

Spoken Language Processing in the Framework of Human-Machine Communication at LIMSI

J. Mariani

LIMSI-CNRS, BP 133
91403 Orsay cedex, FRANCE
mariani@frlim51.bitnet or mariani@limsi.fr

ABSTRACT

The paper provides an overview of the research conducted at LIMSI in the field of speech processing, but also in the related areas of Human-Machine Communication, including Natural Language Processing, Non Verbal and Multimodal Communication. Also presented are the commercial applications of some of the research projects. When applicable, the discussion is placed in the framework of international collaborations.

INTRODUCTION

The initial work on Text-to-Speech synthesis and Word-Based recognition resulted in the development of software and hardware (including an Asic) systems which have since been marketed, and used in different applications such as cockpit design. Some of the problems encountered in their practical use are discussed. The development of a speaker verification system presently under test field trials is also presented.

Our speech research program is now oriented along two main axes: Voice Dictation and Spoken Dialog. The successive steps of the voice dictation project are presented, and the difficulties specific to the French language in this task are highlighted. For the dialog project, the integration of linguistic and pragmatic information with continuous speech recognition in the context of air controller training is described.

While this paper focuses on our speech processing projects, other efforts in related areas, such as character recognition and computer vision with algorithms derived from those developed for speech processing are mentioned. Speech processing is placed in the general framework of Human-Machine Communication. This is reflected in the structure of the laboratory, by the relationship of the *Speech Communication* group with the two other components of the *Human-Machine Communication* department: the *Language & Cognition* and *Non-Verbal Communication* groups. When applicable, the projects of the laboratory are situated in the general context of European/international cooperative actions. Perspectives on the directions of research are given.

LIMSI & CNRS

The LIMSI laboratory is a "full" CNRS (French National Research Agency) laboratory. The acronym stands for "Lab-

oratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur" (Laboratory of Informatics for Mechanical, Chemical and Electrical Engineering).

CNRS is the French National Research Agency. It was created in 1939. With 12,000 permanent researchers and 15,000 permanent technical and administrative staff, it is probably the largest research institution in Europe. It has 375 full laboratories, and 1000 academic "associated" laboratories. The 1991 budget of CNRS was 2,000 M\$ (74% for salaries). CNRS has a general director, and is divided into 6 different scientific departments, each having a scientific director. LIMSI is attached to the "Engineering Sciences" department, with a secondary attachment to the "Social and Human Sciences" Department.

There are also 40 advisory committees, forming the "National Committee for Scientific Research", and corresponding to general research areas. Those committees are responsible for evaluating the quality of each laboratory and of each researcher. LIMSI is principally attached to the committee 7 (Information Sciences and Technologies: Computer Science, Control and Signal Processing), but has secondary attachments to sections 10 (Chemical and Mechanical Engineering), 34 (Language Sciences), and may in the near future be attached to section 29 (Cognition and Psychology).

For 1991, the laboratory staff comprised approximately 180 members, with a total budget of 6 M\$ (60% for salaries). With J. Mariani as director, the laboratory is structured in Departments, Groups and Research Topics, each with its own manager. It has two departments: *Human-Machine Communication*, headed by J. Mariani and *Mechanical and Chemical Engineering*, headed by P. Le Quéré, which will not be presented here.

THE HUMAN-MACHINE COMMUNICATION DEPARTMENT

The Human-Machine Communication department has a total of about 100 persons (38 permanent researchers (CNRS and University, including 30 PhDs), 3 technical and administrative staff, 38 PhD students and 36 postdoctoral, contractual, associate and visiting researchers, or Master Thesis students). There are 3 research groups: the *Speech Communication* group, headed by F. Néel, the *Language & Cognition* group, headed by G. Sabah, and the *Non-Verbal Commu-*

nication group, headed by D. Teil. The groups are divided in research topics, each being headed by a researcher. This structure is flexible, and researchers may participate in different research topics (see Appendix).

