Jornada de Seguimiento de Proyectos, 2004
Programa Nacional de Tecnologías Informáticas

ALIADO: Speech and language technologies for a personal assistant TIC2002-04447-C02

José B. Mariño Acebal
Project Manager
Main researcher of subproject 2

Horacio Rodríguez Hontoria
Main researcher of subproject 1

Abstract

The present project undertakes the study and the development of spoken and written language technologies for the design of personal assistants with a mobile terminal in a multilingual environment. The trade-off between a wish for a more powerful and a smaller terminal can be solved by means of an oral interface. Thus, attention is paid to the design of this interface and the problems related with the environmental noise and the communication channel. We consider two examples of language centred help that can be provided by the assistant: "question answering" and text or speech machine translation. As project goals we can mention a broadcasting speech recognition system suitable for integration with the translator, a flexible dialog controller and a high quality text-to-speech converter with reduced memory requirements. Furthermore, methodologies will be developed for question understanding, and retrieval and summarization of information, so that the questions can be answered. Stochastic corpus-based approaches will be considered in order to cope with the translation problem. The technologies developed will be used to implement two showcases.

1 Objectives

This project is centred on the development of speech and language technologies for a personal portable assistant, provided with mobile telephony, which facilitates user access to information and interpersonal communication in a multilingual environment.

Objective 1: Development of a methodology for the design of the speech processing front-end and the acoustic-phonetic models for a speech and speaker recognition system, compatible with

a) various work situations, such as:
    - relatively quiet room (office)
    - vehicle (with problems of background noise and other speakers)
    - meeting room (with problems of microphone distance and other speakers)
    - public place, etc
b) diverse methods of signal capture
    - quality microphone (close-talk, lapel)
    - array of microphones

Objective 2: Development of a multilingual recognition system for continuous speech and large vocabularies, which include

a) The acoustic-phonetic modelling for various Spanish dialectal varieties, including those from Latin America, and the modelling for English, for European romance languages and for Magrebian Arab.
b) The development of tools for estocastic modelling using finite state automatas (x-grams, word classes) that consider semantic or morpho-syntactic information
c) The specification and programming of dynamic programming algorithms that carry out the recognizer's decision, considering:
    - The coarticulation between words.
    - The generation of useful information for recognition verification.
    - The obtaining of multiple hypotheses organized in a graph of words.
    - The future integration of techniques currently under study for the treatment of spontaneous speech.

Together with the previous objective, the goal is to obtain a system equivalent to the speech recognition systems of the largest European laboratories.

Objective 3: Development of a text-to-speech conversion system emphasizing the quality of the generated speech and the economy of means. In particular, the goals are to:

a) Obtain prosodic modelling close to natural speech.
b) Develop tools for the efficient management of the corpus used in learning models and the generation of speech.
c) Reduce the size of the necessary corpus to carry out the synthesis of the speaker's voice or the incorporation of a new voice.

d) Generate non-neutral speech, including different styles of reading and emotions.

Objective 4: Development of a translation system based on estocastical techniques. Included in this objective are:

a) The obtaining of an estocastical text-to-text translation system.
b) The development of a hybrid speech-to-text translation system.

The translation referred to the Spanish and English languages in a limited application domain (tourist domain).

Objective 5: Construction of a system that provides the most precise answer possible to a question, in written or oral form, asked by the user. The question is formulated in Spainsh and the answer should be found in textual databases previously incorporated into the system or, if not, in Internet. In both cases the document will be available in Spainsh or in English. In the case that the answer is found in English, the translation to Spanish would be carried out automatically.

The particular objectives are the following:

a) Study and solve problems occurring from voice entrance into a answering system.
b) Design and construct a semantic template induction system that reflects the known information about the events that can occur in a domain.
c) Construct a question answering system following conventional architecture, but adapted to the peculiarities of a speech interface.
d) Construction of an automatic summarization system guided by the user's question.

Objective 6: Propose the construction of two demonstrators of the developed technology in the project environment

a) Demo of the technology for "question answering", that includes the dialogue with the user and the eventual translation of texts where the answer can be found.
b) Demo of the speech translation technology that interprets a conversation between two people.

With this it is planned to:

a) Provide a concrete framework for the above technological development, oriented towards the adoption of decisions in a multiple alternative system.
b) Obtain a true test bank of the implemented methodologies.
c) Generate a vehicle for the diffusion of the achieved results.

