COLLATE: Competence Center in Speech and Language Technology

Joanne Capstick*, Thierry Declerck*, Gregor Erbach†, Anthony Jameson*, Brigitte Jörg*, Reinhard Karger*, Hans Uszkoreit*†, Wolfgang Wahlster*†, Tillmann Wegst*

* German Research Center for Artificial Intelligence
Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany
collate@dfki.de

†Saarland University
Im Stadtwald, 66123 Saarbrücken, Germany
collate@coli.uni-sb.de

Abstract

This paper presents the structure and activitities of the recently established Competence Center in Speech and Language Technology in Saarbrücken. The objectives of the Competence Center are to provide a comprehensive information service about speech and language technologies, including live demonstrations of the most important language technology (LT) systems, and to advance the state of the art in the evaluation of LT systems for real-world applications. The Competence Center comprises the following components:

1. the Virtual Information Center "Language Technology World" (www.lt-world.org), the world's most comprehensive information resource about speech and language technology,
2. the Demonstration Center in Saarbrücken, which offers interested parties the possibility to play and experiment with different speech and language technologies, or to attend guided demonstrations,
3. the Evaluation Center, which conducts evaluations of the overall usability of language technology systems and advances knowledge of relevant usability issues and evaluation methods.

The work presented in this paper was carried out by the German Research Center for Artificial Intelligence in collaboration with Saarland University in the context of the project COLLATE (COmputational Linguistics and LAnguage TEchnology for Real Life Applications), funded by the German Federal Ministry of Education and Research (www.bmbf.de).

1. Overview

The German Competence Center for Language Technology was founded as a public service to the global R&D community in language technology and neighboring technology areas, to the German IT industry and to German companies and other organizations planning to employ language technology applications. In addition to a wide variety of free service functions, the Center also offers highly specialized professional services to private customers. Although the Competence Center focusses on the scientific and technological aspects of language technologies, it also builds up and offers competence on the commercial exploitation of LT applications and on their economic, social and ethical impact. The creation of the Center is funded by a grant (01 IN A01 A/B) from the German Federal Ministry of Education and Research in the project COLLATE, jointly carried out by DFKI and Saarland University.

The Competence Center consists of three components: a virtual information center, a demonstration center and an evaluation center. The virtual information center collects, maintains and disseminates information on all aspects of LT. It also serves as a competent host for the fast exploding wealth of collective knowledge in our fast growing subdiscipline of information technology.

The demonstration center gathers research software and commercial products from all over the world that exemplify the functionalities and the potential of our technologies. Representative systems of selected major application areas are installed at the center and skillfully demonstrated to interested parties from industry, research, academia and public administration.

The evaluation center concentrates on the user-centered assessment of LT methods, technologies, applications and specific products. Advanced methodology from cognitive psychology and HCI is adapted and extended to the study of usability and the adequacy of applications with respect to clearly specified tasks. The application of the results reaches from improved and new methodologies for system evaluation to studies of individual systems for specific customers.

2. Virtual Information Center

LT World is a general information service for the global language technology community. The fashionable term for such a Web-based information service is "portal", based on the metaphor of a specific entrance to the internet or to an intranet, leading straightforwardly to the information needed by a certain user community. In this sense LT World is a portal, yet we rather call our service a "virtual information center", because of its complex internal structure and the multitude of functions. LT World is a "virtual" center in the sense that most information physically remains with their creators or with other service providers

The center has been launched on October 25th 2001 under the name "LT World" (for "Language Technology World") and can be consulted at http://www.lt-world.org/.

The levels and tasks of LT-World can be seen in the following table:

	Specification Level	Technical Realization Level	Content Level
underlying logical structure	ontology specifications	concrete architecture	selection of sources, organization of collection
data maintenance structure	XML specifications	DBs, XML pages, HTML pages	content in DB, documents, links
presentational structure	generic design, CI	actual design of pages	presented contents

2.1. Contents of LT World

The wealth of information is divided into four top level areas

1. Knowledge and Information
2. Systems and Resources
3. Players and Teams
4. Communication and News

The first area is dedicated to the accumulated scientific and technological knowledge of our discipline enriched by information on the most important products and projects for each technology area. Entry points are technologies, methods and other basic concepts of the field. They are grouped in larger areas following the chapter outline of the upcoming revised and extended second edition of the Survey of the State in Human Language Technology (Cole et al., 1996; Carbonell et al., to appear). The entry for each technology starts with acronyms, alternative terms and a brief definition. The descriptions of theories, methodologies and technologies have been provided by domain experts. If the specific technology is covered in the first edition of the Survey, a link points to the relevant text. These links will be replaced by links to the corresponding sections in the new edition as they become available. Other links provided in each entry refer to products, projects, people, departments, relevant web pages and literature.