The activities in Speech Communication in the laboratory were initiated in 1968 by Dr J.S. Liénard. The group itself was created in 1981. The *Natural Language Processing* group was created in 1985, when a group headed by Dr G. Sabah at University Paris VI joined a group already situated at University Paris XI, the location of LIMSI. The *Non-verbal Communication* group was created in Fall 1989, by merging teams in the laboratory already working on 3D modeling, Computer Vision and Robotics.

The *Human-Machine Communication* Department conducts research in closely related areas: Speech, Language, Vision and other means of communication between humans and machines. These areas use common methodologies and, having them studied in the same laboratory, allows for an in-depth study of the different communication modes, which, we believe, is mandatory for the development of multimodal communication systems. The research activities are multidisciplinary, and deal with Computer Science, Linguistics, Cognitive Psychology and Neuro-Sciences, and have both theoretical and applied aspects.

The communication process is studied as a triplet Perception - Production - Cognition: Perception by the machine of speech (that could be enlarged to any sound), of vision (with reading as a subpart), of touch or gesture; Production of speech or text, image synthesis (that could be enlarged to Solid Modeling); And all the cognitive aspects related to dialogue, reasoning, knowledge representation and integration of different communication modes. We try as much as possible to link the studies of production and perception (i.e. the emission and reception of information), as it can be observed in speech recognition and synthesis, text analysis and generation, scene analysis and image synthesis, movement sensing and effort feedback.

The relationship between Language and Image processing is becoming more and more important, with the concept of "Intelligent Images", where the different parts of the image are in agreement with their mutual constraints, and with the constraints of the real physical world (Newton Law of gravity, phenomena which occurs in an explosion ...). Those images need advanced Human-Machine Communication, as the task is complex with commands like "Put the ball on the table and make it bounce back".

This opens the perspective of a true multimodal communication system, including oral, written, visual and gestual communication, with the typical problems of multireference processing (like a voice message accompanied by a gesture). It can even be thought that processing multiple communication modes simultaneously is mandatory for each of the modes, as it is a necessity for knowledge acquisition in training within self-organizing approaches.

We find common methodologies in the different domains: signal processing techniques, statistical or structural coding, Vector or Matrix Quantization in Speech and Vision. Pat-

tern recognition techniques such as Hidden Markov Models, Markov Random Fields, Multi Layer Perceptrons, Boltzmann Machines are used in speech, language and vision. Morphological, Lexical, Grammatical, Syntactic, Semantic or Pragmatic analysis are applicable to written and spoken language, with specifics for speaking or writing.

Finally, these domains also have the commonality of requiring a study of Human Factors (Ergonomics) for design of acceptable systems.

THE SPEECH COMMUNICATION GROUP

The *Speech Communication* group produced very early a Text-to-Speech stand-alone synthesis system in French (Icophone). The Icophone V system was marketed in 1975 (and appeared in *Electronics*, June 1975) by the TITN company, and a single unit of this one cubic meter TTS system was sold (to the Iranian Minister of Education...) at that time. The TTS system used a battery of 44 analog oscillators to produce diphone-based synthesis (using 427 diphones). In 1980, it was replaced by a single board digital version, Icolog, which is marketed by the Vecsys company. The grapheme-to-phoneme conversion algorithm uses about 1,000 rules. It has to take into account the difficult problem of liaisons between words in French.

Another early application was formant synthesis with a first synthesizer designed in 1975 (Icophone VI). A parallel formant synthesizer based on segmental rules is now being developed in the framework of the CEC/Espirit Polyglot project, and should be adapted for 7 different languages within the consortium.

Together with the Text-to-Speech effort, research was conducted on segmentation of continuous speech into phonemes and speech recognition. An analog speech segmentator, using a 32 analog filters bank was presented in 1974 (J.S. Liénard et al., 1974). The speech recognition work addressed both analytical phoneme recognition, in the design of a "Speech Understanding System" comparable to those developed in the ARPA-SUR project (J. Mariani et al., 1978), and word-based recognition.

The first approach was also aimed at designing a "phoneme vocoder", with a 50 b/s rate. The experiments conducted in this project showed that a phoneme recognition rate of at least 85%, with no major recognition errors (like vowel/consonant substitutions), was necessary in order to transmit a message that can be understood by the human listener (J.S. Liénard et al., 1977).

