer these questions.

5.1   Lessons Learned

The general impression is that the HLT approach that has been adopted in the Dutch language area has been successful in many respects. The Dutch-Flemish cross-border cooperation within the STEVIN Programme and the HLT Agency has significantly contributed to building a full-fledged Dutch language technology infrastructure. The comprehensive approach that combines stimulating research and development, creating a landing site for the results, and raising awareness of these results among prospective users has turned out to be fruitful and effective. Also the co-operation between government, academia and industry appeared to be successful and the recommendation is that it should be intensified in future initiatives.

The implementation of the STEVIN programme with a Board, a Programme Committee, an International Advisory Panel and a Programme Office might have seemed somewhat heavy at first sight, but in fact it has provided a sound structure for dividing responsibilities and making conscientious decisions.

The availability and maintenance of the developed resources through one central repository, the HLT Agency, is highly appreciated and strongly encouraged in the future.

IPR issues have properly been taken care of, but deserve continuous attention in the future. In this respect we would like to stress that the importance of IPR issues cannot be overestimated. They play a crucial role and should be properly covered through the whole lifecycle of languages resources, from their creation to their distribution. In our initiatives we noticed that sometimes it can be very difficult to explain this to providers of language material or potential users. For instance, the rationale behind the condition that materials developed in STEVIN projects (software, data etc.) should be handed over to the Dutch Language Union was purely to ensure that these materials would become available to the whole language community. However, this was often misinterpreted as

being a specific requirement of the DLU for its own benefit, which of course was not. In addition, when dealing with IPR issues, it is necessary to take account of requirements and desiderata from various stakeholders in the language community, for instance also those of industry.

Finally, another lesson we have learned is the importance of the European dimension when dealing with HLT. Operating in a European context is important to meet with important HLT stakeholders in the various European countries to exchange ideas and best practices, coordinate efforts, reach consensus, and involve the community at large.

In addition, through European cooperation it is possible to create shared multilingual repositories of language data, metadata and tools by linking and aligning already existing resources in various languages, thus promoting the re-use of resources and improving sustainability and portability of language materials and technologies. This form of exchange can facilitate not only the production of multilingual language resources, but also the development of advanced technologies and innovative services in multiple languages.

European projects like Euromap, ENABLER, FLaReNeT, and more specifically CLARIN-EU and META-NET are good examples of how this can be achieved.

CLARIN (www.clarin.eu/external/) aims at facilitating research in the Social Sciences and Humanities by creating an innovative research infrastructure for e-Research and e-Science which employs language and speech technologies and data that are accessible and interoperable. META-NET (www.meta-net.eu/) with its initiative META-SHARE, aims at establishing a sustainable network of repositories of language resources, tools, web services with the corresponding metadata which can be made available for uniform search and access to eventually create a language technology marketplace for HLT researchers and developers, language professionals, and commercial players (see contribution by Uszkoreit, this volume).

For these reasons, the Dutch Language Union intends to take part in the CLARIN ERIC and will continue participating in and contributing to interesting European HLT initiatives like those mentioned above, either directly or indirectly through the HLT Agency.

5.2   Future Perspectives

A number of field studies were carried out by the Dutch Language Union to investigate new sectors in which HLT could play an important role. The surveys looked into the specific needs of different sectors and target groups particularly related to DLU policy such as Dutch language education, public administration and people with communicative disabilities.

A first study, Human language technologies and communicative disabilities, was carried out in 2005 (Rietveld/Stolte 2005). It was aimed at identifying the specific HLT-based tools that language users with communicative disabilities require to improve their communication capabilities, i.e. tools that assist verbal dialogue, reading and writing, and communication with machines. The long-term goal is to try to improve the position of these specific groups of users of Dutch. The study showed a world of very diverse de-

sires, requirements, and possibilities – which helped explain why communicative disabilities arouse so little interest in the business sector. The diversity of disorders and requirements makes it impossible to develop products that everyone can use.

In 2009 a second study was conducted to determine how the development of HLT applications for people with communicative disabilities could be stimulated (Ruiter et al. 2010). Similar studies were carried out for the education sector and government organizations. The results of these investigations will be used as a basis to shape our HLT policy in the coming years.