2 Status of the work

It is important to mention that the project initiation date was December 1st, 2002. Thus, the duration of the project has not yet completed two years. Below is a list of the most relevant activities:

2.1 Objective 1: Robust parameterization of the speech signal

- Noise and reverberation reduction techniques have been tested and compared. For experimental material the Spanish SPEECON database has been used.
- Development and evaluation of a robust beamforming for speech with array of microphones. The principal characteristic is the integration of a modified GSC, which is very effective against directional noises, with a Wiener post-filter, which is very effective against a diffused noise field.
- A voice activity detector (VAD) has been developed based on decision trees, which uses the characteristic resulting from the linear discriminant analysis of the spectral parameters obtained with a frequential filter. The detector was adapted to a room atmosphere using the SPEECON database and it was evaluated with SPEECON's test base as well as with an oral data base from the University of Karlsruhe. The results obtained with this detector during a competitive evaluation as part of the UE's CHIL project were very satisfactory. Likewise, a voice detector was developed that is based on the ADABoost classification technique, which was trained and compared with other classifiers in a vehicle environment, using the Aurora database by SPEECHDATCAR.
- With a larger focus, the classification of acoustic events is being worked on, which includes voice and other types of sound. Classifiers based on Support Vector Machines (SVM) as well as Gaussian Mixture Models (GMM) were developed.
- The Jacobian adaptation was applied in order to adapt sound conditions from the training phase and those from the recognition phase of recognition systems for speech and speakers.

2.2 Objective 2: Spoken language interface

- The performances of our training software for hidden Markov models have been improved. Among the new performances are the trimming of models, efficient quantification techniques, and the tying of probability distributions. Additionally, a technique for the automatic determination of

the optimal assembly of multidialectal acoustic models for an automatic speech recognition system.

- A multidialectal recognition system was developed for Latin-American Spanish based on the previous technique. The recognition precision for each dialect increased by nearly one per cent, decreasing the number of models by 36% with regards to the sum of models necessary for each dialect.
- The multilingual study was extended to English, German, and the languages spoken in Morocco (colloquial and standard Arab and Fench). Tools were prepared to treat the crosslingual cases. Besides, the "Local Codebook Features" method has been developed; this new method automatically provides acoustic phonetic features for sound clustering.
- We have developed a system based on "latent semantic analysis" to generate new trigrams, along with their probability, from trigrams of a given database. We have proved that a statistical dependency exists between trigram appearance frequency and bursty behaviour. This behaviour is related with the fact that the appearance of trigrams follows Levy-Mandelbrot statistics.

2.3 Objetivo 3.- Text-to-speech conversion

- As planned, lexicons developed in the LC-STAR project were incorporated into the system. Additionally, a probabilistic morphological labeller was developed in order to distinguish between words that can have more than one meaning or that are not found in the lexicon. This labeller offers acceptable scores (>95%) at a very limited computational cost, appropriate for real-time text-to-speech conversion
- A new paradigm to model the entonation has been proposed. This paradigm is a top-down model where the representation of pitch contours is led by linguistic characteristics of the text, while satisfying the actual entonation characteristics of an entire corpus. Two models have been particularized, the Fujisaki and the Bezier. Results have been very favourable, both objectively and perceptively.
- New methods were studied that improve the segmentation provided by a reference system based on HMM alignment. We think that the difference in quality between manual and automatic labelling is due to other causes – one of which is not the exact position of the border phonetics- and should be investigated, such as phonetic transcription errors, detection pitch errors, etc. A conclusion that needs to be validated, but of great interest, is that the

human supervision of phonetic borders is not worthwhile, which would greatly decrease production cost.

- Techniques to carry out voice conversion were developed in such a way that an existing synthetic voice can be modified to sound like other voices. Research is centred on converting the parameters that model the vocal tract. The reference system uses a regression based on a join modelling (source voice – target voice) by way of Gaussian mixtures. The developed technique divides the data and applies the reference system on each division. Division is carried out using classification trees based on phonetic characteristics and using the conversion error in a validation corpus as the criteria for tree construction. This methodology provides a significant improvement in quality.
- Voice production models are being explored that allow the manipulation of voice without decrease in quality. This way databases could be smaller and synthesizers could generate richer voices.
- Various synthetic voices were generated for speaker/emotion pairs, following basic emotions defined in MPEG4. These voices were integrated in the selection of segment algorithms for synthesis.