The second part of the Knowledge & Information section is a glossary of acronyms and other abbreviations that will gradually be expanded into a general glossary of the field. It will be linked to relevant texts and in this way serve as an extended subject index of the virtual information center.

The second top-level area, Systems & Resources, offers information on commercial products and research systems containing language technology. It is based on the Natural Language Software Registry maintained by DFKI under the auspices of the Association of Computational Linguistics (Declerck et al. 2000). Whereas the Software Registry does not make a difference between R&D systems and commercial software, LT World tries to distinguish. This is not always easy since many experimental systems and development platforms of R&D labs are also available under commercial licenses.

LT World also provides pointers to digital resources for research and development such as corpora, lexicons and grammars but in this area we mainly rely on the services of organizations dedicated to linguistic resources, mainly ELRA[1], LDC[2] and BAS[3]. We closely collaborate with the Open Language Archive Community (OLAC).

The area Players & Teams contains information on the members of the international R&D community, R&D projects and on all relevant organizations such as companies, academic departments, research centers, professional associations and funding agencies. The starting point of this area was a collection performed by student assistants with the help of state-of-the-art software tools. Each resource has been categorized according to the ontology described in section 2.2, and is described by an average of more than a dozen attributes per URL. As a default we have assumed that any researcher who publishes and any project that publicly announces its existence would be interested in being listed. So far only one researcher has suggested – not demanded – to be removed because of a transition into another professional carreer. Many individuals and projects have registered or provided additional information. For reaching and maintaining completeness, we will have to rely on self-registration. A web-form for registration is provided on LT World.

The fourth area, Communication & Events, is dedicated to short-lived information and to communication within the R&D community. So far two subareas are provided for news and for conference announcements.

Examples of newsworthy information are product announcements, company news, research breakthroughs, important events, and special honors for members of the LT community. A few news items are presented on our front page, others are provided sorted by date on the news page. Conference announcements are also ordered by date.

Currently there are 730 entries for organizations (classified as research institutions, companies and funding organizations), 450 entries on projects (organized along the main technologies they are investigating or implementing), 1500 entries on people of the field, 99 descriptions of theories, methodologies and technologies (also linked to the HLT Survey) and 290 NLP tools and products.

2.2. Conceptual Structure of LT World

Because it has been our goal to create the infrastructure of a virtual information center that can be extended in many directions, mainly in breadth, depth and technological sophistication, the structuring of the complex information has been based on a sophisticated ontology for language technology and computational linguistics. Most parts of the ontology are not visible on the surface of LT World. The highly multidimensionsional nature of scientific-technological information must not be directly reflected in the structure

[1] http://www.icp.grenet.fr/ELRA/
[2] http://www.ldc.upenn.edu/
[3] http://www.phonetik.uni-muenchen.de/Bas/, Bavarian Archive for Speech Signals, which participates in the German Competence Network for Speech and Language Technology together with COLLATE.

of the virtual information center because such a structure would not be apparent to the novice, to many LT experts on the business side of the community or to visiting experts from other fields of science and technology.

Therefore we have employed the ontology mainly for the conceptual design as well as for data collection and maintenance. The ontology has also helped in creating clean interfaces to other information services such as the Software Registry and the LT-Survey. It was even more important for the definition of the OLAC interface, a resource description schema that makes the contents of language archives cross-searchable.

In the future, the ontology will play a more predominant role since we will increasingly apply language and knowledge technologies for area-specific knowledge acquisition and maintenance.

This ontology has three layers (see figure 1). The first layer stems from general schemes for resource description such as the Dublin Core which has been extended by OLAC. Part of this core are such general properties as time, author, geographical place, topic and language.

The second layer stems from the specific types of described objects. It combines ontologies for publications, projects, institutions, products, researchers, conferences. These ontologies are not specific to our trade and in some cases already conventionalized or standardized as in the case of schemes for describing scientific publications. In the area of language resources, this layer overlaps with the OLAC scheme.

