Deliverable D5.2

Final Report
Covering the period from: 1.6.98 to 28.2.99

August 1999

Esprit Long-Term Research Concerted Action No. 24823

Spoken Language Dialogue Systems and Components: Best practice in development and evaluation.

DISC

TITLE	**Final Report** **Covering the period from 1.6.98 to 28.2.99**
PROJECT	DISC (Esprit Long-Term Research Concerted Action No. 24823)
EDITORS	Niels Ole Bernsen and Laila Dybkjær, NISLab
AUTHORS	The DISC Consortium
ISSUE DATE	27 August 1999
DOCUMENT ID	Wp5d2
VERSION	0.2
STATUS	Final
NO OF PAGES	41
WP NUMBER	5
LOCATION	http://www.disc2.dk
KEYWORDS	WP5, spoken language dialogue systems, final report

Document Evolution

Version	Date	Status	Notes
0.1	22/07/99	Draft	First draft published for comments from partners.
0.2	27/08/99	Final	Final version

Spoken Language Dialogue Systems and Components: Best practice in development and evaluation

DISC 24823 Final Report
Basic Details of the Action

Niels Ole Bernsen and Laila Dybkjær
Natural Interactive Systems Laboratory

Co-ordinator

Natural Interactive Systems Laboratory
Science Park 10
5230 Odense M
Denmark

Partners

Daimler-Chrysler Research and Technology, Ulm, Germany
European Language and Speech Network (Elsnet), Utrecht, The Netherlands
Institut für Maschinelle Sprachverarbeitung, Stuttgart University, Germany
Kungliga Tekniska Högskolan (KTH), Stockholm, Sweden
Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI), Paris, France
Vocalis LTD, Cambridge, UK

Co-ordinator Contact Point

Prof. Niels Ole Bernsen
Tel. (+45) 65 50 35 44
Fax (+45) 63 15 72 24
email: nob@nis.sdu.dk

Table of Contents

1. Executive Summary

1.1 Abstract

This DISC project aims to (a) build an in-depth understanding of the state-of-the-art in spoken language dialogue systems (SLDSs) and components development and evaluation with the purpose of (b) developing a first best practice methodology in the field. The methodology will be accompanied by (c) a series of development and evaluation support tools. To the limited extent possible within the duration of the project, the draft versions of the methodology and the tools will be (d) tested by SLDS developers from industry and research, and will be (e) packaged to best suit their needs. (a), (b) and (c) have been accomplished. (d) and (e) have started and are being continued in DISC-2 with the purpose of optimising the usability of the DISC results for developers and deployers of SLDSs.

1.2 Review of Aims

The DISC aims listed in Section 1.1 above, may appear more worthwhile than ever in the light of recent developments. Spoken language dialogue systems are gaining ground in the market place as witnessed by forefront products installed in Europe by, among others, Vocalis, Daimler-Benz, Philips Aachen, Lernout and Hauspie and CSELT, as well as by newcomers in Europe from the US, such as Nuance and Speechworks, and by the rapid spread of known SLDS technologies by European telecoms and others. During the first year of DISC, DARPA in the US decided to start building up strength in the field in view of the massive build-up in speech technology in the US (cf. Business Week 15.2.1998). The DARPA Communicator project is now underway [http://fofoca.mitre.org/]. The book which formed the starting-point of DISC (Bernsen et al. 1998) is selling very well and has been reviewed as constituting essential reading for developers (Elsnews 7.4, 1998). The Commission's FP5 Programme on Human Language Technologies now explicitly emphasises best practice methodologies in speech and language engineering. Imperfect as it still is, we hope that the DISC and DISC-2 best practice work can help show the way for future speech and language engineering best practice efforts.

Starting from their pre-DISC theoretical strengths and experience in development and evaluation of SLDSs and their components, the DISC partners spent the first year testing and revising the lessons learnt from that experience, leading to a deeper and broader understanding of the issues involved in developing and evaluating spoken language dialogue systems and their components. The results of DISC Year 1 gave the Consortium as strong a basis as it is probably possible to have today for drafting a first best practice methodology in the field. This draft best practice methodology is now available for further testing and final packaging in DISC-2.

1.3 Progress and Results

During the first year of DISC, WP1 was completed in draft form and work started on WP2 (tools) and WP3 (best practice draft). During the final 9 months of DISC, the WP2 tools and

the WP3 best practice drafts have been completed. Packaging has started through the establishment of a website dedicated to a structured presentation of the DISC best practice results. WP4 (dissemination) has been active throughout, "changing gear" with the arrival of the first public DISC results in the summer of 1998. The DISC Advisory Panel now includes 50 members. Notable innovative results in the DISC period reported here include a common approach to best practice presentation involving issues, options and pros and cons, and a fundamental approach to the evaluation of spoken dialogue systems (SLDSs) and components based on systematic identification of relevant evaluation criteria from the DISC best practice 'grid' models for the SLDSs aspects investigated in the project (see below). It is expected that these innovations will help making the packaging of DISC results in DISC-2 more useful to developers.

For the field, DISC is highly interdisciplinary, addressing virtually all *aspects* of spoken language dialogue systems, including:

- speech recognition
- speech generation
- natural language understanding and generation
- dialogue management
- human factors
- systems integration

In addition to this "horizontal" spread of expertise, the "vertical" spread of expertise ranges from industrial development (all aspects) to theoretical research into human-human spoken dialogue and the foundations of evaluation.

WP1 Results

The WP1 investigation of current practice following a common methodology developed in DISC and comprising eight deliverables D1.1 through D1.8, was completed during the first year of DISC. Results are summarised in D1.8. See also D5.1. A publication is: Bernsen, Dybkjær and Heid 1999 (Eurospeech).

WP2 Results

In parallel with development of the DISC draft best practice methodology (WP3), the following surveys, concepts and software tools have been developed for testing in industrial and academic environments:

LIMSI

D2.1 Survey of existing and easily available platforms and development methods for testing and enhancing the performance of Speech Recognition components. *Completed.*

This report was added to ensure that Disc takes into account existing methodologies used by the speech recognition community, and serves as a basis for D2.2. The survey covers those systems that are available either as freeware, as freeware for research purposes, or from groups with close working ties in the research community. Major industrial suppliers of products and turn-key systems are not included.

D2.2 Guidelines and testing protocols for the development of speech recognition components for SLDSs. *Completed.*

This report reviews the fundamental issues in using a corpus to evaluate a speech recogniser, including the "why", "when" and "how" of evaluation. Less formal methods specific to the development of speech recognisers for Spoken language Dialogue Systems are discussed as supplementary to the corpus methods.

KTH

D2.3: A survey of existing methods and tools for development and evaluation of speech synthesis and speech synthesis quality in SLDSs. *Completed.*

This deliverable gives short overviews of commercial TTS systems and of freeware speech synthesis systems. For nearly all the described systems, a website address that gives further information is supplied or, for some freeware systems, an ftp-address is supplied from which the system can be downloaded. The deliverable also contains information about speech synthesis evaluation, especially intelligibility tests.

D2.4: Software tool for evaluation of speech synthesis components in SLDSs. *Completed.*

This is a simple tool to help developers choose the right type of synthesis for their application. The tool will be implemented on the web.

IMS

D2.5: Concepts, Methods and Tools for the acquisition of lexical data in Spoken Language Dialogue Systems. *Completed.*

The report contains an overview of approaches, methods and tools for lexical acquisition in Dialogue Systems; it describes environments such as the acquisition scenario at DaimlerChrysler's research laboratories and in the Verbmobil project, but also commercial tools such as temic's LDS (Lexicon Development System). A brief summary of the main features of the Java Speech Grammar Format, JSGF, is given. The surveyed techniques are positioned within the scenario of SLDS developent, including the interrelationship with other components of SLDSs. User groups and functions of acquisition tools (and their interrelationships) are discussed.

D2.6: Towards guidelines for the acquisition of lexical data in Spoken Language Dialogue Systems. *Completed.*

This short report contains a few recommendations for lexical acquisition, both with respect to functions of acquisition tools, and with respect to relevant properties of the SLDSs for which the acquisition is supposed to take place. As tool support for acquisition is only upcoming, ‘guidelines’ sensu strictu cannot be formulated yet: too little experience is available from which to draw; thus the report is more meant to provide a checklist of problems to be kept track of than to furnish a full guideline. Focus is on types of data to be acquired, support functions of acquisition tools, user groups and on consitency control (in terms of definitions of lexical classes, and in terms of the concrete implementation); the use and usefulness of transcribed spoken data along with (variants of) tools and methods from (text)corpus linguistics is assessed as well. The relevant degrees of user involvement in acquisition work are discussed, in particular against the move towards more involvement of users (not only deployers) in lexicon maintenance.