The word-based approach took advantage of the idea that the stationary parts of the signal convey less, and more variable information than transition parts (J.L. Gauvain, J. Mariani, J.S. Liénard, 1983). This led us to non-linear fixed-length compression for isolated word recognition (Moïse system in 1980), and non-linear variable-length compression for connected word recognition (Mozart system in 1982) (J.L. Gauvain, J. Mariani, 1982). Both systems used template matching via dynamic programming. Both systems resulted in single-board products which were marketed by the Vecsys company.

These systems were the first IWR and CWR systems made in France and have been used in several applications, like Voice-activated telephone dialing (Jeumont-Schneider, 1985), or packet sorting (NMPP, 1983). Most of these applications provided the opportunity for experimenting with voice input as a new communication means, and the market stayed limited to a few units. The application where most efforts have been devoted was pilot-plane dialog. In collaboration with the Crouzet company, a program was conducted for the "Research and Technology Agency" (Dret) of the French DoD. A flight with an IWR voice command system with actual commands to the plane was made in July 1982, and was reported as the first voice-controlled flight. The conclusions of the flight trials specified the need of continuous speech recognition, the necessity of a stable level of performances, whoever the pilot, and that voice input was especially interesting in critical conditions (high Gs, stress), which unfortunately correspond to adverse environments which tend to lower the recognition rates.

The computing power needed for the Dynamic Programming algorithm led us to develop an Asic chip. The MuPCD (Microprocessor for Dynamic Programming) was developed at Limsi, together with the Bull company on a contract of the Ministry of Telecommunications. It was available in 1989 (G. Quénot et al., 1989). It has 120,000 transistors in Cmos 2 micron technology, and a power of 70 MOPS. It allows for the recognition of up to 5,000 isolated words, or 300 connected words. This chip is used in the last generation of Vecsys Datavox recognition systems.

The work on Language modeling was also an early project at LIMSIS. It started with a set of experiments on the problem of phoneme-to-grapheme conversion in French. A casual 9 phoneme string can generate more than 32,000 possible combinations of segmentation into words, and spelling of those words. In many cases, the absence of pronunciation of the mark of the plural (-s at the end of the nouns, -nt at the end of verbs), generates many of those homophones. First, a simple heuristic was tried with a 20,000 word lexicon which segmented the phoneme string into the smallest number of words. It gave good results for the segmentation task, but also demonstrated the necessity of using a language model to improve the quality. This resulted in a collaboration with researchers using stochastic language modeling based on grammatical categories for document retrieval, and results were reported in 1979 on phoneme-to-grapheme conversion with a 270,000 word full-form lexicon (A. Andreewski et al., 1979). This approach was also applied to stenotype-to-grapheme conversion.

Another area of research is speaker verification (J. Mariani et al., 1983). The algorithm used for word-based recognition was adapted to speaker verification, with dynamic adaptation of the reference templates of the speaker to their daily variations. The *Sesame* system was tested "live" at the *Machines Parlantes* exhibition in 1985, and had an impostor acceptance of 4 per 1000 obtained with informal test conditions. The system is currently in use in everyday operational conditions as the entry system at LIMSIS by about 100 users

since 1987.

PRESENT PROJECTS

In the Speech Communication group, work is now conducted around two main projects: the Dictation project, and the Dialog project.

In the dictation project, several steps have been taken in the design of a Voice-Activated typewriter (VAT) in French since the beginning of the project 10 years ago. Continuing the study on phoneme-to-grapheme conversion, for continuous, error-free phonemic strings, using a large vocabulary and a natural language syntax, LIMSIS participated in the ESPRIT project 860 "Linguistic Analysis of the European Languages". In this framework, the approach for language modeling developed at LIMSIS has been extended to 7 European languages. The link between the language model and the acoustic recognition was made, and resulted in a complete system (*Hamlet*), for a limited vocabulary (2,000 words), pronounced in isolation, and then to a 5,000 word VAT system taking advantage of the existence of the specialized MuPCD DTW chip. The complete system was demonstrated in Spring 1988. Now, work is being conducted within the ESPRIT Polyglot project, with the goal of designing speech-to-text and text-to-speech systems for the 7 languages. In this framework, the methods first developed at Olivetti for dictation in isolated mode is adapted to French, and other methods are being developed, based on discrete & tied-mixture HMMs, and TDNNs & TDNNs-HMMs combinations, for continuous speech recognition. Comparative tests are being conducted on part of the DARPA Resource Management Database.