2.4 Objective 4: Speech translation

- A flexible and complete DTD was defined that allows codification in XML of any type of corpus (monolingual, multilingual, text, dialogue, eatc.) with linguistic information.
- Improvements in alignment of parallel corpus were incorporated into our tools. The effect of symmetrization of alignments and the treatment of difficult alignments has been evaluated.
- Linguistic knowledge, for now morpho-syntactic information, was incorporated into the training algorithm of the estocastical translation automata.
- Progress was made in the training of the translation automata. Technical improvements were studied for the generation of “bilingual units (tuples)” after aligning parallel corpus, the problem of embedded words (words that alone have no translation, since they only appear inside tuples) was tackled and the estimation of the probabilities of the translation automata was smoothened out.
- Participation in a world wide campaign for the evaluation of automatic translation systems. The pair of languages chosen for the evaluation was made up of Chinese (source language) and English (target language).

2.5 Objective 5: Question answering

- Participation with the question answering baseline system for the English language in the Question Answering competition of the TREC-2003 (August 4-11, 2003). Participation with a substantially improved version in CLEF-2004 competition for the Spanish language and TREC-2004 for the English language. The first system only treated factual questions, during CLEF-2004 we touched upon definition questions, and during TREC-2004 factual questions, definition and list questions were addressed.
- A special effort has been made with regards to the fine-tuning and enrichment of resources that facilitate the process: names of places, terminology, WordNet, acronyms, etc.
- Although the main focus has been put on the improvement of the question answering system, tasks such as automatic summarization or Web information access have not been ignored.

2.6 Objective 6: Demos

Currently, a first translation demo for the tourism domain is available and implementation of the question answering demo is ahead of schedule. For the latter, the task is the answers to questions about Spanish geography and the speech recognizer for the questions is already available.

3 Main achievements

3.1 PhD Students

List of PhD students that participate in the activities of the project:

Objective 1

Tenko Andrey Alexandrovich, FPI granted by Ministerio de Ciencia y Tecnología, grade: Computer Science Engineer (Dniepropetrovsk National University, Ukranie), date of incorporation: 1/05/2003.

Alberto Abad, FI granted by Generalitat de Catalunya, grade: Telecommunication Engineer, date of incorporation: 9/04/2003.

Pere Pujol Marsal, FPI granted by Ministerio de Ciencia y Tecnología, grade: Telecommunication Engineer, date of incorporation: 1/05/2003

Jan Anguita, ALIADO grant, grade: Telecommunication Engineer, date of incorporation: 1/10/2003.

Objective 2

Frank Diehl, IGSOC granted by Generalitat de Catalunya, grade: "diplom ingenieur" (Universidad de Kaiserslautern, Germany), date of incorporation: 01/01/2003.

Mónica Caballero, ALIADO grant, grade: Telecommunication Engineer, date of incorporation: 12/06/2003

Marta Casar, ALIADO grant, grade: Telecommunication Engineer, date of incorporation: 1/10/2003.

Mireia Farrús, FPU granted by Ministerio de Educación, Cultura y Deportes, grade: Diplome of Physics (UB), date of incorporation: 1/02/2004.

Objective 3

Helenca Duxans, FPI granted by Ministerio de Ciencia y Tecnología, grade: Telecommunication Engeenier, date of incorporation: 1/07/2001.

Pablo Agüero, IGSOC granted by Generalitat de Catalunya, grade: Electronic Engeenier (Universidad Nacional de Mar de la Plata), date of incorporation: 1/01/2003.

Javier Pérez ALIADO grant, grade: Telecommunication Engeenier, date of incorporation: 01/03/2003.

Jordi Adell, UPC grant, grade: Telecommunication Engeenier, date of incorporation: 1/03/2004.

Objective 4

Adrià de Gispert, FI granted by Generalitat de Catalunya, grade: Telecommunication Engeenier, date of incorporation: 1/01/2003.

Josep Maria Crego, UPC grant, grade: Computer Science Engeenier, date of incorporation: 1/10/2003.

Patrick Lambert, ALIADO grant, grade: Diploma of Physics, date of incorporation: 1/12/2002

Objective 5

Daniel Ferrés, UPC grant, grade: Computer Science Engineer, date of incorporation: 1/12/2002.

3.2 Publications

Objective 1:

- D. Moreno, P. Jackson, J. Hernando, M. Russell, "*Improved ASR in noise using harmonic decomposition*". 15th International Congress of Phonetic Sciences(ICPhS´03), Barcelona, August 2003.
- H. Bauerecker., C. Nadeu, J. Padrell, "*On the Advantage of Frequency-Filtering Features for Speech Recognition with Variable Sampling Frequencies. Experiments with SpeechDatCar Databases*", Proc. EUROSPEECH 2003, Geneva, Sept. 2003.
- Abad, C. Nadeu, J. Hernando, J. Padrell, "*Jacobian adaptation based on the frequency-filtered spectral energies*". Proceedings of 8th European Conference on Speech Communication and Technology (EUROSPEECH'03), September 2003.