In addition, we have being exploring the possibilities for a new Dutch-Flemish HLT programme that should address fundamental research, strategic research and application-oriented research and at the same time promote the re-use and re-purposing of available language resources and technologies for the Dutch language.

6. Concluding Remarks

In this paper we have outlined the HLT policy that has been developed and adopted in the Dutch language area by the Dutch Language Union together with a number of crucial Dutch and Flemish partners. We have explained the vision, rationale, approach and results of this 10-year endeavour. We hope that our efforts in developing a model to make a language "technology-ready" and the lessons we have learned from this experience may prove useful to other language communities that face the challenge of surviving in the digital era.

7. Acknowledgements

We would like to express our gratitude to our colleagues of the Dutch Language Union, the HLT Agency and the whole STEVIN team for their valuable cooperation

8. References

Beeken, J.C./van der Kamp, P. (2004): The Centre for Dutch Language and Speech Technology (TST Centre). In: Lino, M.T./Xavier, M.F./Ferreira, F./Costa, R./Silva, R. (eds.): *Proceedings of the 4th International Conference on Language Resources and Evaluation*. Paris: ELRA, 555-558.

Binnenpoorte, D./Cucchiarini, C./D'Halleweyn, E./Sturm, J./de Vriend, F. (2002): *Towards a roadmap for Human Language Technologies: the Dutch-Flemish experience. Proceedings of LREC2002*. Las Palmas de Gran Canaria.

Boekestein, M./Depoorter, G./van Veenendaal, R. (2006): Functioning of the Centre for Dutch Language and Speech Technology. In: *Proceedings of the 5th International Conference of Language Resources*. Genoa: LREC, 2303-2306.

Cucchiarini, C./Daelemans, W./Strik, H. (2001): *Strengthening the Dutch Language and Speech Technology Infrastructure, notes from the Cocosda Workshop 2001, Aalborg, 2 Sept. 2001*. 110-113.

Cucchiarini, C./D'Halleweyn, E. (2002): *A Human Language Technologies Platform for the Dutch language: awareness, management, maintenance and distribution*. Proceedings of LREC 2002, Canary Islands, Spain.

Cucchiarini, C./D'Halleweyn, E. (2004): The new Dutch-Flemish HLT Programme: a concerted effort to stimulate the HLT sector. In: *Proceedings of LREC 2004, Lisbon, Portugal*. 105-108.

Cucchiarini, C./Van Hoorde, J./D'Halleweyn, E. (2000): NL-Translex: Machine Translation for Dutch. In: *Proceedings of LREC 2000, Athens, Greece*. 1775-1780.

D'Halleweyn, E./Odijk, J./Teunissen, L./Cucchiarini, C. (2006): The Dutch-Flemish HLT Programme STEVIN: Essential Speech and Language Technology Resources. In: *Proceedings of LREC 2006, Genoa, Italy*. 761-766.

Rietveld, T./Stolte, I. (2005): *Taal- en spraaktechnologie en communicatieve beperkingen*. Den Haag: Nederlandse Taalunie.

Ruiter, M./Rietveld, T./Cucchiarini, C./Krahmer, E./Strik, H. (2010): Human Language Technology and communicative disabilities: requirements and possibilities for the future. In: *Proceedings of the 7th International Conference on Language Resources and Evaluation (*LREC *2010), Valletta, Malta*. 2839-2846.

Spyns, P./D'Halleweyn, E./Cucchiarini, C. (2008): The Dutch-Flemish comprehensive approach to HLT stimulation and innovation: STEVIN, HLT Agency and beyond. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation, Marrakesh, Morocco*. 1511-1517.

Spyns, P./D'Halleweyn, E. (2010): Flemish-Dutch HLT Policy: evolving to new forms of collaboration. In: *Proceedings of the 7th International Conference on Language Resources and Evaluation (*LREC *2010), Valletta, Malta*. 2855-2862.

Van Veenendaal, R./Van Eerten, L./Cucchiarini, C. (2010): The Flemish-Dutch HLT Agency: a comprehensive approach to language resources lifecycle management & sustainability for the Dutch language. In: *Proceedings of the Workshop Language Resources: From Storyboard to Sustainability and LR Lifecylce Management at the 7th International Conference on Language Resources and Evaluation (*LREC *2010), Valletta, Malta*. Valletta: LREC, 1-5.