NIS

D2.7a. State-of-the-art survey of existing dialogue management tools. *Completed.*

This deliverable provides an overview of existing dialogue management tools.

D2.7b. State-of-the-art survey of existing human factors tools. *Completed.*

This deliverable provides an overview of existing human factors tools.

D2.8. Concepts and a diagnostic methodology for the identification of user-system interaction problems, their typology, severity and remedies. Software tool in support of cooperative system dialogue design. *Completed.*

This deliverable includes a report as well as a software tool. The report provides a specification of the tool and has been completed. The software tool which supports developers and testers of SLDSs to improve dialogue cooperativity exists in a first version. This version will be iterated and improved throughout DISC-2. Seven publications are: Bernsen, Dybkjær and Dybkjær 1997 (IEEE Computer), Bernsen, Dybkjær and Dybkjær 1997 (book chapter), Bernsen, Dybkjær, Dybkjær and Zinkevicius 1997 (Eurospeech), Bernsen, Dybkjær and Dybkjær 1998 (book), Dybkjær, Bernsen, and Dybkjær 1997 (ACL Workshop), Dybkjær, Bernsen, and Dybkjær 1997 (book chapter), Dybkjær, Bernsen, and Dybkjær 1998 (International Journal of Human Computer Studies/Knowledge Acquisition).

D2.9. Software tool in support of speech functionality decisions in early design. *Completed.*

This deliverable includes a report as well as a software tool. A DISC Working Paper which specifies the tool has been completed. This paper was not planned in the Project Programme and has been given the identifier D2.9. The software tool which supports developers in deciding whether or not to use speech in their applications, exists in a first version. This version will be iterated and improved throughout DISC-2. Five publications are: Bernsen 1997 (Speech Communication), Bernsen, Dybkjær and Dybkjær 1998 (book), Bernsen and Dybkjær 1998 (ICSLP), Bernsen and Dybkjær 1999 (ESCA), Bernsen and Luz 1999 (IEEE).

D2.10. Working Paper on Speech Functionality. *Completed.*

This is a second DISC Working Paper which was not planned in the Project Programme. The paper is a comprehensive study of data on Speech Functionality and has been given the identifier D2.10.

WP3: Results

Based on the results of WP1, a working document on the detailed DISC best practice methodology (D3.1) was completed by the end of the first year of DISC. The following results can be noted during the final phase of DISC:

LIMSI

D3.2 Draft proposal on best practice methods and procedures in speech recognition. *Completed.*

This report summarizes key aspects of the best practice in speech recognition in the design and implementation of commercial and research spoken language dialogue systems. The discussion of this topic is structured around a sequence of design issues, the options for their implementation, and their pros and cons. This information is discussed in the context of the particular types of task and development environments likely to be encountered when building current state-of-the-art speech recognisers for spoken language dialog systems. A publication is Lamel 1998 (Int'l Symposium on Spoken Dialogue).

KTH

D3.3: Draft proposal on best practice methods and procedures in speech generation. *Completed.*

The deliverable gives a first version of a best practice method for speech generation in SLDSs. It will be tested and developed further in DISC-2. It contains mainly the grid part of the DISC model where experiences gained in WP1 and general knowledge has been implemented. Part of the deliverable discusses evaluation of different parts of speech generation. A publication is Karlsson and Thornton 1999 (Computer Telephony Europe).

IMS

D3.4: Draft proposal on best practice methods and procedures in natural language understanding and generation. *Not completed.*

The report summarizes the best practice proposals for the major design and evaluation issues in NL components of SLDSs. At the difference of other SLDS components, NL components differ massively in size, approach and sophistication, from one SLDS to the next. A broad classification is operated, along with a detailed discussion of the criteria underlying it; as a result, systems with keyword spotting (no NL), systems with (partial) grammars (similar to Information Extraction), systems with grammars covering full sentences, and systems based on sophisticated syntactic, semantic and pragmatic processing are distinguished. For each type, options available for the major design and evaluation issues are discussed, mainly againts the background of the major constraints on NL components in SLDSs.

A publication is Thornton 1999 (Computer Telephony Europe).

D3.8b: From a description of spoken language dialogue systems to their evaluation and best practice. *Completed.*

This deliverable presents what is called IDEM (integrated DISC evaluation model). The model takes as its point of departure a grid description and generates on that basis a set of evaluation questions. The idea is in this way to facilitate the easy creation of a uniform and covering set of evaluation questions tailored to the specific application.

NIS

D3.5: Draft proposal on best practice methods and procedures in dialogue management. *Completed.*

This deliverable gives a first version of a best practice methodology for dialogue management in SLDSs. A publication is: Bernsen and Dybkjær 2000 (Advances in Computers).

D3.8a: DISC Dialogue Engineering Best Practice Methodology. *Completed.*

This deliverable reviews the first complete version of the DISC dialogue engineering best practice methodology, synthetising the DISC WP2 and WP3 results.

D3.10: Working Paper on Dialogue Management Evaluation. *Completed.*

This DISC Working Paper was not planned in the Project Programme and has been given the identifier D3.10. The paper has been done in response to comments from the DISC Advisory Panel at the DISC First Year Review. A publication is Bernsen and Dybkjær 1999 (MIT).

Vocalis

D3.6: Draft proposal on best practice methods and procedures in human factors. *Completed.*

This deliverable gives a first version of a best practice methodology for human factors in SLDSs. A publication is Thornton and Dybkjær 1999 (Computer Telephony Europe).

D3.7: Draft proposal on best practice methods and procedures in systems integration. *Completed.*

This deliverable gives a first version of a best practice methodology for systems integration in SLDSs. A publication is Thornton 1999 (Computer Telephony Europe, Communications MEA).

WP4 Results

See 1.6 below.

WP5: Results

WP5 is the Management Work Package. The following items are reported here (see also D5.1):

(1) NIS has sub-contracted the creation of a DISC Results Packaging website to Mrs. C. J. Wilson. As Mrs. C. J. Wilson has had obvious problems meeting the contract requirements on time, the site set-up has been delayed and NIS will have to step in to do the work. The site is being set up at the time of writing.

(2) Lin Chase has left LIMSI in February 1999. Wolfgang Minker has been added to the LIMSI DISC group.

(3) Klaus Failenschmid has left Vocalis. Mike Thornton has been added to the Vocalis DISC group.

(4) Due to the AP's timely request for additional work on evaluation as part of the DISC best practice effort, most WP3 deliverables have been delayed with the result that the synthesis of WPs 2 and 3 (D3.8) has been delayed. As Klaus Failenschmid has left Vocalis, D3.8 has been done by NIS. The request for more work on evaluation has occasioned DISC (NIS) to propose a fundamental approach to evaluation which has so far been applied to dialogue management evaluation (D3.10) and human factors evaluation (D3.6). Deliverable D3.10 has been done in addition to those planned in the Project Programme. The approach presented therein was discussed at the DISC-2 May 1999 Workshop in Paris and was adopted as a common approach for packaging the DISC results in DISC-2.

(5) Elsnet will organise the planned Best Practice Conference for industry and end-users (D4.4) towards the end of DISC-2.

(6) One deliverable in addition to those originally planned has been done on speech functionality (D2.10).

(7) IMS has had considerable delays in producing Deliverables D2.5, D2.6 and D3.4. The delays are due to excessive internal workload at IMS.

(8) The planned D2.7 has been split into two separate deliverables: D2.7a, a state-of-the-art survey of existing dialogue management tools, and D2.7b, a state-of-the-art survey of existing human factors tools.

(9) Due to the delay of D3.8, the AP report on the results of DISC (D3.9) has been delayed. A time schedule for its production has been agreed on the DISC May 1999 Workshop. D3.9 will be finalised before the end of June 1999.

1.4 Internal Collaboration

Three workshops have been organised, in Stuttgart, Germany, 8-10 June 1998 (3 days), Ulm, Germany, 18-20 November 1998 (3 days), and Cambridge, UK, 21-22 January 1999 (2 days), respectively. In addition, the DISC work was closed at the joint DISC-DISC-2 workshop in Paris, France 5-7 May 1999.

Cross-site authorship frequently turns out to force theoretical discussions which might not have happened otherwise within the project. From the period reported here, the following six DISC Working Papers have cross-site authorship (cf. Section 1.3):

D2.7b: NIS, Vocalis.

D3.5: NIS, IMS.

D3.6: Vocalis, NIS.

D3.7: Vocalis. LIMSI.

D3.8a: NIS, all partners.

D5.2: NIS, all partners.