- P.J.B. Jackson, D. Moreno, M. Russell, J. Hernando, "*Covariation and weighting of harmonically decomposed streams for ASR*". Proceedings of 8th European Conference on Speech Communication and Technology (EUROSPEECH'03), September 2003.
- A. Abad, J. Hernando, "*Speech enhancement and recognition by integrating adaptive beamforming and Wiener filtering*". IEEE Sensor Array and Multichannel Signal Processing Workshop, SAM 2004, Sitges, July 2004
- A. Abad, J. Hernando, "*Integrated adaptive beamforming and wiener filtering for a robust microphone array*". Proceedings of 3rd IEEE Sensor Array and Multichannel Signal Processing Workshop, Sitges (Spain), July 2004.
- J. Anguita & J. Hernando, "*Inter-Phone and Inter-Word Distances for Confusability Prediction in Speech Recognition*". XX Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural, SEPLN. Barcelona, Spain, July 2004.
- A. Abad, J. Hernando, "*Integrated adaptive beamforming and Wiener filtering for a robust microphone array*". Proceedings of International Conference on Spoken Language Processing, ICSLP 2004, Jeju Island, Corea, October 2004.
- P. Pujol, C. Nadeu, D. Macho & J. Padrell, "*Speech Recognition Experiments with the SPEECON Database Using Several Robust Front-Ends*" . Proceedings of the International Conference in Speech and Language Processing (ICSLP'04), Jeju Island, Korea, October 2004.
- J. Anguita, J. Hernando & A. Abad, *"Jacobian Adaptation with Improved Noise Reference for Speaker Verification"*. Proceedings of International Conference on Spoken Language Processing, ICSLP 2004, Jeju Island, Korea, October 2004.
- J., Stephane Peillon, J. Hernando & A. Bramoulle, *"Word Confusability Prediction in Automatic Speech Recognition*". International Conference on Spoken Language Processing, ICSLP'04, Jeju, Korea. October 2004.
- Josep M. Solà, Yin H. Lam, Thomas Kemp. Climent Nadeu, "*Environmental Robust Features for Speech Detection"*. International Conference on Spoken Language Processing, ICSLP'04, Jeju, Korea. October 2004.
- J. Anguita, J. Hernando & A. Abad, "*Improved Jacobian Adaptation for Robust Speaker Verification*". Submitted to IEICE Transactions on Information and Systems.
- P. Pujol, S. Pol, C. Nadeu, A. Hagen, H. Bourlard, "*Comparison and combination of features in a hybrid hmm/mlp and a hmm/gmm speech recognition system*", IEEE Trans. on Speech and Audio Processing, Accepted, 2004.

Objective 2

- José A. Rodríguez Fonollosa, "*Fast vector quantization based on subcodebook selection and its application to speech recognition*", II Jornadas en Tecnologías del Habla. Granada. December 2002.
- J. R. Saeta, J. Hernando, "*Estimación a priori de umbrales dependientes del locutor*", II Congreso Nacional de la Sociedad Española de Acústica Forense (SEAF'2003), Barcelona, April 2003.
- J. R. Saeta, J. Hernando, "*Automatic Estimation of a priori Speaker Dependent Thresholds in Speaker Verification*". 4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA'03), Guildford (UK), June 2003.