The third layer is specific to the discipline of language science and technology. It contains dimensions such as computational/mathematical methods, linguistic models/theories, level of linguistic description/processing (e.g., OCR, speech recognition, morphology, syntax, semantics, dialogue), application functionalities/types, relevant technologies, linguality (monolingual, bilingual, multilingual, crosslingual, language-independent), languages or language pairs.

2.3. User Interface

The user has access to this information through a portal (see figure 2) containing a header, a navigation bar to get to the top and secondary levels of information, forms for entering new data, and "Frontpage News" with links to more news and areas.

All parts of the virtual information center are searchable. As the search keys differ among areas, the search interface automatically adapts to the currently visited area. A global search facility permits the search of all areas on any substring.

The data in the different sections have been automatically linked to each other so that users can navigate easily between sections, for example to find out about the organizations and projects which a person has been involved with (see figure 3), or about the products which a company has developed.

2.4. Relation to other Initiatives

To enable both interchange with other databases as well as arbitrary post-processing of the data, the system is equipped with an XML interface.

We are working with the OLAC[4] (Open Language Archive Community, Bird & Simons 2001), for which one of our information services, the ACL Natural Language Software Registry, is a data provider (Declerck 2001). The sections which describe particular language technologies in depth will be developed further in close coordination with the production of the Survey of the State of the Art in Human Language Technology (2nd edition). We are also in close contact with the IMDI (Isle Metadata Initiative) Initiative[5] of the European Project "ISLE" (Broeder and Wittenburg, 2001).

2.5. Usage of LT World

Between October 2001 and March 2002, LT-World has had 4742 visitors (3002 from outside DFKI), that is an average of 26 (17) visitors per day. Visitors used the site extensively: each visitor spent an average of 592 (366) seconds on the site, and viewed 19 (15) pages. The total number of HTTP requests (hits) was over 160000.

LT World was accessed from sites all over the world. The ten most common top-level domains from which LT World was accessed are given below.

.de (Germany)	20.0 %
.net	11.1 %
.com	3.5 %
.edu	2.6 %
.se (Sweden)	2 %
.nl (Netherlands)	1.6 %
.dk (Denmark)	1.2 %
.fi (Finland)	1.1 %
.uk (United Kingdom)	1.0 %
.fr (France)	1.0 %

The most popular searches were for projects, persons, technologies and organisations. The following table gives a breakdown of the information requested by users.

Projects	5320	47.2 %
Persons	1530	13.6 %
Technologies	1359	12.1 %
Organisations	1260	11.2 %
Systems	705	6.3 %
Products	611	5.4 %
Conferences	459	4.1 %
News	33	0.3 %

2.6. Future Development of LT World

LT World will be extended in breadth, depth, functionality and technological sophistication. Although we will continue and improve the acquisition of new information, we also will have to rely on the contributions from the user community for extending the breadth of the information base. We already receive quite a number of additions and updates. All types of information can easily be entered through forms. Regional professional organizations have started to improve coverage of their

[4] http://www.language-archives.org/, Open Archives Iinitiative: http://www.openarchives.org/

[5] http://www.mpi.nl/ISLE/

respective geographical regions. We also received the product announcements from private enterprises.

Valuable improvements in depth will come through the second edition of the LT-Survey which will be published in the second half of this year on the Internet before a book edition is produced in 2003. Another source of increased depth (as well as breadth) will be the anthology of all ACL and COLING conferences and workshops that is currently produced by an initiative lead by Steven Bird under the auspices of ACL. This anthology will help us to greatly improve the bibliographic component of LT World.

Functionality will be added through the inclusion of thematic discussion lists and similar fora in the area Communication&Events. We also plan to offer a facility for submitting questions that will be sent to experts. Another planned addition to the functionality will come from collections of frequently and infrequently asked questions. The knowledge in a field can best be chunked as answers to questions. This technique constitutes the basis for most knowledge testing in the academic world. We will collect large sets of questions together with answers or hyperlinks to answers in digital texts.

The technological sophistication of LT World will be increased by applying advanced language and knowledge technology to the virtual information center. For maintaining and exploiting the ontology, we plan to utilize methods and tools from the semantic web development. These methods will also facilitate the creation and maintenance of interfaces to other information services on the WWW.

Language technology will be employed to a greater degree in the processing of gathering, filtering and integrating available information on the Internet. Our own systems for categorization, indexing, automatic hyperlinking and information extraction will be utilized. In order to maintain the quality of the integrated information, we foresee human quality control at the end of the acquisition process.