We have recorded BREF, a large read-speech corpus for French, containing over 36 GBytes of speech data from 120 speakers. The text materials were selected verbatim from the French newspaper *Le Monde*, so as to provide a large vocabulary (over 20,000 words) and a wide range of phonetic environments. Separate text materials, with similar distributional properties, were selected for training, development test and evaluation purposes. A series of experiments for vocabulary independent phone recognition has been recently carried out using this corpus. A baseline phone accuracy of 60% was obtained with context-independent phone models, and no phone grammar, and a phone accuracy of 68.6% with context dependent phone models and a bigram phone language model (J.L. Gauvain, L. Lamel, this conference).

The dialog project has been explored within the framework of the application of air-controller training. Currently, the training sessions are limited by the availability of the human instructor who plays the role of a pilot. Our goal is to replace him by a spoken dialog system. This allows for more availability of the system, and also to have several voices corresponding to different pilots in the synthesis module. In this project, speech understanding uses speech recognition in conjunction with a representation of the semantic and pragmatic knowledge related to the task. While the language is supposed to follow a pre-defined "phraseology", it is not the case most of the time. The language model is a bigram model based on grammatical categories. The

probabilities for word successions are changed depending on the previous step in the dialog (prediction), and corrections of the recognized sentence can be made, using redundancy within the sentence, and a word confusion matrix. An evaluation test involving 6 speakers, and 5 scripts of an average of 20 sentences each, prediction improved the results by 10%, and correction added an extra 18.5% (the sentence understanding rate improved from 68% to 96.5%) (A. Matrouf et al., 1991). Prior to this work, large "Wizard of Oz" experiments were conducted (D. Luzzati, 1984), and the linguistic analysis of the resulting corpus in a train timetable enquiry system simulation was realized (D. Luzzati, 1987). In general, all the implementations used linguistic analysis of real corpora in order to meet the user's needs. The recording of a spontaneous speech database (Spot) has also been started.

Other research topics concern speech signal processing, with both basic research on wavelet analysis, and development of a real time PC-Based speech analysis tools (the Unice package which is marketed by Vecsys). The study of phonological variations has been pursued on a text ("La Bise et le Soleil") pronounced by several speakers, and will continue with the analysis of BREF. Another area of interest is the use of symbolic coding for improving cochlear implants. Also, apart from the classical Multilayer perceptron or TDNN approaches, an original "Guided Propagation" connectionist model is experimented. In the hardware domain, the use of several MuPCD Asic chips in parallel is now being implemented, and the design of a new chip, taking advantage of improved technology is envisioned.

APPLICATION OF THE ALGORITHMS DEVELOPED FOR SPEECH TO OTHER AREAS

The algorithms developed for speech recognition have been applied in other areas. The DP matching process developed for connected word recognition, with the MuPCD Asic, has been adapted to the problem of optical character recognition. Instead of considering the recognition of individual characters after segmentation (including eventually segmentation errors), the complete line of characters is considered, and segmentation is included in the recognition process (M. Khemakhem, 1987). The algorithm has also been extended successfully in 2 dimensions in Computer Vision for matching similar images, with application to stereovision, and to movement analysis (G. Quénot, 1992). Preliminary studies for gesture recognition (throwing away, taking, holding tight...) using a Data Glove™ have also been conducted. It is expected that the increase of quality obtained by stochastic modeling instead of template matching in speech recognition can also be obtained in the fields of character recognition and computer vision, and we are now considering applying these techniques.