- L. Villarejo, N. Castell, J. Hernando, "*Dialogue management in an automatic meteorological information system*". 16th International Conference on Industrial & Engineering Applications of Artificial Intelligence and Expert Systems (IEA/AIE'03), Loughborough (UK), June 2003.
- José A. Rodríguez Fonollosa, "*Nearest-neighbour search algorithm based on subcodebook selection and its application to speech recognition*", Proc. EUROSPEECH 2003, Geneva, Sept. 2003.
- L. Villarejo, J. Hernando, N. Castell, "*VoiceXML in a real automatic meteorological system*". Proceedings of Internet Technologies for Language Applications, Berlin, Germany, October 2003.
- J. R. Saeta, J. Hernando, O. Manso, M. Medina, "*Securing certificate revocation through speaker verification: the CertiVeR project*". Proceedings of Second COST275 Workshop on Biometrics on the Internet: Fundamentals, Advances and Applications, Vigo, Spain, March 2004.
- J. R. Saeta, J. Hernando, "*Applying speaker verification to certificate revocation*". Proceedings of Odyssey 2004, The Speaker and Language Recognition Workshop, Toledo (Spain), June 2004. ISBN: 84-7490-722-5.
- J. R. Saeta, J. Hernando, "*On the use of score pruning in speaker verification for speaker dependent threshold estimation*". Proceedings of Odyssey 2004, The Speaker and Language Recognition Workshop, Toledo (Spain), June 2004.
- F. Diehl, A. Moreno, "*Local Codebook Features for Mono- and Multilingual Acoustic Phonetic Modelling*", Proceedings of Advances in Speech Technology 2004, AST 2004, Maribor, Slovenia.
- A. Zgank et al, "*The COST 278 MASPER initiative – crosslingual speech recognition with large telephone databases*", LREC 2004, Lisbon, Portugal.
- A. Zgank et al., "*Crosslingual transfer of source acoustic models to two different target languages*". ROBUST 2004, COST278 and ISCA Tutorial and Research Workshop, Norwich, Great Britain.
- Marta Casar, José A. R. Fonollosa, "*An approach to a new architecture for continuous speech recognition*". First results.. Talp Research Center. Internal Report. Barcelona, Spain. July 2004.
- M.Caballero, A. Moreno, A. Nogueiras. "*Data Driven Multidialectal Phone Set for Spanish Dialects* ".Proceedings 8th International Conference on Spoken Language Processing , ICSLP 2004 . Jeju Island, Korea . October 2004.
- F. Diehl, A. Moreno, "*Acoustic Phonetic Modelling using Local Codebook Features*". Proceedings of ICSLP 2004, 8th International Conference on Spoken Language Processing, ICSLP 2004, Jeju Island, Korea.
- X. Anguera, J. Hernando, "*Evolutive Speaker Segmentation using a Repository System*". Proceedings ICSLP04, Jeju, Korea (October 2004.
- J. R. Saeta, J. Hernando, "*Model quality evaluation during enrollment*". Proceedings ICSLP04, Jeju, Korea (October 2004.
- José B. Mariño, A. Moreno, A. Nogueiras, "*A first experience on multilingual acoustic modelling of the languages spoken in Morocco*". Proceedings ICSLP04, Jeju, Korea (October 2004).

- M. Farrús, J. Anguita, X. Anguera, J.M. Crego, A. Gispert, J. Hernando, C. Nadeu, "*Els sistemes de reconeixement de veu i traducció automàtica en català: present i futur*". II Congrés d'Enginyeria en Llengua Catalana, Andorra, November 2004.
- X. Anguera, J.Hernando, "*Segmentació de locutor per a l'indexat automàtic de bases de dades multimèdia en català*". II Congrés d'Enginyeria en Llengua Catalana, Andorra, November 2004.
- J. Padrell, A. Abad, J.Hernando, A. Bonafonte, J.B. Mariño, C. Nadeu, J.A.R. Fonollosa, N. Castell, H. Rodríguez, L. Villarejo, "*aTTemps: Un sistema de diàleg oral per donar informació Meteorològica*". II Congrés d'Enginyeria en Llengua Catalana, Andorra, November 2004.

Objective 3

- H. Duxans, A. Bonafonte, "Estimation of GMM in voice Conversation including unaligned data", Proc. EUROSPEECH 2003, Geneva, Sept. 2003 (CDROM).
- D. Escudero, V. Cardeñoso, A. Bonafonte, "*Experimental evaluation of the relevante of prosodia features in Spanish using machine learning techniques*", Proc. EUROSPEECH 2003, Geneva, Sept. 2003 (CDROM).
- Agüero Pablo Daniel, Bonafonte Antonio, "*Intonation Modelling for TTS using a Joint Extraction and Prediction Approach*". 5th ISCA Speech Synthesis Workshop, SSW5 . Pittsburgh, EEUU . June 2004
- Agüero Pablo Daniel, Wimmer Klaus, Bonafonte Antonio, "*Automatic Analysis and Synthesis of Fujisaki's Intonation Model for TTS*". Speech Prosody 2004, SP2004 . Nara, Japan. March 2004
- P-D. Agüero, K. Wimmer, A. Bonafonte, "*Joint Extraction and Prediction of Fujisaki's Intonation Model Parameters*". International Conference on Spoken Language Processing, ICSLP-04, Jeju, Korea, 2004.
- Pablo Daniel Agüero, A. Bonafonte, "*Intonation modelling using a joint extraction and prediction approach*". 11th International WS: Advances in Speech Technology 2004
- Antonio Bonafonte, Pablo Daniel Agüero, "*Phrase Break Prediction Using a Finite State Transducer*". 11th International WS: Advances in Speech Technology 2004
- David Suendermann, A. Bonafonte, H. Hooge, H. Ney, "*Voice Conversion Based on Time Domain VTLN*". 11th International WS: Advances in Speech Technology 2004
- H. Duxans, A. Bonafonte, A. Kain, J. van Santen, "*Including dynamic information in voice conversion systems*", XX Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural , SEPLN 2004 . Barcelona, Spain. July 2004
- H. Duxans, A. Bonafonte, A. Kain, J. van Santen, "*Including dynamic and phonetic information in voice conversion systems*". International Conference on Spoken Language Processing, ICSLP 2004 . Jeju Island, Korea. October 2004
- D. Suendermann, A. Bonafonte, H. Hoege, H. Ney, "*A First Step towards Text-Independent Voice Conversion*". International Conference on Spoken Language Processing, ICSLP-04, Jeju, Korea, 2004.
- Jordi Adell, Antonio Bonafonte, "*Towards Phone Segmentation for Concatenative Speech Synthesis*". 5th ISCA Speech Synthesis Workshop, ISSW . Pittsburgh, USA, pages 139-144 . June 2004.

- I. Esquerra, A. Bonafonte, "*Habla emocional mediante métodos de re-síntesis y selección de unidades*", URSI'2004, Barcelona, September 2004.
- I.Esquerra, J.Adell, P.D.Agüero, A.Bonafonte, H.Duxans, A.Moreno, J.Pérez, D.Sündermann, "*Els talps també parlen. Línies de recerca en síntesi de la parla al Centre TALP* ", II Congrés d'Enginyeria en Llengua Catalana, Andorra, November 2004.

Objective 4

- de Gispert, A. & Mariño, José B, *"Análisis de las relaciones cruzadas en el alineado estadístico para la traducción automática"*, II Jornadas en Tecnología del Habla, Granada (Spain), Dec 2002.
- de Gispert, A. & Mariño, José B, "Experiments in word-ordering and morphological preprocessing for transducer-based statistical machine translation", sent to IEEE ASRU2003 Automatic Speech and Understanding Workshop.
- D. Conejero, J. Giménez, V. Arranz, A. Bonafonte, N. Pascual, N. Castell, A. Moreno, "Lexica and Corpora for Speech-to-Speech translation: a trilingual approach", Proc. EUROSPEECH 2003, Geneva, Sept. 2003 (CDROM).
- A. de Gispert, José B. Mariño, "*Experiments in word-reordering and morphological preprocessing for transducer-based statistical machine translation*", Proc. IEEE ASRU'03,, USA (November 2003), ISBN 0-7803-7981-0.
- Arranz, Castell, Giménez (2003) "*Development of Language Resources for Speech-to-Speech translation*" Proceedings of RANLP'03.
- V. Arranz, N. Castell, J.M. Crego, J. Giménez, A. de Gispert & P. Lambert, "*Bilingual Connections for Trilingual Corpora: An XML Approach*".In Proc. of the LREC 2004, Lisbon, Portugal, May 26-28.
- Patrik Lambert & Núria Castell, "*Alignment of parallel corpora exploiting asymmetrically aligned phrases*". In Proc. of the LREC 2004 Workshop on the Amazing Utility of Parallel and Comparable Corpora, Lisbon, Portugal, May 25.
- Folkert de Vriend, Núria Castell, Jesús Giménez & Giulio Maltese, "*LC-STAR: XML-coded Phonetic Lexica and Bilingual Corpora for Speech-to-Speech Translation*". In Proceedings of the Papillon Workshop on Multilingual Lexical Databases, Grenoble, France. 2004.
- Josep M. Crego, José B. Mariño, A. de Gispert, "*Finite-state-based and Phrase-based Statistical Mahine Translation*". Aceptado en ICSLP04, Jeju, Korea (October 2004)
- de Gispert , José B. Mariño, Josep M. Crego, "*Phrase-based alignment combining corpus cooccurrences and linguistic knowledge*". Proceedings International Workshop on Spoken Language Translation, Kyoto, Japan (September 2004).
- de Gispert, José B. Mariño, "*TALP: Xgram-based Spoken Language Translation System*". Proceedings International Workshop on Spoken Language Translation, Kyoto, Japan (September 2004)
- A. de Gispert , José B. Mariño, Josep M. Crego, "*Linguistic knowledge in statistical phrase-based word alignment*". Remitido para su publicación en "Natural Language Engineering ", Cambridge University Press.
- V. Arranz, E. Comelles, D. Farwell, C. Nadeu.,J. Padrell, A. Febrer, D. Alexander & K. Peterson, "*A Speech-to-Speech Translation System for Catalan, Spanish and*

English". 6th Conference of the Association for Machine Translation in the Americas (AMTA-2004), Washington DC, September-October 2004.

- V. Arranz, E. Comelles, D. Farwell, C. Nadeu.,J. Padrell, A. Febrer, D. Alexander & K. Peterson, "*A Speech-to-Speech Translation System for Catalan, Spanish and English*". Lecture Notes in Computer Science, vol. 3265, Springer-Verlag Berlin Heidelberg.
- V. Arranz, E. Comelles, D. Farwell, C. Nadeu & J. Padrell: "*Sistema de Traducció Oral per al Català,.Castellà i Anglès*". II Congrès d'Enginyeria en Llengua Catalana.(CELC-2004), Andorra, November 2004.
- V. Arranz, N. Castell & J. Giménez: "*Creació de Recursos Lingüístics per a la Traducció Automàtica*". II Congrès d'Enginyeria en Llengua Catalana (CELC-2004), Andorra, November 2004.

Objective 5

- L. Alonso, I. Castellón, S. Climent, M. Fuentes, L. Padró, H. Rodríguez, "*Approaches to Text Summarization: Questions and Answers*". Special Issue "Acceso a Información Multilingüe" (Noviember 2003) of "Revista Iberoamericana de Inteligencia Artificial".
- M. Bertran, M. Gatius, H. Rodriguez, "*FameIr, a Multimedia Information Retrieval Shel*l". Actas de 2ª Jornadas de Tratamiento y Recuperación de la Información, JOTRI-2003, Madrid, September del 2003.
- M. Fuentes, M. Massot, H. Rodriguez, L. Alonso, "*Mixed Approach to Headline Extraction for DUC 2003*". Proceedings of DUC 2003 Workshop on Text Summarization. Edmonton, Canada,. May 2003.
- Tuells, T, Rigau, G., Rodríguez, H., "*Off-line Compilation of Chains for Head-driven Generation with Constraint-based Grammars*" Proceedings del Fourth International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2003). México, February 2003.
- J.L. Vicedo, H. Rodriguez, M. Massot, A. Peñas, "*Los Sistemas de Búsqueda de Respuestas desde una Perspectiva Actual*" Actas de SEPLN-2003, Alcalá de Henares, September 2003.
- Marc Massot, Daniel Ferrés, Horacio Rodríguez, "*QA UdG-UPC System at TREC-12*". Proceedings of the Twelfth Text Retrieval Conference (TREC), Gaithersburg, Maryland, United States, 2003.
- Alicia Ageno, Daniel Ferrés, Edgar González, Samir Kanaan, Horacio Rodríguez, Mihai Surdeanu, Jordi Turmo, "*Talp QA System for Spanish at CLEF-2004*". Proceedings of the 2004 Cross Lingual Evaluation Forum Workshop (CLEF 2004), Bath, United Kingdom, September 2004.
- Laura Alonso, Maria Fuentes, Marc Massot, Horacio Rodríguez, "*Re-using high-quality resources for continued evaluation of automated summarization systems*". Proceedings of the 4th International Conference on Languages Resources and Evaluation (LREC 2004). Lisbon, Portugal, 2004.
- Javier Farreres, Horacio Rodríguez, "*Selecting the Correct English Synset for a Spanish Sense*". Proceedings of the 4th International Conference on Languages Resources and Evaluation (LREC 2004). Lisbon, Portugal, 2004.
- Daniel Ferrés, Marc Massot, Muntsa Padró, Horacio Rodríguez, Jordi Turmo, "*Automatic Building Gazetteers of Co-referring Named Entities*". Proceedings of the

4th International Conference on Languages Resources and Evaluation (LREC 2004). Lisbon, Portugal, 2004.

- Daniel Ferrés, Marc Massot, Muntsa Padró, Horacio Rodríguez, Jordi Turmo, "*Automatic Classification of Geographical Named Entities*". Proceedings of the 4th International Conference on Languages Resources and Evaluation (LREC 2004). Lisbon, Portugal, 2004.
- Jorge Vivaldi, Horacio Rodríguez, "*Automatically Selecting Domain Markers for Terminology Extraction*". Proceedings of the 4th International Conference on Languages Resources and Evaluation (LREC 2004). Lisbon, Portugal, 2004.
- Maria Fuentes, Edgar Gonzàlez & Horacio Rodríguez, "*Resumidor de noticies en català del projecte Hermes*". 2n Congrés d'Enginyeria en Llengua Catalana. (CELC'04). Andorra, 2004.
- Marta Gatius, Manuel Bertran and Horacio Rodríguez, "*Multi-language and Multimedia Information Retrieval from Web Documents*". In 15 th Internacional Workshop on Database and Expert Systems Applications (DEXA), 2004, University of Zaragoza.

3.3 Transfer of technology

Some of the technologies developed in the framework of the current project have been transferred to private companies. The agreements below were carried out with the following companies:

Applied Technologies on Language and Speech S. L. (ATLAS), spin-off of the UPC with whom there are diverse agreements established for commercial exploitation of speech recognition and synthesis technologies developed in the UPC.

SouthWing: development of an embedded front-end for distributed speech recognition

Biometrics Technologies S. L: agreement for the exploitation of speaker identification software.

Servicio Metorológico de Cataluña (SMC): incorporation of oral interface for the forecasting service.

3.4 Participation in international research projects

BIOSEC: Biometrics and Security" (IST-2002-001766), http://www.biosec.org

CHIL: Computers in the Human Interation Loop (FP6-506909), http://chil.server.de, IP project.

COST 275: Biometrics-based Recognition of People over the Internet.

COST 278: Spoken Language Interaction in Communications.

ECESS: European Center of Excellence for Speech Synthesis. (http://www.ecess.org)

FAME: Facilitating Agent in Multiculture Exchange, http://isl.ira.uka.de/fame/

HOPS: Enabling an Intelligent Natural Language Based Hub for the Deployment of Advanced Semantically Enriched Multi-channel Mass-scale Online Public Services (IST-2002-507967) http://www.bcn.es/hops/

LC-STAR: Lexica and Corpora for Speech to Speech Translation Components (IST-2001-32216), http://www.lc-star.com

MASPER initiative: Multilingual and crosslingual speech recognition, http://masper.uni-mb.si

MEANING: Developing Multilingual Web-scale Language Technologies (IST-2001-34460), http://www.lsi.upc.es/~nlp/meaning/meaning.html

ORIENTEL: Multilingual access to interactive communication services for the Mediterranean and the Middle East (IST-2000-28373), http://www.orientel.org

SIMILAR: The European taskforce creating human-machine interfaces SIMILAR to human-human communication (IST-2002-507609), http://www.openinterface.org/

TC-STAR: Technology and Corpora for Speech to Speech Translation (FP6-506738) http://www.tc-star.org, IP project.

3.5 Collaboration with other research groups

On top of collaborations and exchanges that imply the participation in research projects, excellence networks, etc. one can site the following:

Conferences organization: participation in the organizing committees for the following conferences that took place in Barcelona:

- II Congreso Nacional de la Sociedad Española de Acústica Forense, SEAF'2003, April 2003.
- "15th International Congress of Phonetic Sciences" ICPhS'03, August 2003.
- "42nd Annual Meeting of the Association for Computational Linguistics", ACL'04, July 2004.

Participation in the "Forum de les Cultures, 2004". The techniques presented include the multimodal access to multimedia documents in different languages, speech translation, navigation by means of speech and information retrieval, etc. The other groups that have collaborated in the exhibition belong to the University of Karlsruhe (Germany), l'INRIA and the University J. Fourier de Grenoble (France), l'Istituto Trentino di Cultura (Itàlia), Carnegie Mellon University (U.S.A.) and the companies Sony (Germany) and Atlas (Barcelona).

Collaborations are maintained with research groups for the aforementioned technologies:

- University of Alicante: natural language processing, dialog management.
- University of Barcelona: natural language processing, definition of a multilingual corpus of the languages spoken in Spain for translation.
- University of Girona: automatic summarization and question answering.
- University of Maribor (Slovenia): proposals of methods for evaluation of prosodia, comparison of voice synthesis architecture.
- University of Basque Country: definition of a multilingual corpus of the languages spoken in Spain for translation.
- Technical University of Valencia: Comparison of phonetic segmentation methods developed by both universities in an application and through a common corpus; definition of a multilingual corpus of the languages spoken in Spain for translation.
- University Pompeu Fabra: definition of a multilingual corpus of the languages spoken in Spain for translation.
- University of Valladolid: development of common entonation models.
- University of Vigo: definition of a multilingual corpus of the languages spoken in Spain for translation