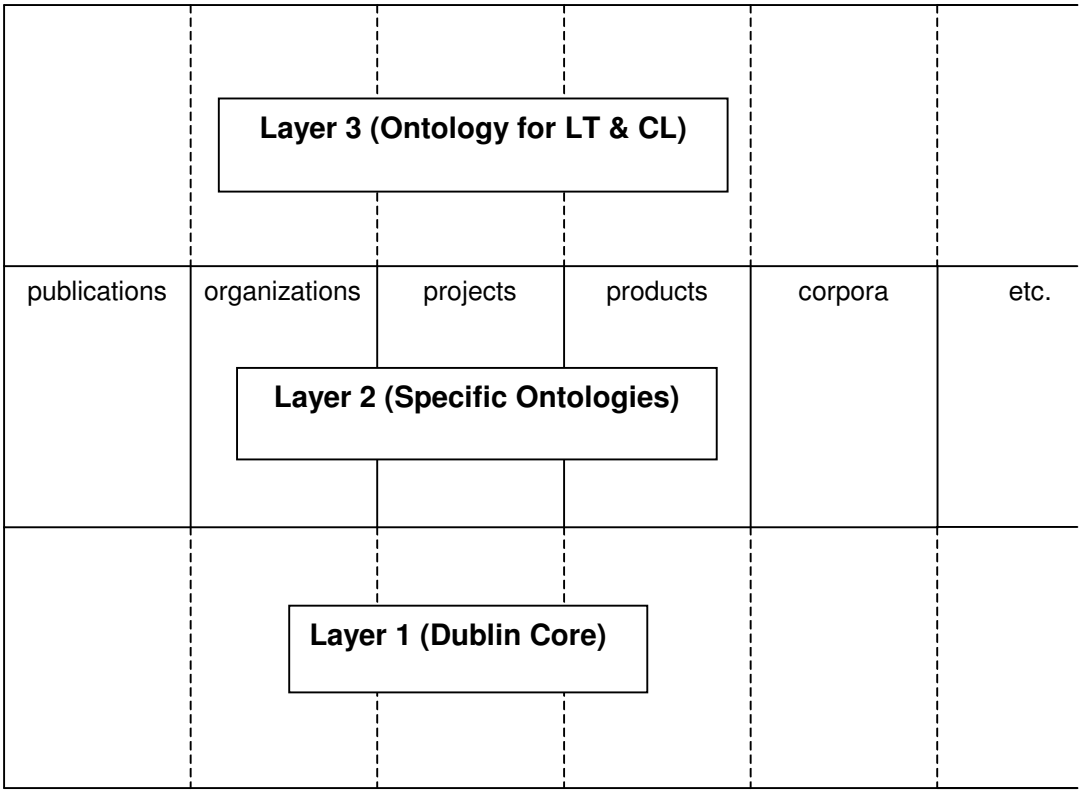

Figure 1: Ontology of LT World

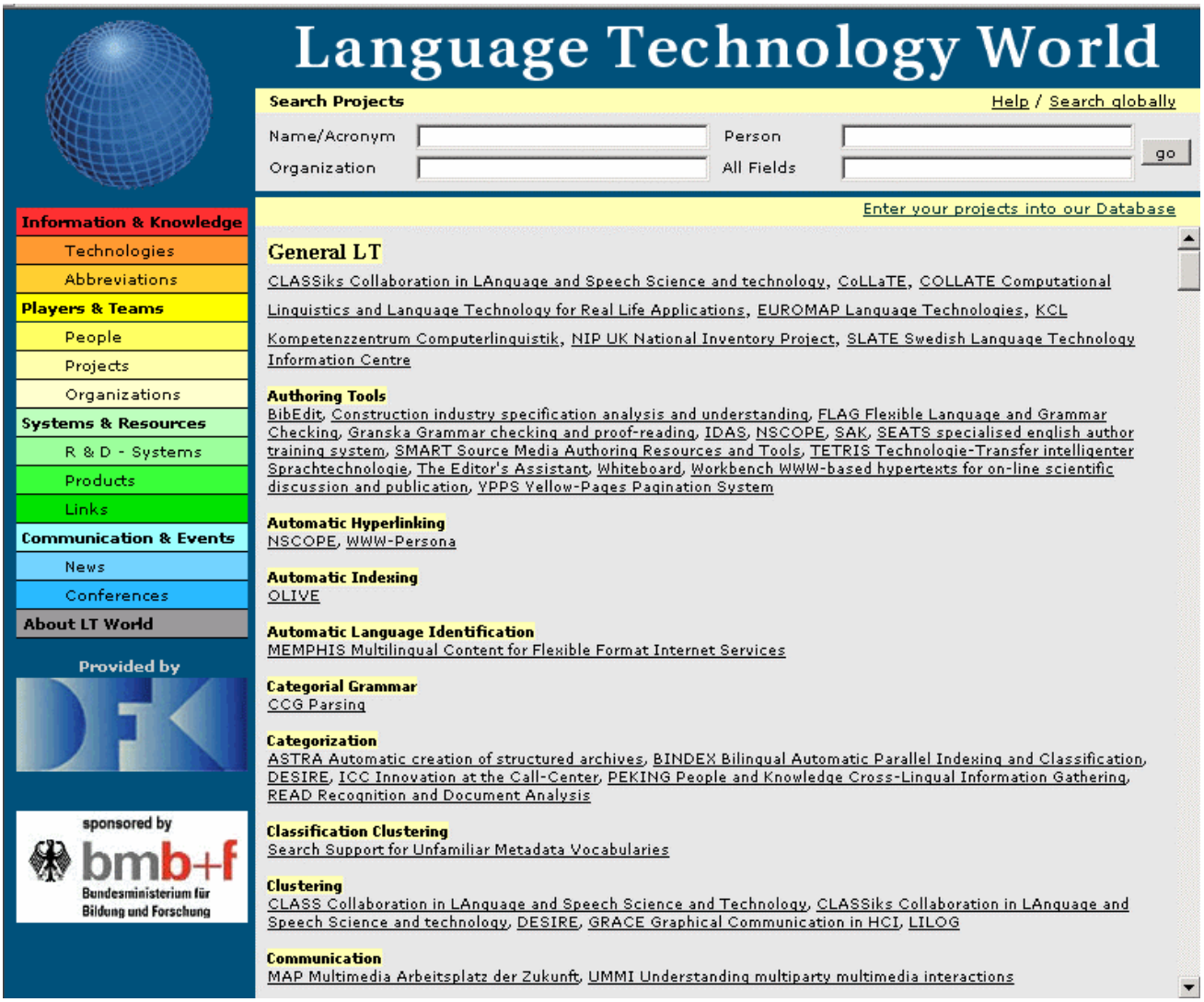

Figure 3: User interface of LT World

Figure 4: Automatic Hyperlinking in LT World

3. Demonstration Center

We have set up a cluster of PC-based demonstration kiosks on which a variety of speech and language technology products from different vendors can be demonstrated and tried out. The demo kiosks have been set up in a publicly accessible area in the lobby of DFKI, so that interested persons can try out LT products on their own. In order to enable guided demonstrations, the kiosks are linked with a data and audio-visual network, through which the input and output of any kiosks can be redirected the other kiosks, as well as to video projector and sound system. This makes it possible to give demonstrations for any number of people, from one individual to a group of up to 100 persons. This setup has been successfully used during the official opening of the demonstration and competence center to give a demo to a group of decision-makers from academia, government and industry, and for subsequent groups of visitors.

For in-depth demonstrations of language technologies, a special room has been set up which is equipped with two projection screens so that systems can be demonstrated in parallel to enable direct comparison.

Among the LT products installed on the demo kiosks are

- Voice Dialling for 300 phone numbers (Voicedirector)
- Telephone-based spoken dialogue system for stock market information (Sympalog Stocki)
- Finite-State Toolkit for NLP applications (XEROX XLE)
- morphological analysis and generation in three languages (mmorph)
- Text-to-speech synthesis (MARY, L&H RealSpeak)
- Platforms for Information Extraction (Lassie, SPPC)
- Multimedia indexing and search engine (MUMIS)
- Grammar checking for technical writers (SKATE)
- Dictation Systems (L&H Dragon NR, L&H ASR 16)
- Machine Translation (Linguatec Personal Translator)
- Automatic e-mail categorization and auto-response (Xtramind MailMinder)
- Spellchecking for professional users (CLT Corrigo)

Further systems are being selected and installed. Interested developers are invited to submit their systems for installation in the demonstration center.