THE LANGUAGE & COGNITION GROUP

The activities of the *Language & Cognition* group have of course many interactions with the *Speech Communication* group. Several common research projects have been, or

are being, conducted, such as the use of Conceptual Graphs to represent semantic information in a speech understanding system, or stochastic modeling of semantic information. Also, the grapheme-to-phoneme conversion software has been used to correct errors. There are also many interactions between the Connectionist Systems and the Connectionist Models research topics in the two groups. The "Time & Space representation" topic has also close relationship with the *Non-verbal Communication* group.

A new activity is now starting in the field of cognitive psychology. A group is being created, integrating the former *Center for Cognitive Psychology* (Cepco), a university Paris XI laboratory within Limsi. It includes researchers in the field of the psychology of reading and text comprehension, visuo-spatial mental representation and cognitive ergonomics.

SPEECH IN THE FRAMEWORK OF MULTIMODAL COMMUNICATION

Speech can be used with other communication modes in order to obtain the more versatile and reliable means for human-machine communication. A study on automated telematic (voice + text) switchboard has been conducted by both the *Speech Communication* and the *Language & Cognition* groups. Speech recognition together with gesture (touch screen) showed that using both together allows for better efficiency and better comfort (D. Teil et al., 1991), with gestural communication being preferable anytime there is a need to give a low level analogous information. Timing and co-reference are the difficult problems to solve with the integrated system.

PERSPECTIVES FOR FUTURE RESEARCH

A more ambitious project is now starting, including computer vision and 3D modeling, natural language and knowledge representation, and speech and gestural communication. This project aims at examining the theoretical problems of model training in the framework of multimodal information, how non-verbal (visual, gestural) information can be used in building a language model, and how linguistic information can help in order to build models of objects.

REFERENCES

- "French Ready terminal that speaks English", *Electronics*, June 26, 1975.
- A. Andreevski, J.P. Binquet, F. Debili, C. Fluhr, Y. Hlal, B. Ponderoux, J.S. Liénard, J. Mariani, "Les dictionnaires en forme complète et leur utilisation dans la transformation lexicale et syntaxique de chaînes phonétiques correctes", 10èmes JEP du GALF, Grenoble, Mai-Juin 1979
- J.L. Gauvain, J. Mariani, "A method for connected word recognition and word spotting on a microprocessor.", Proc. IEEE ICASSP 82. Paris, 3-5 mai 1982.
- J.L. Gauvain, J. Mariani, J.S. Liénard, "On the use of time compression for word-based recognition.", ICASSP 83. Boston, April 14-16, 1983.

J.L. Gauvain, L.F. Lamel, "Speaker-Independent Phone Recognition using BREF", DARPA Speech and Language Workshop, Arden House, February 1992

M. Khemakhem, J.L. Gauvain, J. Rivallier, "Reconnaissance de caractères imprimés par comparaison dynamique.", 6ème Congrès AFCET-INRIA "Reconnaissance des Formes et Intelligence Artificielle". Antibes, 16-20 novembre 1987.

J.S. Liénard, M. Mlouka, J. Mariani, J. Sapaly, "Time segmentation of speech", Speech Communication Seminar, Stockholm, Août 1974

J.S. Liénard, J. Mariani, G. Renard, "Intelligibilité de phrases synthétiques altérées : application à la transmission phonétique de la parole", ICA, Madrid, Juillet 1977

D. Luzzati, "ORSO. Projet pour la constitution et l'étude de dialogues homme-machine.", LIMSI internal report, Septembre 1984.

D. Luzzati, "ALORS : a skimming parser for spontaneous speech processing.", Computer Speech and Language, Vol.2, 1987

J. Mariani, J.S. Liénard, "ESOPE 0 : un programme de compréhension de la parole continue procédant par prédiction-vérification aux niveaux phonétique, lexical et syntaxique", 1er Congrès AFCET "Reconnaissance des formes et Intelligence Artificielle, Chatenay-Malabry, Février 1978

J. Mariani, J.L. Gauvain, J.L. Soury, "Un système de vérification du locuteur", 13èmes JEP du GALF, 28-30 Mai 1984

K. Matrouf, F. Néel, "Use of Upper Level Knowledge to Improve Human-Machine Interaction", Venaco Workshop & ETRW on "The structure of Multimodal Dialogue", Maratea, September 1991