Three cross-site publications have been produced so far: in the first year of DISC, one cross-site DISC publication was produced by the entire DISC team for the Granada Conference in May 1998. In the second year of DISC, a cross-site NIS/IMS paper has been produced for Eurospeech'99 and a second NIS/IMS paper has been submitted. More cross-site publications are in preparation.

A final mechanism of collaboration has been the DISC website established by Elsnet, which is being used extensively for uploading draft working papers for easy access by the partners. Each uploading of a paper onto the "Partners only" web pages is accompanied by a message from the website robot to all DISC partners that a paper with a specified identity is now available on the website.

1.5 External Collaboration and Use of Results

For the purpose of this section, "external collaboration" includes any DISC-related collaboration undertaken by a DISC partner with non-DISC partners, including site-internal collaboration.

NIS

DISC results are being used in the following projects:

ELSE Telematics/Language Engineering project (http://www.limsi.fr/TLP/ELSE). The ELSE project has produced a blueprint for evaluation in language and speech engineering. The blueprint was originally focused solely on black-box evaluation of language and speech technologies, but has recently been augmented to include usability evaluation.

MATE Telematics/Language Engineering project (http://mate.nis.sdu.dk). In MATE, a coding scheme for communication problems is being developed and implemented in the MATE Workbench. The coding scheme is based on the DISC-tooled cooperativity principles and can help evaluators annotate communication problems in spoken language dialogues in a state-of-the-art annotation environment.

In addition, DISC results are being proposed for use in FP5 projects.

References

Paroubek, P. and Blasband, M. (Eds.), Bernsen, N. O., Blasband, M., Calzolari, N., Chanod, J.-P., Choukri, K., Dybkjær, L., Gaizauskas, R., Krauwer, S., Lamberterie, I. de, Mariani, J., Netter, K., Paroubek, P., Rajman, M. and Zampolli, A.: *Evaluation in Language and Speech Engineering.* Executive Summary Draft of a Blueprint for a General Infrastructure for Natural Language Processing Systems Evaluation Using Semi-Automatic Quantitative Black Box Approach in a Multilingual Environment. Telematics - Language Engineering project ELSE Final Deliverable. Paris: LIMSI, April 1999.

Klein, M., Bernsen, N. O., Davies, S., Dybkjær, L., Garrido, J., Kasch, H., Mengel, A., Pirelli, V., Poesio, M., Quazza, S. and Soria, C.: *Supported Coding Schemes.* Telematics/Language Engineering Project MATE Deliverable D1.1. Odense University, Denmark, July 1998.

Isard, A., McKelvie, D., Cappelli, B., Dybkjær, L., Evert, S., Fitschen, A., Heid, U., Kipp, M., Klein, M., Mengel, A., Møller, M.B. and Reithinger, N.: *Specification of Workbench Architecture.* MATE Deliverable D3.1. Odense University, Denmark, August 1998.

Talks and seminars:

SALT workshop 1997: Presentation of DISC given by L. Dybkjær.

ACL/EACL 1997: Presentation of NIS WP2-related results given by L. Dybkjær.

Eurospeech 1997: Presentation of NIS WP2-related results given by L. Dybkjær.

ICSLP 1998: Presentation of NIS WP2 related results given by L. Dybkjær.

Advanced Voice-enabled Applications for the Car, Brussels, 26.10.1998. Presentation of DISC given by N. O. Bernsen.

IST'98 Workshop on The Language of Business – the Business of Language, Vienna, 2.12.1998. Presentation of DISC-related results given by N. O. Bernsen.

In addition, DISC results are being used in teaching computer science students.

LIMSI

DISC results have been or are being used in the following projects:

LE ARISE project.

Aupelf B1 and B2 activities.

ELSE Telematics/Language Engineering project.

DARPA Communicator (preliminary contacts).

References

Paroubek, P. and Blasband, M. (Eds.), Bernsen, N. O., Blasband, M., Calzolari, N., Chanod, J.-P., Choukri, K., Dybkjær, L., Gaizauskas, R., Krauwer, S., Lamberterie, I. de, Mariani, J., Netter, K., Paroubek, P., Rajman, M. and Zampolli, A.: *Evaluation in Language and Speech Engineering.* Executive Summary Draft of a Blueprint for a General Infrastructure for Natural Language Processing Systems Evaluation Using Semi-Automatic Quantitative Black Box Approach in a Multilingual Environment. Telematics - Language Engineering project ELSE Final Deliverable. Paris: LIMSI, April 1999.

Talks and seminars:

LREC'98: Keynote L. Lamel 30/05/98 Granada: On the Role of Evaluation in Spoken Language System Development.

LREC'98: Disc talk (for the paper on the DISC approach [Dybkjær et al. 1998]) given by L. Lamel 28/05/98.

Disc presentation by L. Lamel at the Arise final workshop, October 1998, Torino.

KTH

DISC results are being used in teaching.

There has been some informal co-operation at KTH concerning identification of existing speech synthesis systems. This resulted in the following seminar: I.Karlsson gave an internal seminar at KTH about existing speech synthesis. Attending were about 20 doctoral students and staff.

I. Karlsson made a short presentation of the DISC project on 3 September 1998 at the Center for PersonKommunikation, Aalborg University at a meeting where people from Aalborg University, Trondheim Science University and KTH participated.

IMS

Site-internal collaborations using DISC results:

SFB 340 (Theoretical Foundations for Computational Linguistics): DISC survey results used in the work on the semantic and pragmatic modelling of dialogues (Jan v. Kuppevelt).

Verbmobil, Workpackage on Transfer: DISC results (especially from DISC analysis of Natural Language and of Dialogue Managament aspects) used in work on the formulation of lexical and (dialogue-context-dependent) grammar rules for transfer; similarly: the Verbmobil system implementation available through STRs involvement in the Verbmobil project has been used in the DISC WP-1 work.

MATE (LE-4): The DISC survey results are used as a background for the definition of requirements with respect to the functionalities required in a multilevel-annotated corpus: SLDS developers are seen as a major client of the type of corpora to be produced in MATE (Ulrich Heid, Andreas Mengel (MATE collaborator in STR)); Experience from the presence of explicitly modeled interactions between different levels of description in dialogues (esp. as visible from DISC WP-1 analyses of Verbmobil and Waxholm Dialogue Management) is used in MATE as an exemplification of the types of interactions between levels necessary from the point of view of SLDSs (Jan v. Kuppevelt); this also applies to cases where a thorough description of the facts would require such interaction to be explicitly modeled, but it isn't in the exemplars analysed in DISC.

Ongoing work towards a proposal for dialogue analysis in terms of the question/answer approach, integrated with DRT: work in close connection with SFB-340 (Theoretical Foundations for Computational Linguistics): J. van Kuppevelt, Hans Kamp

Exchange with the VERBMOBIL group (IMS is the main node in the Transfer sub-project of Verbmobil), on dialogue modelling for translation (UH, JvK, Michael Schiehlen, H. Kamp) and on lexical acquisition (UH, Martin Emele).

Regular exchange with the STR part of the MATE (LE-4/8377) project (A. Mengel), on dialogue modelling, query of dialogue corpora and lexical acquisition (WP-2, UH and JvK).

Participation of JvK in the PROCOPE 2000-proposal `"Topik" im Dialog und in der Logik' (Cooperation between IMS-FLS (HK, Peter Krause, JvK, Uwe Reyle) and IRIT, Toulouse; proposal is submitted; intention to integrate results of the PROCOPE exchange in DISC-2 work and, vice versa, to have DISC as a provider of concrete cases for the PROCOPE proposal).

External Collaborations

Close collaboration with the MATE (LE-4/8377) project:

- dialogue modelling (WP-1/5), dialogue acts: definition and annotation (MATE, WP-2); a meeting with MATE partner DFKI was held in STR on 08-02-99, to cover these issues (UH, JvK)

- querying multilevel corpora (MATE, WP-3) vs. acquisition of lexical and grammatical information (WP-2); a meeting with MATE AP members Melina Alexa and Lothar Rostek was held in Stuttgart on 07-12-98, to cover these aspects.

DISC/MATE interaction with NIS Odense: MATE query tool will be used in the CODIAL tool of DISC (UH, A.Mengel; L. Dybkjær)

Collaboration with Department of Computational Linguistics and Department of Philosophy of the University of Amsterdam, in particular within the framework of Amstelogue'99 (Jan van Kuppevelt).

Talks and seminars:

Heid, Ulrich: Objectives and working procedure of DISC, WP-1, at the DISC/Verbmobil joint working meeting in Saarbruecken, 28-11-1997.