4. Evaluation Center

Evaluation has rightly long been viewed as a crucial aspect of the success of LT systems (see, e.g., Hirschman & Thompson, 1997; and the report of the EAGLES Evaluation Working Group[6]). Methods have been developed and applied for evaluating the performance of LT system components of various types (e.g., the recognition accuracy of speech recognizers). But it is widely agreed (e.g., by several contributors to the chapter by Hirschman & Thompson, 1997) that methods for evaluating the *overall usability of LT systems for real users in typical contexts* have received much less attention than they deserve.

The COLLATE Evaluation Center (http://dfki.de/evaluation) aims to advance the state of the art in this area so as to encourage and facilitate this type of evaluation. An integral part of its work involves conducting evaluation studies of specific LT systems. For concreteness, we now present two small-scale examples of such studies.

4.1. Example 1: A Natural Language Music Retrieval System

The system BEAGLE, currently being developed by the firm Sonicson (http://www.sonicson.de), allows a user to enter into a Google-style interface a variety of natural language queries when searching for music to download, such as: “I'd like some fast songs by Fil Kollins”, “old German love songs”, or “snappier [than the songs currently shown on the screen]”. The LT capabilities (and the other novel capabilities) of BEAGLE are discussed by Baumann, Klüter, and Norlien (2002). What is important here is that the system can process a wide range of queries and that its capabilities are continually being expanded on the basis of analyses of queries entered by users. But there is a bottleneck that language technology itself cannot remove: Some users do not realize that they can express themselves flexibly: Transferring habits acquired from web search engines, they tend simply to enter names of artists and titles of songs. Since they obtain reasonable results in this way, they may never experience a need to express themselves more flexibly. Hence they may never benefit from the expressive power afforded by BEAGLE.

The COLLATE Evaluation Center looked into ways of presenting to users a few carefully chosen example queries that will quickly suggest to the user the kinds of queries that can be entered. At the Evaluation Center's stand at the CeBIT 2002 trade fair in Hannover, each of several users was shown the basic BEAGLE interface augmented with some example queries. With a remote eye tracker, each user's eye movements were recorded as he or she formulated each of several queries and examined the results returned.

Figure 4 shows how one user looked at an early version of the example-enhanced interface while making one query. The example section comprised two parts: five example queries (on the left) and five corresponding sets of concepts (on the right) that could be used in place of the highlighted words in the example queries. It can be seen that this user did not even look at the right-hand column. A retrospective interview with him revealed that he assumed that this column would just contain more examples like those on the left. Moreover, the distinction between examples and

[6] http://www.issco.unige.ch/ewg95/ewg95.html

possible variations on them seemed to be too complex for a user who was eager to enter his first query. This user also pointed out that two of the five examples (e.g., "I would like something by Fil Kollins") did not demonstrate any added value of natural language queries relative to the simple entry of names.

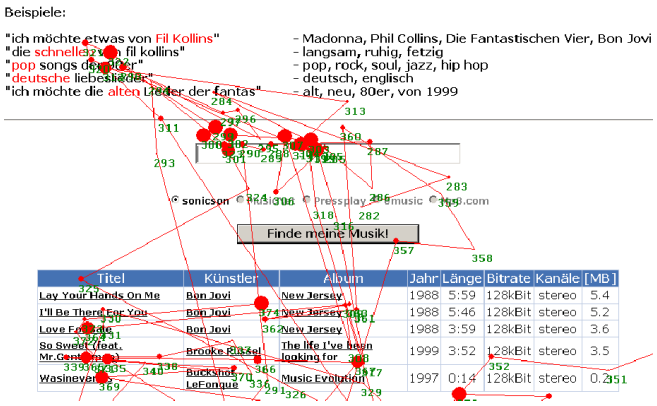

Figure 4: Eye movements of a user presenting a query to an early version of the example-enhanced interface of BEAGLE. (Each dot in the figure represents one fixation by the user, the size of the dot reflecting the length of the fixation. Numbers indicate the order of the fixations.)

On the basis of similar results with other users, the screen shown in Figure 5 was designed. The first of the five short examples (which means roughly "Something snappier from 2002!") was printed in huge orange letters so that users would be sure to read at least this one example. As the fixations (reproduced in the figure) of one user for one query illustrate, the examples did indeed receive close attention. On the other hand, the one extremely salient example was ignored: The user stated afterward that he assumed that it must be just a general heading or an advertising slogan.