G. Quénot, J.L. Gauvain, J.J. Gangolf, J.J. Mariani, "A Dynamic Programming Processor for Speech Recognition", IEEE Journal of Solid-State Circuits, Vol. 24, N. 2, Avril 1989

G. Quénot, "The "Orthogonal Algorithm" for optical flow detection using Dynamic Programming", IEEE ICASSP'92, San Francisco, March 1992

D. Teil, Y. Bellik, "Multimodal Dialogue interface on a workstation", Venaco Workshop & ETRW on "The structure of Multimodal Dialogue", Maratea, September 1991

APPENDIX

Speech Communication Group (Head: F. Néel)

17 permanent researchers, 17 Phd thesis students

6 Research topics:

Speech Analysis and Synthesis: This topic includes Short term analysis, Wavelets, Instantaneous frequency analysis, Speech signal editor (Unice product marketed by Vecsys) and coding (SNCF contract), Recognition in noisy environments (Dret (MoD) contract), Text-to-speech synthesis, High-quality synthesis, Multivoice and multidialect synthesis (cooperation with Montreal University (MoFA), CEC-

Esprit Polyglot project), Models in Prosodics and diagnostic (paralinguistic) aspects.

Assessment and Variability: In this topic, Recording of large vocabulary recognition evaluation data base (Gdr-Prc CHM "Bref" project) and spontaneous speech database (Spot), Standardization of speech recognition systems assessment (contribution to Afnor), Use of phonotactic constraints, Grapheme-to-phoneme conversion with phonological variations, Regrouping of wavelets, improvement of cochlear implants.

Recognition: Word-based recognition (hardware and software) (Vecsys Datavox product, Sextant-Dret contract), Syllable-based recognition, Diphone-based recognition, Large vocabulary recognition (10,000 words), Continuous and discrete HMMs, Custom Dynamic Programming VLSI (DGT-DGA/ Bull (VTI) contract), Application to printed character recognition, Speaker-independent recognition, Vocabulary-independent recognition, Discriminant recognition, Speaker adaptation, Recognition in adverse environments (DRET/Sextant contract), Speaker verification (MRT contract with Fichet-Bauché and Vecsys), Evaluation of speech recognition systems (participation in CEC/Esprit-SAM project).

Dialog Structures: Models for dialogue structures (Gdr/PRC CHM programme), Speaker models, Task-oriented oral dialogue system for air controller training (Steria/Sextant Avionique/Vecsys contract for CéNA), Linguistic study of man-machine dialogues, Selective syntax parser, Multimodal dialogue (Pilot's assistant application (Dret / Sextant) and Simulation of a telephone operator.

Spoken language modeling: Linguistic Models (Esprit Polyglot project), Phoneme-to-grapheme conversion (Esprit 291/860 project), Stenotype-to-grapheme conversion (Cecet-Systex project), Continuous speech automatic phonetic labelling and recognition through learning (Inserm collaboration), Morphosyntactic analysis, Automatic syntactic classification, Use of linguistic models for handwritten character recognition.

Connectionist systems: Guided Propagation Models (CEC-Stimulation Brain programme, Dret (MoD) contract), Application to noisy speech analysis (cocktail party effect), to continuous speech recognition (contract with French Philips Research Labs (LEP)) and to the modeling of reading activity, Back Propagation Models, Application to character recognition and grapheme-to-phoneme conversion, Feature Maps, Parallel architectures, Integration of the connectionist approach and of the symbolic approach (AI) into the same formalism, Perception-to-Action modeling (Dret contract).

Langage & Cognition Group (Head: G. Sabah)

14 Permanent Researchers, 13 PhD Thesis students

8 Research topics:

Automatic analysis of sentences and texts: Study and implementation of a general architecture for automatic language processing (Distributed AI: communicating multi-expert system structure); Application to the automatic con-

struction of internal representations, Trope and anaphora processing, Semantic flexibility and context influence, Pragmatics (CEC/Esprit PLUS project).

Written Dialog: Implementing a real dialogue in question-answer processes, Speaker modeling, Direct and indirect speech act processing, Adapting an answer to the speaker, Application to computer-aided education (chess playing), and to documentation database query (MRT contract with the Resoudre company).

Flexibility: Elaborated lexical search, Justification or opposition to presupposed concepts (question-answer systems), Interpretation of unforeseen events (orthographic, syntactic errors... (Mannesman-Tally prize)).

Generation: Development of a Computer-Aided language training system (SWIM (See What I Mean)), Development of a multilingual sentence generator: Arabic-French-Russian (Collaboration between France and Bulgaria), Development of a text generator which produces scene descriptions according to psycholinguistic principles.

Learning: Automatic parsing rules creation from examples, Syntactic and semantic aspects, Automatic pragmatic knowledge learning from texts using frames, Fine grain parallel architecture (connection machine) for knowledge representation.

Connectionist Models: Neural networks for Natural Language Processing, Disambiguation and syntactic structure, Relationship with symbolic approaches, Epistemology of connectionism.

Semantic representations: Study of the most adequate lexical representations for automatic processing, Relationship between semantic and connectionist networks, Text semantics.

Time and space representation: Modeling time and space in text analysis, Temporal Logic, Use of generalized intervals, Parallel complexity of temporal constraint networks.

Non-Verbal Communication Group (Head: D. Teil)

7 Permanent Researchers, 8 PhD Thesis students

6 Research Topics:

Tri-dimensional Modeling: 3D Modeling software "Sculptor", Steady and animated image synthesis. Multiparametric varieties, Application to acquisition and modeling for life and earth sciences and for cartoons; Human Factor study of interactivity in graphic interfaces.

Computer Vision: Stereo Vision, Neural and Symbolic learning in computer vision, Application of Genetic Algorithms to image analysis, Scene analysis (contours and texture). Use of a parallel architecture (Transputers). Application to the analysis of road traffic (Inrets contract).

Real Time Architecture: Medium and fine grain parallel architecture. Use of Object Oriented Languages, Hardware specialised for image synthesis. Representation of multimodal knowledge, learning, decision making. Use of AI tech-

niques; link with mobile robots, and Computer Integrated Manufacturing, Real Time modeling and formulation.

Character recognition and coding: Graphic encoding of characters (Eco system, Anvar licence), Printed character recognition by training and template matching (using the MuPCD VLSI chip).

Gestual and multi-modal Communication: Use of tactile screens, touch analysis, mouvement analysis and synthesis, Effort feedback (3D mouse, Data Glove™). Integration of different communication means, perceptive sensor fusion.

Human Factors: Human Factors for system design, Study of the vocal interface for dialog and voice-activated dictation. Study of multimodal communication. Application to Air Traffic Control, to the Pilot's Associate, to an automated telecommunication switchboard.

Industrial Cooperation: Amd, Bull, Cap-Gemini, Ccett, Crai, Cselt, Daimler-Benz, Erli, Fichet-Bauche, Hewlett-Packard, Jutland Telephone, Logica, Olivetti-Syntax, Philips, Resoudre, Sextant Avionique, Siemens, Steria, Systex, Vecsys.

International Cooperation: Universities of Aalborg, Amsterdam, Athens, Bochum, Cambridge, Copenhagen, Delft, Dublin, Edinburgh, Hamburg, Madrid, McGill, Nijmegen, Patras, Pisa, Rabat, Roskilde, Saarlandes, Sheffield, Stuttgart, Tilburg, UCL London, Umist, Inesc, IUT Luxembourg, IPO-Eindhoven, DRA (RSRE), TNO.

LIMSI is a Managing node in the "Speech and Language" Esprit Basic Research Network of Excellence. This network has the goal to promote the integration of Speech and Natural Language. It has now about 40 nodes around Europe. LIMSI is also a node in the French Coordinated Research Project (Prc/Gdr) on "Human-Machine Communication." The Prc has 4 poles on Speech Communication, Natural Language Processing, Vision and Multimodal Communication.