Seminars attended:

Heid, Ulrich: Session on Spoken Language Dialogue System Evaluation (1st of 2), in the framework of the 1st Intl. Conference on Linguistic Resources and Evaluation, Granada, 28-5-1998.

In addition, DISC results are being used in teaching.

van Kuppevelt, Jan and Heid, Ulrich: Seminar on Spoken Language Dialogue Systems, in the framework of the curriculum in Computational Linguistics at IMS Stuttgart (13 * 90 min): for 3rd year students in Computational Linguistics (Spring 1999).

van Kuppevelt, Jan: Talk at the International Workshop on the Relationship Between Prosody and Meaning, Prague, November 1998. (Invited speaker).

van Kuppevelt, Jan: Talk on Focus in Linguistics and Artificial Intelligence, Discourse Processing (DIP) Colloquium, Department of Computational Linguistics, University of Amsterdam, December 1998.

van Kuppevelt, Jan: Program chair and co-organizer of the Amstelogue'99 Workshop on the semantics and pragmatics of dialogues, dialogue systems and dialogue analysis (DISC supported invited speakers: Hans Kamp (also Stuttgart University) and Bonnie Webber (University of Edinburgh)).

Vocalis

'Guidelines for Advanced Voice Dialogues' Project (ESRC Project ref: L127251012 by the Department of Sociology, University of Surrey, UK and by Vocalis Ltd.).

This project is concerned with establishing guidelines for the development of advanced voice dialogues.

The work of the Guidelines project focuses on a traditional framework for dialogue design. 'Hanging off' this framework are a variety of guidelines. The guidelines are based on:

- Linear and iterative approaches
- Empirical investigation of *de facto* standards in commercial dialogues.

Central to the life-cycle is the applicability and uptake of the guidelines within an industrial context. To this end, the project began with an investigation of the current design practice in commercial environments (See Cheepen, 1996). With this as a starting point, the work has focused on identifying the different audiences within the company which are involved in dialogue design, i.e. developers, marketing/sales and researchers. Each group has a stake in the system but very different requirements in terms of the information they require in order to follow a suitable design process. Only by addressing each group in their own terms, can a process description hope to succeed.

In addition to the iteration of the process within a commercial setting, the project also challenges *de facto* standards within the industry. This includes:

- The use of human-like tokens in (English) system output prompts, e.g. 'please', 'thank you', 'I', 'you' (See Williams and Cheepen, 1998).
- The use of only verbal aural output in dialogue.

References

Cheepen, C. (1996) "Designing advanced voice dialogues - what do designers do and what does this mean for the future?", *http://www.soc.surrey.ac.uk/research/reports*

Williams, D. M. L., C. Cheepen (1998) "Just Speak Naturally: Designing for Naturalness in Automated Spoken Dialogues". In: Proceedings of ACM SIGCHI'98, Los Angeles.

Talks and seminars:

Klaus Failenschmid attended Eurospeech Conference in Rhodes (22-25 September 1997).

Klaus Failenschmid attended COST Workshop (Speech Technology in the Public Telephone Network: Where are we today?), Rhodes (26-27 September 1997).

Klaus Failenschmid gave Presentation of DISC-relevant issues (Title: Spoken Language Dialogue Systems - Dreams and Reality) at Jetai '97: New Ways of Communicating in Glasgow (12-15 November 1997).

1.6 Information Dissemination

Mailing lists, web sites, flyers

ELSNET has managed the DISC email lists (disc-all@let.uu.nl for the project partners, and disc-ap@let.uu.nl for the AP members).

The DISC website was maintained throughout period 2. In collaboration with the DISC-partners, DISC relevant information was collected and stored on the DISC internal webpages for the project partners (http://www.elsnet.org/disc-internal), the AP pages for the Advisory panel Members (http://www.elsnet.org/disc/ap), and in part made available via the public pages (http://www.elsnet.org/disc).

NIS is presently setting up the DISC-2 website which will exhibit the packaged results of DISC based on concepts developed at the Cambridge workshop in January 1999.

The DISC Flyer has been distributed by mail and at various conferences and other events. A joint DISC and DISC-2 flyer has been produced.

DISC Advisory Panel

The DISC Advisory Panel of experts has increased from 23 members at the end of period 1 to the following 50 members by the end of period 2:

Contactperson	Organisation	Country
Josef Psutka	University of West Bohemia	Czech Republic
Mr Finn Barrett	Tele Danmark Research and Development	Denmark
Lars Bo Larsen	Aalborg University	Denmark
Nancy Underwood	Center for Sprogteknologi (CST)	Denmark
Mr Peter Boda	Nokia Research Centre	Finland
Klaus Failenschmid	SpeechWorks International, Inc.	France
Jean Caelen	CLIPS-IMAG	France
David G. Novick	EURISCO/European Inst. of Cognitive Sciences & Engineering	France
Latifa Taleb	Holistique communication	France
Jan Alexandersson	DFKI GmbH	Germany
Andreas Kellner	Philips Research Laboratories Aachen	Germany
Bernd Ludwig	FORWISS - (Univ. Erlangen)	Germany
Prof. Dr. Jaap Hoepelman	IBM	Germany
Robert Neumann	Institut für Deutsche Sprache LDV	Germany
Dr. Phil. Gerhard Hanrieder	DaimlerChrysler Aerospace AG	Germany
Dr. Maria Aretoulaki	FORWISS-Bavarian Res. Centre for Knowledge-Based Systems	Germany
Dr. Harald Aust	Philips Research Laboratories	Germany
Jens-Uwe Möller	Universität Hamburg	Germany
Ramin Assadollahi	ExB Communication Systems GbRmbH	Germany
Eli Hagen	Deutsche Telekom Berkom	Germany
Dr. Jörg Schütz	IAI	Germany
Dr. Carl Vogel	Trinity College O'Reilly Institute	Ireland
Maurizio Omologo	IRST	Italy
Cristina Delogu	Fondazione Ugo Bordoni	Italy
Dr. Nick Campbell	ATR Interpreting Telecommunications	Japan
Prof. Michael McTear	University of Ulster	Northern Ireland
Mircea Giurgiu	Technical University of Cluj-Napoca	Romania
Maleev G. Oleg	Balti State Universitaty	Russia
Olga Krivnova	Moscow Lomonosov State University	Russia
Daniel Tapias	Telefónica I+D	Spain
Dr Scott McGlashan	Ericsson radio systems AB	Sweden
Elisabeth Maier	Swiss Bank Corporation SBC / IT Camp	Switzerland
M.F. Weegels	Eindhoven University of Technology, IPO	The Netherlands
Dr. Els den Os	KPN Research (Dutch PTT)	The Netherlands
Marc Blasband	Compuleer	The Netherlands

Mike Craig	The Speech Recognition Co Ltd	UK
Tony G. Rose	Canon Research Centre Europe Ltd	UK
Brian Mellor	DERA Malvern	UK
John Gillett	Unisys Ltd	UK
Jim McGhie	Logica UK Ltd.	UK
Dr Susann LuperFoy	Information Extraction and Transport, Inc. (IET)	USA
Carl Uhrik	Berdy Medical Systems	USA
Matthew Lennig	Nuance Communications	USA
Mike Phillips	Applied Language Technologies	USA
John Yuchimiuk	Unisys Corporation	USA
Bruce Pollock	VOCI Corporation	USA
Daryle Jean Gardner-Bonneau	Michigan State University	USA
Dr. Lynette Hirschman	The MITRE Corporation, DARPA Communicator project	USA
Dr. Daythal L. Kendall	Unisys Corporation	USA
Deborah Dahl	Unisys Corporation	USA

The Advisory Panel contributes to DISC's work on SLDSs best practice by commenting on intermediate results, providing access to products, SLDS prototypes or components under development, and/or making the DISC partners aware of practices, theories and tools in current use. An AP report on DISC results is in progress.

In Period 2 of DISC, two workshops were organised with external participation:

- On 8-9 June 1998, a workshop was held in Stuttgart, Germany. The workshop attracted 22 participants (11 project partners, 8 Advisory Panel members, 2 reviewers, 1 Esprit LTR representative).
- On 18-20 November 1998, a workshop was held in Ulm, Germany. This workshop attracted 22 participants (12 project partners, 10 Advisory Panel members).

All members of the Advisory panel were invited to the workshops. DISC covered the costs for accommodation and meals. Transportation costs were paid by the Panel Members themselves.

It turned out that the Advisory Panel was very interested to participate in the workshops. For the future, we suggest to announce DISC workshop earlier in order to allow more Advisory Panel members to make the necessary arrangements for their participation.

Publication

As shown in the list of 34 DISC-related scientific papers and books in Section 4 below, the DISC partners have been active participants in international research on SLDSs and components during the project. Seven of those publications specifically present the DISC agenda (Dybkjær and Bernsen 1997, Elsnews; Bernsen and Dybkjær 1997, ELRA Newsletter; Bernsen and Dybkjær 1997, SALT; Bernsen and Dybkjær 1997, DISC Flyer; Dybkjær et al. 1998, Granada; Bernsen and Dybkjær 1999, DISC and DISC-2 Flyer; Bernsen, Dybkjær and Heid 1999, Eurospeech.

The DISC AP has recommended that DISC make a particular effort to reach out beyond the relatively narrow circle of researchers in speech and natural language processing, to the much

wider communities of developers, deployers and marketeers of SLDSs and components. To this end, DISC (Vocalis) has commissioned a series of presentations of DISC results for the journal Computer Telephony Europe. At this point, four presentations have been written by technical writer David Boothroyd with input from the DISC Consortium.

Also on the point of creating wide awareness of the DISC results, DISC (IMS) has negotiated an agreement with the journal Natural Language Engineering to publish a Special Issue on Best Practice in the Development and Evaluation of Spoken Language Dialogue Systems. The call went out in March 1999 and a strong international Review Committee has been established to review submissions. The deadline for submissions is 20 August 1999.

Bonnie Webber was a DISC invited speaker at the Amstelogue'99 workshop in May 1999. Elsnet organised a DISC stand at the Workshop.

Final DISC Conference

After consultation with the DISC partners, it was decided to postpone the DISC Best Practice Conference which was originally planned for the end of the DISC project, until the end of DISC-2. By that time full and tested documentation and demonstrators will be available.

2. Update of Worldwide State-of-the-Art

2.1 Speech Recognition

The state-of-the-art in speech recognition for spoken language dialog systems consists of real-time, speaker-independent, continuous speech systems with mid-size vocabularies (up to several thousand words). Most of these systems make use of statistical models of speech production. Acoustic models are typically continuous density Hidden Markov Models (HMMs) with Gaussian mixture, of phones in context. Different techniques are used to select contexts, such as decision trees, frequency of occurrence in the training data, clustering or generalized smoothing. Language models for the recognizer are ususally N-gram or class N-grams, where the statistics for these models are estimated on the language model training material. Search is based on frame synchronous dynamic programming, and beam-search pruning techniques are used to reduce the search space. Multiple decoding passes may be carried out so as to allow more accurate models to be used in later passes thus improving performance while minimalizing the additional computation time.

Most speech recognizers are written in C or C++ and can run on standard platforms without special hardware. Only signal capture and communication are particularly device dependent.

References

"Let's Talk", *Business Week,* Feb 23, 1998.

Gibbon Dafydd, Roger Moore and Richard Winski (1997). *Handbook of Standards and Resources for Spoken Language Systems.* Mouton de Gruyter. Berlin, New York. 1997.

Steve Young, "A Review of Large-Vocabulary Continuous-Speech Recognition," *IEEE Signal Processing Magazine,* Sept 1996, pp. 45-57.

2.2 Speech Generation

Many different types of speech synthesis systems exist today and are commercial products. They range in complexity from waveform coded spoken messages to full text-to-speech systems containing many languages. The choice between different systems will be decided by the intended use. Coded speech gives a natural speech quality while the number of messages is very restricted. Text-to-speech systems are much more flexible, but the quality of the speech varies between systems. Today, most systems demonstrate a high degree of comprehension while the quality is often perceived as unnatural.

Speech synthesis systems can normally be split up into two parts. One part contains text-to-speech rules and the other part is the speech synthesiser that produces the sound. Two main methods of speech synthesisers exist, concatenative synthesis where shorter or longer speech segments are concatenated according to rules, and formant synthesis where a speech production model consisting of sources and filters produce the speech. Speech synthesis is sometimes combined with a talking head where the mouth movements are synchronised with the acoustic speech signal to enhance comprehension. A third synthesis method, articulatory synthesis, is still only on the research stage. Text-to-speech rules in combination with lexicons decide the pronunciation of the speech both on the segment and on the sentence level. The

text-to-speech rules used to be derived from linguistic knowledge about the language in question but in later years statistical methods have more and more come into use for rule generation. Likewise, with the growing size of computer memory and faster computers an increasing proportion of the word pronunciation is derived from lexicon transcripts.

References

http://www.speech.kth.se/info/ext_speech.html/

http://www.speech.cs.cmu.edu/comp.speech/FAQ.Packages.html/

Carlson R., Granström B., Speech Synthesis, in W Hardcastle & J Laver (editors): *The Handbook of Phonetic Sciences,* Blackwell Publishers Ltd, Oxford 1997, pp.768-788.

Dutoit, T.: *An Introduction to Text-To-Speech Synthesis,* Kluwer Academic Publishers, Dordrecht, 1997.

van Santen, J., Sproat, R., Olive, J., Hirschberg, J. (Eds.) Progress in Speech Synthesis, Springer, New York, 1997.

Syrdal, A., R. Bennett, S. Greenspan (Eds.) *Applied Speech Technology,* CRC Press, Boca Raton, 1995.

2.3 Natural Language Understanding and Generation

In the first yearly report, we had pointed to the following new developments in evaluation.

Polifroni et al. 1998 describe evaluation methods for the JUPITER system, a weather forecast SLDS for Northern America. The evaluation of the NL understanding component is very similar to the "concept accuracy measure" used in the evaluation of the DaimlerBenz SLDS. JUPITER produces semantic frame representations of user queries. Evaluation of parser coverage is simply done by counting the sentences from both user queries and weather report sources (used for knowledge acquisition in JUPITER) which are analysable.

Evaluation of NL understanding is based on semantic representations and on a comparison of representations generated for new sentences with known-to-be-correct representations of analogous sentences from previous sessions. Insertions, deletions and substitutions are flagged, as in the work of DaimlerBenz (see summary report D-1.4), and counted. The procedures are automatic and use a continuous flow of user queries. Generation quality can only be assessed on the basis of human judgement.

These are still relevant, but no major new aspects have come up in that area.

As far as the development of NL components in SLDSs is concerned, the major new development is that the Java speech grammar format, JSGF, has been defined and described in detail by a group of industrial companies, led by Sun Microsystems, and that it has been accepted as a de-facto standard (or at least a proposal for a standard) by the ECTF.

Details about JSGF have been described in D.2.5. However, it should be noted that the fact that JSGF is accepted by many companies as a proposal for a de facto standard, will certainly have an impact in the short to medium term: we expect that an increasing number of spoken language dialogue systems will have their natural language understanding component designed and implemented in the framework of JSGF.

References:

Joseph Polifroni, Stephanie Seneff, James Glass and Timothy J. Hazen: Evaluation Methodology for a Telephone-Based Conversational System', in: Proceedings of the 1st

International Conference on Language Resources and Evaluation, Granada, 28-30/5/1998: 43-49.

2.4 Dialogue Management

Dialogue managers have so far typically been designed as part of an entire spoken language dialogue system (SLDS). Focus has rarely been on dialogue management for its own sake rather than on dialogue management as one of several components which have to work together to form an SLDS. In some cases, focus has actually not been on dialogue management at all. The dialogue manager was built only because it is a necessary part of an SLDS and hence a necessary evil if one wants to study, e.g., speech recognition in an SLDS environment. However, there are now ongoing efforts world-wide to develop dialogue managers which are more general-purpose and which may fairly easily be adapted to a new task domain, such as the Daimler-Benz dialogue manager.

There are not many tools available in support of building dialogue managers. A few dedicated languages exist. Common to most of these languages is that they are event-based and have a range of primitive operations that support the speech and language layers. It remains, however, an open question to which extent primitives of any of the languages just mentioned, scale up beyond relatively simple, well-structured tasks. If the language primitives needed are not provided by a specialised dialogue specification language, it may often be preferable to use a general programming language, such as Lisp, Prolog or Java, which usually also means that the resulting software will be easier to port.

2.5 Human Factors

Human factors cover all aspects of the interactive system design which are related to the end-user's abilities (perceptual, cognitive and motor), experience (system specific, domain specific and common sense), goals (both interactional and transactional) and organisational/cultural context (del Gado and Neilson, 1996). Whilst the remit of this field is broad, the theoretical and practical work tends to occupy a variety of small niches with few unifying approaches defining interactive system design on all of the dimensions noted.

The analysis with respect to Human Factors used a number of exemplar systems to evaluate current design practice against a best practice framework. These were Vocalis Operetta, Waxholm, the Danish Dialogue System and Verbmobil.

As well as specific areas of research and practical experience, the evaluation also addressed the unique requirements for the design life-cycle of interactive systems. This includes prototyping and descriptive methodologies.

This current practice evaluation has provided an overview of the Human Factors-related aspects of commercial and research spoken-dialogue systems. Those areas of particular interest have been elucidated and placed in the context of an engineering life-cycle for interactive system development. In order to provide an indication of the state-of-the-art, both individual aspects and life-cycle best practice were used to evaluate the exemplars. From this analysis the following deficiencies were identified in current practice:

Documentation

- There was little formal documentation at any stage of the design and implementation process except for Danish Dialogue System.

Ethnology

- Little consideration given to organisational effects of systems and how to design for these.
- Few real users were used in evaluations.
- Context analysis was limited. The reason given for this in Waxholm and Danish Dialogue system was that the goal of the project was to research some aspect of the system ,e.g. speech recognition., rather than meet user needs. The Danish Dialogue System was a little bit different in that it focused on Human Factors. However, since the systems carried out evaluations with real subjects, it seems possible that results for the component under study will be confounded by the negative impact of a system ill-fitted to its context.

Help

- In NLP systems there was little explicit help capability. This could be in the form of example dialogue flows or typical utterances.

This work has provided a good understanding of current practice with respect to Human Factors in SLDSs, both research systems and commercial systems.

2.6 System Integration

The exemplars evaluated in this Workpackage were: Verbmobil, Waxholm, Vocalis Operetta and Vocalis VAD. Evaluations were carried out by Vocalis and LIMSI.

A common feature of SLDSs is that they are highly complex systems which make use of a number of more or less distinct functional modules (e.g. speech recogniser, language understanding, speech synthesiser) to achieve the overall goal of providing a service or information. Depending on the application and task domain these functional modules require very different contextual information (e.g. acoustic speech models, dialogue models) and different levels of 'intelligence' (e.g. user models, dialogue history). The complexity of SLDSs stems from the need of exchanging and sharing large amounts of data between functional modules. Not only is it difficult to clearly distinguish tasks for different modules, but also is it difficult to specify the exact flow of data between two or more modules.

In currently available systems, module tasks and data exchanged between modules are defined by the context of the application. SLDS-internal systems integration of the various functional blocks is done such that the complete SLDS performs well in the task and application domain specified. Unfortunately, extending, modifying or adapting such a system can be extremely time and labour extensive. Also, the development of a new SLDS for a different task domain using pre-existing functional modules often involves re-engineering the fundamental concepts that were defined when the initial system was developed.

Results of the current practice evaluation of this aspect showed that although SLDSs are tightly integrated software systems with numerous (semi-) autonomous functional modules, they tend to make use of proprietary standards and protocols. This makes modification and adaptation of the systems to a new target domain time and cost extensive. Furthermore, the systems integration life-cycles for research systems differ from the ones for commercial systems. The individual stages in the life-cycle are identical for the two types of systems, however systems integration for research systems tends to be driven by the need for integration

of existing functional modules. By contrast, systems integration for commercial system tends to be driven by the need of achieving certain functionality as described by the client.

The evaluation and analysis on this aspect has benefited from the grid and life cycle work on other aspects. Systems integration is very much concerned with the whole SLDS from an 'information-exchange' point of view. Therefore, the analysis of relationships between different modules which has been added to the analysis of each individual aspect has provided valuable input to the systems integration evaluation. The importance of this has been recognised at the second workshop in Ulm.

During the current practice evaluation it has emerged that not only is it important to focus on 'SLDS-internal' communication, but also to concentrate on the way in which an SLDS is integrated with other systems such as the telephone system or LAN/WANs. This aspect is an extremely important one especially for commercial systems.

3. Deliverables Overview

3.1 Deliverable Summary Sheet

Milestone M1 (concluding month 12):

Delive-rable	Task	Respon sible	Due mth.	Acc.	Description
D4.1	T4.3	Elsnet	1	R	Proposal for Industrial Advisory Panel in collaboration with the DISC Management Board. **Done.**
D4.2	T4.1	Elsnet	1	P	DISC email list and DISC WWW pages established. Elsnet dissemination plan. **Done.**
D4.3	T4.2	Elsnet	1, 4, 12,18	R/R/ P/P	Four workshops. **4 workshops done in Year 1.**
D1.1	T1.1	IMS	3	P	Report describing the first DISC dialogue engineering best practice model. **Done.**
D1.2	T1.2	LIMSI	10	R	Working paper on speech recognition current practice. **Done.**
D1.3	T1.3	KTH	10	R	Working paper on speech generation current practice. **Done.**
D1.4	T1.4	IMS	10	R	Working paper on language understanding and generation current practice. **Done.**
D1.5	T1.5	NIS	10	R	Working paper on dialogue management current practice. **Done.**
D1.6	T1.6	Vocalis	10	R	Working paper on human factors current practice. **Done.**
D1.7	T1.7	Vocalis	10	R	Working paper on systems integration current practice. **Done.**
MD1.8	T1.8	IMS	11	P	Integrated report on current development and evaluation practice of SLDSs and components and its deficiencies based on the results of T1.1-T1.7 and W2. **Done.**
D2.1	T2.1	LIMSI	8	P	Survey of existing and easily available platforms and development methods for testing and enhancing the performance of Speech Recognition components. **Done**.
D3.1	T3.1	KTH	12	R	Working document on the detailed DISC best practice methodology. **Done.**

D5.1	WP5	NIS	12	P, A	Annual progress report. **Done.**

Milestone M2 (concluding month 18):

Delive-rable	Task	Respon sible	Due mth.	Acc.	Description
D2.2	T2.1	LIMSI	13	R	Guidelines and testing protocols for the development of speech recognition components for SLDSs. **Done.**
D2.3	T2.2	KTH	13	P	A survey of existing methods and tools for the development and evaluation of speech synthesis and speech synthesis quality in SLDSs. **Done.**
D2.4	T2.2	KTH	13	R, S	Software tool for evaluation of speech synthesis components in SLDSs. **Done.**
D2.5	T2.3	IMS	13	R	Survey of concepts, methods, and tools for the acquisition of lexical data in SLDSs. **Done.**
D2.6	T2.3	IMS	13	R, S	Towards guidelines for the acquisition of lexical data in SLDSs. **Done.**
D2.7a	T2.3	NIS	13	P	State-of-the-art survey of existing dialogue management tools. **Done.**
D2.7b	T2.3	NIS	13	P	State-of-the-art survey of existing human factors tools. **Done.**
D2.8	T2.4	NIS	13	R, S	Concepts and a diagnostic methodology for the identification of user-system interaction problems, their typology, severity and remedies. Software tool in support of cooperative system dialogue design. **Done.**
D2.9	T2.4	NIS	13	R, S	SMALTO: Speech Functionality Advisory Tool Specification. Software tool in support of speech functionality decisions in early design. **Done.**
D2.10	T2.4	NIS	N/A	R	Working Paper on Speech Functionality. **Done.**
D3.2	T3.2	LIMSI	17	R	Draft proposal on best practice methods and procedures in speech recognition. **Done.**
D3.3	T3.3	KTH	17	R	Draft proposal on best practice methods and procedures in speech generation. **Done.**
D3.4	T3.4	IMS	17	R	Draft proposal on best practice methods and procedures in language understanding and generation. **Not completed.**
D3.5	T3.5	NIS	17	R	Draft proposal on best practice methods and procedures in dialogue management. **Done.**
D3.6	T3.6	Vocalis	17	R	Draft proposal on best practice methods and procedures in human factors. **Done.**
D3.7	T3.7	Vocalis	17	R	Draft proposal on best practice methods and procedures in systems integration. **Done.**
MD3.8	T3.8a	NIS	18	P	DISC Dialogue Engineering Best Practice Methodology. **Done.**
D3.8b	T3.8b	IMS	18	P	From a description of spoken language dialogue systems to their evaluation and best practice. **Done.**
D3.9	N/A	IAP	15	R	Assessment report on the DISC best practice methodology and toolbox. **Done.**
D3.10	T3.5	NIS	N/A	R	Working Paper on Dialogue Management Evaluation. **Done.**
D4.4	T4.2	Elsnet	18	P	Best practice conference for industry and end-users. **Postponed.**
MD5.2	WP5	NIS	18	P, A	Final report. **Done.**

Figure 1. Milestones and deliverables.

3.2 Deliverable Details Forms

Project No.: 24823 **Acronym:** DISC

DELIVERABLE DETAILS FORM

Deliverable No.: D1.1 **Due date:** 31.8.1998 **Date of finalisation:** 15.9.1998

Short description: The first DISC dialogue engineering best practice model.

Partner responsible: IMS.

Partners who contributed: IMS **with input from all partners**.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP1 in Year 1.

2. Outside the project: -

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Used in publications on DISC in Year 1.

DELIVERABLE DETAILS FORM

Deliverable No.: D1.2 **Due date:** 31.3.1998 **Date of finalisation:** 16.5.1998

Short description: Working paper on speech recognition current practice. Based on exemplars analyses.

Partner responsible: LIMSI.

Partners who contributed: LIMSI, KTH.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D1.3 **Due date:** 31.3.1998 **Date of finalisation:** 15.5.1998

Short description: Working paper on speech generation current practice. Based on exemplars analyses.

Partner responsible: KTH.

Partners who contributed: KTH, LIMSI.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D1.4 **Due date:** 31.3.1998 **Date of finalisation:** 19.5.1998

Short description: Working paper on natural language understanding and generation current practice. Based on exemplars analyses.

Partner responsible: IMS.

Partners who contributed: IMS, LIMSI.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D1.5 **Due date:** 31.3.1998 **Date of finalisation:** 25.5.1998

Short description: Working paper on dialogue management current practice. Based on exemplars analyses.

Partner responsible: MIP.

Partners who contributed: MIP, IMS, Linköping.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D1.6 **Due date:** 31.3.1998 **Date of finalisation:** 18.5.1998

Short description: Working paper on human factors current practice. Based on exemplars analyses.

Partner responsible: Vocalis.

Partners who contributed: Vocalis, MIP.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D1.7 **Due date:** 31.3.1998 **Date of finalisation:** 22.5.1998

Short description: Working paper on systems integration current practice. Based on exemplars analyses.

Partner responsible: Vocalis.

Partners who contributed: Vocalis, LIMSI, all.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D1.8 **Due date:** 30.4.1998 **Date of finalisation:** 25.5.1998

Short description: Current practice in the development and evaluation of spoken language dialogue systems. Based on D1.1-D1.7.

Partner responsible: IMS.

Partners who contributed: IMS, MIP, all.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.1 **Due date:** 30.6.1998 **Date of finalisation:** 15.5.1998

Short description: Survey of platforms and methods for speech recognition.

Partner responsible: LIMSI.

Partners who contributed: LIMSI.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D2.2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.2 **Due date:** 30.6.1998 **Date of finalisation:** 25.2.1999

Short description: Guidelines and testing protocols for the development of speech recognition components for spoken language dialogue systems.

Partner responsible: LIMSI.

Partners who contributed: LIMSI.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into D3.2 on best practice in speech recognition.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.3 **Due date:** 30.6.1998 **Date of finalisation:** 8.4.1999

Short description: A survey of existing methods and tools for development and evaluation of speech synthesis quality in SLDSs.

Partner responsible: KTH.

Partners who contributed: KTH.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D2.4.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.4 **Due date:** 30.6.1998 **Date of finalisation:** 8.4.1999

Short description: Software tool for evaluation of speech synthesis components in SLDSs.

Partner responsible: KTH.

Partners who contributed: KTH.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into D3.3 on best practice in speech generation.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.5 **Due date:** 30.6.1998 **Date of finalisation:** 25.7.1999

Short description: The report contains an overview of approaches, methods and tools for lexical acquisition in Dialogue Systems.

Partner responsible: IMS.

Partners who contributed: IMS.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into D3.4 on best practice in natural language understanding and generation.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.6 **Due date:** 30.6.1998 **Date of finalisation:** 25.7.1999

Short description: This report contains a few recommendations for lexical acquisition, both with respect to functions of acquisition tools, and with respect to relevant properties of the SLDSs for which the acquisition is supposed to take place.

Partner responsible: IMS.

Partners who contributed: IMS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into D3.4 on best practice in natural language understanding and generation.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.7a **Due date:** 30.6.1998 **Date of finalisation:** 9.2.1999

Short description: State-of-the-art survey of dialogue management tools.

Partner responsible: NIS.

Partners who contributed: NIS.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.5.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.7b **Due date:** 30.6.1998 **Date of finalisation:** 15.2.1999

Short description: State-of-the-art survey of existing human factors tools.

Partner responsible: NIS.

Partners who contributed: NIS, Vocalis.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.6.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.8 **Due date:** 30.6.1998 **Date of finalisation:** 9.4.1999

Short description: CODIAL, a tool in support of cooperative dialogue design. Software tool.

Partner responsible: NIS.

Partners who contributed: NIS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into D3.6 on best practice in human factors.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.9 **Due date:** 30.6.1998 **Date of finalisation:** 30.4.1999

Short description: SMALTO: Speech functionality advisory tool specification. Software tool.

Partner responsible: NIS.

Partners who contributed: NIS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into D3.6 on best practice in human factors.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D2.10 **Due date:** Not planned **Date of finalisation:** 22.4.1999

Short description: Working paper on speech functionality.

Partner responsible: NIS.

Partners who contributed: NIS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D2.9.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.1 **Due date:** 31.5.1998 **Date of finalisation:** 1.6.1998

Short description: Working document on the detailed DISC best practice methodology.

Partner responsible: KTH.

Partners who contributed: KTH.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on WP3 in Year 2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.2 **Due date:** 31.10.1998 **Date of finalisation:** 25.2.1999

Short description: Draft proposal on best practice methods and procedures in speech recognition for SLDSs.

Partner responsible: LIMSI.

Partners who contributed: LIMSI.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.8. Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.3 **Due date:** 31.10.1998 **Date of finalisation:** 6.4.1999

Short description: Draft proposal on best practice methods and procedures in speech generation.

Partner responsible: KTH.

Partners who contributed: KTH.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.8. Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.4 **Due date:** 31.10.1998 **Date of finalisation:** not completed.

Short description: Draft proposal on best practice methods and procedures in natural language understanding and generation.

Partner responsible: IMS.

Partners who contributed: IMS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.8. Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.5 **Due date:** 31.10.1998 **Date of finalisation:** 27.8.1999

Short description: Draft proposal on best practice methods and procedures in dialogue management.

Partner responsible: NIS.

Partners who contributed: NIS, IMS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.8. Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.6 **Due date:** 31.10.1998 **Date of finalisation:** 28.4.1999

Short description: Draft proposal on best practice methods and procedures in human factors.

Partner responsible: Vocalis.

Partners who contributed: Vocalis, NIS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.8. Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.7 **Due date:** 31.10.1998 **Date of finalisation:** 28.4.1999

Short description: Draft proposal on best practice methods and procedures in systems integration.

Partner responsible: LIMSI.

Partners who contributed: LIMSI.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.8. Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.8a **Due date:** 30.11.1998 **Date of finalisation:** 31.7.1999.

Short description: DISC dialogue engineering best practice methodology.

Partner responsible: NIS (responsibility taken over from Vocalis).

Partners who contributed: NIS, all partners.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.8b **Due date:** 30.11.1998 **Date of finalisation:** 1.5.1999

Short description: From a description of spoken language dialogue systems to their evaluation and best practice.

Partner responsible: IMS.

Partners who contributed: IMS.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.9 **Due date:** 30.11.1998 **Date of finalisation:** 07.1999

Short description: Assessment report on the DISC best practice methodology and toolbox.

Partner responsible: Advisory Panel members.

Partners who contributed: Advisory Panel members.

Made available to: Public.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

DELIVERABLE DETAILS FORM

Deliverable No.: D3.10 **Due date:** 31.10.1998 **Date of finalisation:** 28.4.1999

Short description: Working paper on dialogue management evaluation.

Partner responsible: NIS.

Partners who contributed: NIS.

Made available to: Restricted.

Description of further use, including exploitation of the deliverable results:

1. Inside the project: Basis for work on D3.8. Feeds into DISC-2.

2. Outside the project: Discussion with APs.

Impact of the deliverable (publication, product, patent, contribution to standard, exhibition, technology transfer, etc.): Publications, technology transfer.

4. DISC-related Publications

Bernsen, N.O.: Language and speech technologies in natural interactive systems. *Elsnews* 6, 3, 1997, 4-5.

Bernsen, N.O.: Towards a tool for predicting speech functionality. *Speech Communication* 23, 1997, 181-210.

Bernsen, N.O., Dybkjær, H. and Dybkjær, L.: What should your speech system say to its users, and how? Guidelines for the design of spoken language dialogue systems. *IEEE Computer,* Vol. 30, No. 12, December 1997, 25-31.

Bernsen, N.O., Dybkjær, H. and Dybkjær, L.: *Designing Interactive Speech Systems. From First Ideas to User Testing.* Springer Verlag 1998.

Bernsen, N.O. and Dybkjær, L.: The DISC project. *ELRA Newsletter* 2, 2, 1997, 6-8.

Bernsen, N.O. and Dybkjær, L.: The DISC Concerted Action. In R. Gaizauskas (Ed.): *Proceedings of the SALT Club Workshop on Evaluation in Speech and Language Technology,* Sheffield, June 1997, 35-42.

Bernsen, N.O. and Dybkjær, L.: DISC. Spoken Language Dialogue Systems and Components. Best practice in development and evaluation. Esprit Long-Term Research Concerted Action No. 24823. 1 June 1997 - 30 November 1998. Flyer on the DISC project. September 1997.

Bernsen, N.O. and Dybkjær, L.: DISC and DISC-2. Spoken Language Dialogue Systems and Components. Best practice in development and evaluation. Esprit Long-Term Research Concerted Actions No. 24823 and 29597. 1 June 1997 - 31 December 1999. Flyer on the DISC and DISC-2 projects. February 1999.

Bernsen, N. O. and Dybkjær, L.: Is speech the right thing for your application? *Proceedings of the International Conference for Spoken Language Processing, ICSLP'98,* Sydney. Sydney: Australian Speech Science and Technology Association 1998, 3209-3212.

Bernsen, N. O. and Dybkjær, L.: Speech in multimodal systems. Paper to appear in the *Proceedings of the ESCA Tutorial and Research Workshop on Interactive Dialogue in Multi-Modal Systems,* Irsee, Germany, June 1999.

Bernsen, N. O. and Dybkjær, L.: Evaluation of spoken language dialogue systems. Invited book chapter for MIT Press, 1999.

Bernsen, N. O. and Dybkjær, L.: From single word to natural dialogue. Article to appear in Marvin V. Zelkowitz (Ed.): *Advances in Computers,* Vol. 50. Academic Press, 2000.

Bernsen, N. O., Dybkjær, L. and Dybkjær, H.: User errors in spoken human-machine dialogue. In Maier, E., Mast, M. and LuperFoy, S. (Eds.): *Dialogue Processing in Spoken Language Systems.* Lecture Notes in Artificial Intelligence, Springer Verlag 1997, 14-28.

Bernsen, N. O., Dybkjær, L., Dybkjær, H. and Zinkevicius, V.: Generality and Transferability: Two Issues in Putting a Dialogue Evaluation Tool into Practical Use. In *Proceedings of Eurospeech '97,* Rhodes , September 1997, 1911-1914.

Bernsen, N. O., Dybkjær, L. and Heid, U.: Current practice in the development and evaluation of spoken language dialogue systems. Paper to appear in the *Proceedings of Eurospeech'99.*

Bernsen, N. O. and Luz, S.: SMALTO: Advising interface designers on the use of speech in multimodal systems. Paper to appear in the *Proceedings of the IEEE International Workshop on Multimedia Signal Processing,* Copenhagen 1999.

Dybkjær, L. and Bernsen, N.O.: Concerted Action in pursuit of Best Practice. DISC: development and evaluation for dialogue engineering. *Elsnews* 6, 4, 1997, 6-7.

Dybkjær, L., Bernsen, N.O., Carlson, R., Chase, L., Dahlbäck, N., Failenschmid, K., Heid, U., Heisterkamp, P., Jönsson, A., Kamp, H., Karlsson, I., Kuppevelt, J.v., Lamel, L., Paroubek, P., Williams, D.: The Disc Approach to Spoken Language Systems Development and Evaluation. Proceedings of the *First International Conference on Language Resources and Evaluation,* Granada, 1998.

Dybkjær, L., Bernsen, N. O. and Dybkjær, H.: Generality and objectivity. Central Issues in Putting a Dialogue Evaluation Tool into Practical Use. In J. Hirschberg, C. Kamm and M. Walker (Eds.): *Proceedings of the ACL Workshop on Interactive Spoken Dialog Systems: Bringing Speech and NLP Together in Real Applications.* Madrid, July 1997, 17-24.

Dybkjær, L., Bernsen, N.O. and Dybkjær, H.: Designing Co-operativity in Spoken Human-Machine Dialogues. In K. Varghese and S. Pfleger (Eds.): *Human Comfort and Security of Information Systems. Advanced Interfaces for the Information Society.* Springer Verlag 1997, 104-124.

Dybkjær, L., Bernsen, N. O. and Dybkjær, H.: A methodology for diagnostic evaluation of spoken human-machine dialogue. *International Journal of Human Computer Studies,* 48, 1998, 605-625.

Failenschmid, Klaus (1998) Spoken Dialogue System Design - The Influence of the Organisational Context on the Design Process. Proceedings of the 4th IEEE Workshop on Interactive Voice Technology for Telecommunications Applications. Torino September 1998.

Karlsson, Inger and Thornton, Simon: Really natural Speech Generation poses some tough challenges. Computer Telephony Europe 1999.

Lamel, L.: Spoken Language Dialogue System Development and Evaluation at LIMSI. Proc. 1998 International Symposium on Spoken Dialogue, Sydney, Australia, Nov 30, 1998 (invited keynote).

Mariani, Joseph and Lamel, Lori: An Overview of EU Programs Related to Conversational/Interactive Systems. Proc. DARPA Broadcast News Transcription and Understanding Workshop, Landsdowne, VA, Feb 1998.

Thornton, Simon and Dybkjær, Laila: Designing voice-based interfaces: the Human Factor. Computer Telephony Europe, April 1999.

Thornton, Simon: Speak, and you will be understood. Computer Telephony Europe, May 1999.

Thornton, Simon: Putting it all together: system integration is a key to designing spoken language dialogue systems. Communications MEA July-August 1999, and Computer Telephony Europe, July 1999.

van Kuppevelt, J.: Directionality in Discourse: Prominence differences in subordination relations", in: Journal of semantics 13 (1997): 361-393.

van Kuppevelt, J.: Topic and Comment. In: P.V. Lamarque (Ed.): The concise encyclopedia of the philosophy of language, (Elsevier Science Publishers) 1997.

van Kuppevelt, J.: Context and Inference in Topical Structure Theory. In S. Buvac, L. Iwanska (Eds.): Proceedings of the AAAI Fall Symposium on Context in Knowledge Representation and Natural Language, (Cambridge, MA: MIT Press) 1997.

van Kuppevelt, J.: On the semantic nature of the different types of conversational implicatures", in: R. Cooper, G. Chicoidze (Eds.): Proceedings of the 2nd Tbilisi Symposium on Language, Logic and Computation, to appear.

van Kuppevelt, Jan, Heid, Ulrich, Bernsen, Niels Ole, and Dybkjær, Laila: The DISC Method for Describing, Comparing and Evaluating Spoken Language Dialogue Systems. Submitted (accepted with changes) to IEEE Transactions on Speech and Audio Processing, Special Issue on Language Modeling and Spoken Language Dialogue Systems, to appear 1999.

van Kuppevelt, Jan (co-ordinator), Heid, Ulrich, and Kamp, Hans (guest eds.), Preparations for a special issue on Best Practice in Spoken Language Dialogue Systems Engineering, to be published by the journal Natural language Engineering (CUP) in the beginning of 2000. An international Call for Papers has been sent out in May 1999.

Williams, D. M. L., C. Cheepen: Just Speak Naturally: Designing for Naturalness in Automated Spoken Dialogues", In: Proceedings of ACM SIGCHI'98, Los Angeles, 1998.

Williams, D. M. L., C. Cheepen, N. Gilbert: Designing for Naturalness in Automated Spoken-Dased Dialogues: 'All you gotta do is act naturally'. Submitted to the 'Interacting with Computers Journal', 1998.

5. Aggregated information on resources used

	NIS	LIMSI	IMS	KTH	Vocalis	Daimler-Chrysler	Elsnet
WP1							
Planned PM	8.25	7.5 +1 from KTH	11	8.5 -1 to LIMSI	9	0.25	
PM used year 1	8	11	10.5	8	3	0.25	
PM used year 2	0.25	0	0.5	0	0	0	
WP2							
Planned PM	9	7 -1 to IMS	7 +1 from LIMSI	6			
PM used year 1	5	5	3.25	4			
PM used year 2	4	4.5	4.75	2			
WP3							
Planned PM	6.75	6.5	6	6.5	9	0.75	
PM used year 1	0.25	0	0.75	0.5	0	0	
PM used year 2	6.50	8.5	5.25	3.3 + 5 uncosted	4.5	0.75	
WP4							
Planned PM							4
PM used year 1							2.2
PM used year 2							3
WP5							
Planned PM	4.5						
PM used year 1	3						

PM used year 2	1.5						