This result suggests further optimizations of the display of the examples. It is hoped that the final version will not only be processed efficiently but will also be sufficiently compact for presentation on the screens of smaller devices such as PDAs.

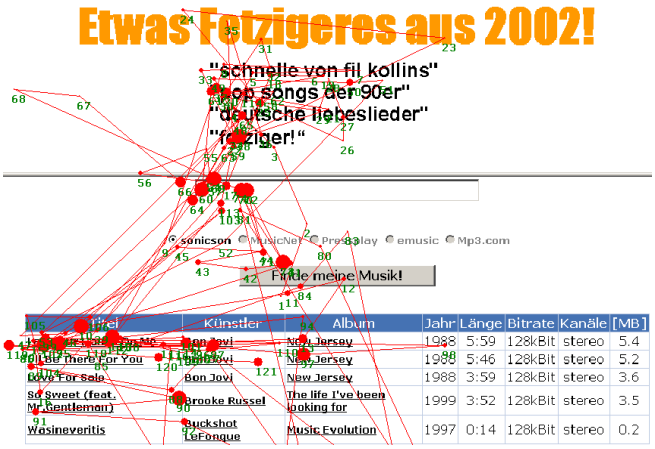

Figure 5: Eye movements of a user working with an improved version of the interface.

4.2. Example 2: Automatic Hyperlinking

The SHOW system (http://show.dfki.de/) automatically enhances web pages with hyperlinks that lead to explanations of preselected terms that occur in the text. The LT challenges raised by this system include those of (a) automatically finding the terms that have been singled out for explanation within a page even though they may appear in various grammatical forms; and (b) automatically determining which occurrences of a term (if not all of them) should be made into hyperlinks.

A key usability issue with this system concerns the way in which hyperlinks created in this way should be marked in the text (cf. Weinreich, Obendorf, & Lamersdorf, 2001, for a more general discussion of issues involving the appearance of hyperlinks). The type of marking of textual hyperlinks that is used by default in web pages (i.e., coloring and underlining) would be distracting, because of the relatively high proportion of words that would be highlighted. A minimally distracting solution was presented to the Evaluation Center for testing: Terms that have automatically been made into hyperlinks look just like all other words; but when the cursor passes over such a term, the form of the cursor changes to signal a hyperlink.

Quick usability tests were conducted with three users under the same conditions as those for the evaluation of the BEAGLE interface. All three users sooner or later began to search actively for hyperlinks by moving the cursor across each line of text – a behavior that was not intended by the system's designers and that presumably does not lead to natural or effective processing of the information in the text. One user, commenting retrospectively on the trace of his eye movements, explained that he at first checked for a hyperlink only when he encountered a term that interested him. But he became frustrated when, on two successive occasions, a word for which he had expected an explanation turned out not to have one. In order to avoid further frustration, he began searching actively for hyperlinks.

On the basis of these results, a version was tested in which the hyperlinked terms were printed in a dark brown color that is moderately distinguishable from the normal black text color. Initial feedback indicates that this form of marking is not distracting and that it allows users to make use of the automatic hyperlinks in the way that was intended by the designers. But since the hyperlinks are now visually identifiable, it becomes more important to be selective in the choice of terms to make into hyperlinks. For example, when a term occurs for the third time within a paragraph, in an unfocused part of a sentence, highlighting it seems relatively unnatural and distracting. Accordingly, further development of SHOW now includes efforts to apply linguistic criteria for the selection of term occurrences to make into hyperlinks.

4.3. General Points Illustrated

The specific results of these studies cannot yet be viewed as replicable and generalizable, because of the small samples involved; but they do illustrate several general points concerning usability evaluation for LT systems:

1. An LT system can fail, at least with some users, because of usability problems that are not due to any inadequacy of the LT components themselves.
2. These usability issues concern not only individual interface design flaws but also more interesting general issues that may apply to an entire class of LT systems. For example, the challenge of conveying quickly to users what sort of natural language input is appropriate is one that arises with just about any system that accepts such input. Consequently, it should be possible to synthesize and generalize results on the usability of LT systems, adding to the relatively limited knowledge base of relevant issues, findings, and guidelines that already exists (see, e.g., Dybkjaer & Bernsen, 2000, for a synthesis of this type that concerns spoken dialog systems).
3. An LT system can typically be viewed as a member of one or more classes of interactive system that do not necessarily involve language technology. For example, Sonicson's BEAGLE can be seen as a query-based information retrieval system, while SHOW can be seen as a system that generates web pages. For each of these classes of system, a good deal of knowledge has been acquired about usability problems and ways of investigating and combating them. The fact that a system uses language technology is of course no reason not to take such knowledge into account.
4. As is the case with any interactive system, attention to usability issues should not be left to a late stage in the design process. In particular, even usability problems that do not directly concern LT capabilities may have important consequences for the LT aspects of the system. For example, with SHOW, the need for more selective hyperlinking raises significant new challenges for the language technology component of the system.
5. Straightforward indices of usability such as task completion times and subjective ratings by users are not always adequate to support rapid, iterative improvement of an LT system's usability. In both of the two examples presented above, the method of having each user comment retrospectively on their eye tracking records yielded a fine-grained picture of the way in which users perceived and interacted with the system. This detailed understanding in turn made it possible to eliminate the problems uncovered quite quickly. This methodological result was not obvious in advance: Eye tracking has traditionally been seen as an expensive, cumbersome technique; and many people have expressed surprise at the idea of using it for the evaluation of LT systems.

4.4. Activities of the Evaluation Center

At the time of writing, the Evaluation Center is beginning to conduct larger-scale evaluations than the ones summarized above, for companies developing various types of LT system. It is expected that each study will yield some general insights into usability problems, possible solutions, and methodological issues with regard to the type of system in question. These insights will be synthesized and made generally available in publications and in web-based information resources.

5. Conclusion

The Competence Center for Language Technology fills an urgent need in the language technology community by providing comprehensive information services for an area of technology which is becoming ever more complex, and which is increasingly being used in a wide range of applications such as document and knowledge management, customer relationship management, search engines, office software, telephony applications etc. The Competence Center covers all areas of language technology in unprecedented breadth as well as depth. The demonstration center, evaluation center and specialized professional services are designed to transfer expert knowledge to developers, integrators and users of language technologies, and to technical journalists and political decision-makers with an interest in the area.

6. References

Baumann, S., Klüter, A., & Norlien, M. (2002). *Using natural language input and audio analysis for a human-oriented MIR system.* (Manuscript submitted for publication)

Bird, S. & Simons, G. (2001) The OLAC Metadata Set and Controlled Vocabularies, in Proceedings of the ACL/EACL Workshop *Sharing tools and Resources*, Toulouse, 2001.

Broeder, D. & Wittenburg, P. (2001) Interaction of Tools and Metadata-Descriptions for Multimedia Language Resources, in Proceedings of the ACL/EACL Workshop *Sharing tools and Resources*, Toulouse, 2001.

Carbonell, J., Mariani, J. , Ostendorf, M., Uszkoreit, H., Varile, G., Zaenen, A. & Zampolli, A. (to appear) *Language Technology - A Survey of the State of the Art.* to appear, 2002

Cole, R. A. , Mariani, J. , Uszkoreit, H., Zaenen, A. & Zue, V. (1996) *Survey of the State of the Art in Human Language Technology.* Cambridge University Press, 1996.

Declerck, T. (2001) Extending NLP Tools Repositories for the Interaction with Language Data Repositories, in Proceedings of the ACL/EACL Workshop *Sharing tools and Resources*, Toulouse, 2001.

Declerck, T., Jachmann, A.W. & Uszkoreit, H. (2000). The New Edition of the Natural Language Software Registry (an Initiative of ACL hosted at DFKI), in *Proceedings of the Second Conference on Language Resources and Evaluation*, Athens, 2000.

Declerck, T., Krauwer, S. & Rosner, M. (chairs) (2001): Proceedings of the ACL/EACL Workshop *Sharing tools and Resources*, Toulouse, 2001.

Dybkjaer, L., & Bernsen, N. O. (2000). Usability issues in spoken dialogue systems. *Natural Language Engineering*, *6*(34), 243--271.

Hirschman, L., & Thompson, H. S. (1997). Evaluation. In R. A. Cole, J. Mariani, H. Uszkoreit, A. Zaenen, &V. Zue (Eds.), Survey of the state of the art in human language technology (pp. 475–518).

Weinreich, H., Obendorf, H., & Lamersdorf, W. (2001). The look of the link – concepts for the user interface of extended hyperlinks. In *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia*