AGENCE NATIONALE DE LA RECHERCHE ANR	**Programme BLANC**	Réservé à l'organisme gestionnaire du programme N°de dossier : ANR-08-XXXX-00 Date de révision :
	Document scientifique associé	**Edition 2008**

Acronyme/short title	**OTIM**
Titre du projet *(en français)*	**Outils de traitement d'information multimodale**
Titre du projet/Proposal title *(en anglais)*	**Tools for Multimodal Information Processing**

Les pages seront numérotées et l'acronyme du projet devra figurer sur toutes les pages du document en pied de page.
Un sommaire du document est bienvenu

S'il s'agit d'un projet déposé dans le cadre d'un accord de coopération internationale*, préciser avec quelle agence étrangère :
☐ National Natural Science Foundation of China (NSFC)
☐ Japan Society for the Promotion of Science (JSPS)
☐ Japanese Science and Technology Agency (JST)
☐ National Science Council of Taiwan (NSC)

Sommaire

Programme scientifique / Description du projet
Technical and scientific description of the proposal

1.1 Problème posé/Rationale

Multimodality has become one of today's most crucial challenges both for linguistics and computer science, entailing theoretical issues as well as practical ones (verbal interaction description, human-machine dialogues, virtual reality etc…). Understanding interaction processes is one of the main targets of these sciences, and requires taking into account the whole set of modalities and the way they interact.

From a linguistic standpoint, language and speech analysis are based on studies of distinct research fields, such as phonetics, phonemics, syntax, semantics pragmatics or gesture studies. Each of them have been investigated in the past either separately or (in the best cases), in relation with another field that was considered as closely connected (e.g. syntax and semantics, prosody and syntax, etc.). The perspective adopted by modern linguistics is a considerably broader one: even though each domain reveals a certain degree of autonomy, it cannot be accounted for independently from its interactions with the other domains. Accordingly, the study of the interaction between the fields appears to be as important as the study of each distinct field. This is a pre-requisite for an elaboration of a valid theory of language.

However, we have to face gaps of knowledge in this area, concerning empirical data as well as theoretical approaches. No theory of language has yet integrated the fact that information comes from the different linguistic domains. The description of relevant data illustrating prototypical interactions between the distinct domains and modalities (e.g. interactions between speech, gesture and syntax in the description of deictics, or interactions between semantics, pragmatics and prosody in the description of non-declarative utterances, etc.), is crucial in trying to reduce this gap. In order to be performed, such descriptions need to be based on ***annotated corpora***, encompassing all of the distinct types of data previously mentioned.

However, as important as the needs in this area might be, high level multimodal resources and adequate methods in order to construct them are scarce and unequally developed. It should be noted that ***there exists at the present moment no such annotated multimodal corpus in French*** (they are rare and incomplete for the English language). Ongoing projects, including international ones, mainly focus on one modality as a main target, with an alternate modality as an optional complement. Moreover, coding standards in this field remain very partial and do not cover all the needs in terms of multimodal annotation. Last, existing annotation and querying tools are still at a very preliminary level. Our project develops linguistics and technological targets and aims at treating these issues following four distinct steps:

- Designing of a generic multimodal coding format
- Designing of an annotation platform
- Designing and use of a request tool adapted to the new format
- Designing of new annotated resources in the specified format

One of the first issues we have to face is the definition of a coding scheme providing adequate responses to the needs of the various levels encompassed, from phonetics to pragmatics or syntax. While working in the general context of international coding standards, we plan to create a specific coding standard designed to supply proper responses to the specific needs of multimodal annotation, as available solutions in the area do not seem to be totally satisfactory. Once we settle on a coding standard, our next step will be the adaptation of existing analysis tools, and more specifically the adaptation of the tools developed by our research partner teams (signal editors, POS taggers, parsers, etc.) to output productions matching those standards. We plan to use or develop data manipulating and processing tools matching our specific needs. There again, we will base our work on existing tools and adapt them to our needs.

Accordingly, the aim of this project is to gather in a unique platform a resource base consisting in raw or enriched data, for the description of French (and more specifically for spoken French) together with a set of available tools (whether developed in our own labs, or from alternate sites) for annotation, manipulation and query.

In order to achieve these goals, we will rely on the experience and the resources developed by the project partners: a number of audio and video corpora of verbal interactions already exist in our labs. Some of the data have already been enriched of several annotations, whereas others are still a rough material. Our second step should be the automatic or semi-automatic analysis of the data resulting in an enriched corpus of linguistic information sources. In a first stage, this operation will take place on a certain number of samples, fully transcribed and annotated, extracted from the corpus. Then, we propose to create entirely new corpora responding to the needs of the different linguistics fields, and annotated by means of the OTIM platform.

1.2 Contexte et enjeux du projet/*Background, objective, issues and hypothesis*

The description of the spoken form of language is one of the main goals of general linguistic research, but it is also an important issue for applications, which is especially true in the case of human-computer communication or information processing. We have to admit that research in the fields of resources and tools for spoken French are considerably behind in those matters. No such thing as a reasonably sized homogeneous resource for the joint treatment of acoustic and symbolic aspects of Spoken French are yet available. Moreover, current coding standards and tools are not completely fit for multimodal information processing (typically in the case of audio and textual input).

Our project plans to bring along solutions, in terms of both designing resources and designing techniques, and in terms of resource management. The target is the creation of the most important database on Spoken French, including information at the physical level (signal analysis) or the symbolic level as well, while displaying a set of tools facilitating annotation and information retrieval. The first step of the project consists in establishing a precise state of the art on existing resources, projects and tools in the area. Some of the main directions of this review may already be sketched.

1.2.1. Current projects in multimodal linguistics

Western languages have been largely studied in linguistic fields as diverse as syntax, semantics, phonetics etc, first in their formal aspects. The descriptions were then based on written texts or constructed examples which served to show the underlying mechanisms of the language under study. The development of tools which responded to the need of linguists allowed them to begin a collection of audio recordings. Such collections were useful to keep trace of a language at a certain state of its evolution for diachronic studies. They were also the medium to start working on the oral side of languages. Linguists then realized that working on a modality that had been ignored so far, they could improve the description of each language. They then started creating new corpora in controlled situations, e.g. in lab settings. At about the same time, researchers, mostly ethnologists began to come back from distant missions with video clips which offered new perspectives: languages could be studied outside the laboratory environment, and linked to a sociological system. This opened up a large road to multidimensional linguistics which has been at the core of interactional studies. Encompassing multimodality implies introducing several new dimensions into the study of language, dimensions linguists have not been accustomed to deal with: continuous vs. digital phenomena, simultaneity vs. sequential organization of relevant data. One of the consequences of this change in the scope of language science was a strong need for adequate tools in order to be able to spot and to treat such qualitatively new parameters.

Among the earliest approaches, the Palo Alto communication studies brought along both theoretical (*Pragmatics of Human Communication*, 1967) and practical approaches of nonverbal behaviour in relation to speech (studies of filmed interactions). Following up D. Efron's (1954) field work and typology of gestures, P. Ekman and W.V. Friesen introduced a broader typology (1969) and designed an efficient tool for the description and analysis of facial movements: the Facial Action Coding System "F.A.C.S." (1978), more recently adapted as "FEAT" for a computerized use by Kaiser & Wehrle (1992). Whereas J. Haviland has kept to the ethnological perspective, C. Goodwin started to work on videotaped interactions of natural situations such as family dinners, work interactions, etc. His work fell within the scope of conversation analysis, to which he started to add a systematic analysis of the gesture behaviour of the participants. Even more systematic was the work of A. Kendon who contributed much to the field of gesture studies in a linguistic perspective.

Their pioneer work has had much influence on what may be termed mutimodality today: D. McNeill, a psycholinguist began an important and systematic study of the interactions between verbal, prosodic and gesture phenomena. Researchers in the McNeill Gesture Lab obtained encouraging results in their studies of the English language which have triggered more research on other languages and in other fields of study. While various researchers (J. Cosnier, J. Bavelas, I. Poggi.) developed models intertwining both communicative and speaker-utility functions of gestures, McNeill and his followers, although sharing the latter interactive perspective, mainly focused on manual gestures, and their relationship to speech in an integrated system of verbal language and gesture production. Eventually, numbers of findings in the fields of both theoretical and practical aspects of gesture studies were facilitated by advances in the study of Sign Languages, even though sign languages as strongly digitalised systems clearly have distinct functions and their kinship to co-speech gestures is limited and has more to do with shared visuo-spatial substance than with structural similarities. However, researches in the field of Sign Language (Gallaudet studies, Scott Lidell) significantly contributed to advances in the general description and understanding of gestures.

Below are listed a few projects which main aim is to contribute to the enrichment of multimodal studies on languages of the world:

ECHO: sign languages

- Together with the acknowledgement of their status of languages, sign languages have seen in the evolution of multimodal studies the possibility to describe their morphological and syntactic structure. The European Cultural Heritage Online is a European Project which aims at transcribing several sign languages. The videos are transcribed with the ELAN tool developed at the MPI. Both videos and transcriptions are freely available online.

CHILDES: first language acquisition

- CHILDES is a large corpus for the study of language acquisition. It is freely available online and the database contains both audio and video recordings together with their transcription. Somme of the transcriptions are quite rich since they contain some prosodic or gesture information as well. Many languages are represented in this databases. Yet, the annotations which are made with the CLAN tool do not really allow a multimodal treatment of the recordings which transcriptions are rather oriented towards morphology.

CLAPI: interactions in natural settings

- The CLAPI (http://clapi.univ-lyon2.fr/feuilleter.php) database is a large corpus of on-line recordings and annotations much in the spirit of Goodwin's interactional work. It is composed of circa 300 hours of audio recordings of French interactions together with their orthographic transcriptions. Part of the corpus are videos, these however are not available online, the corpus still being under annotation process at the present time. The transcriptions can be queried so as to obtain cross-information on morphological phenomena in relation to interactional structure of the recordings.

NIST/VACE and AMI: meetings

- These two projects are quite similar in their aim which is to help understand the participants' interactions during meetings. The NIST/Vace project (http://www.nist.gov/speech/test_beds/mr_proj/meeting_corpus_1/) is a corpus of online videos together with enriched transcriptions of the interactions. The AMI project is slightly more advanced in that it is composed of circa 100 hours of video recordings entirely annotated. This corpus is a European project which counts 15 members so that several languages are represented.

1.2.2. Multimodal corpora

To come back to traditional linguistic description, several English corpora contain the transcribed and annotated acoustic signal. For example, the Switchboard corpus (see [Godfrey92]) constituted of telephone conversation recordings or the corpus made up within the framework of the MapTask project (see [Anderson91]). The annotation kinds are variable and contain information such as speech turns. No equivalent resource is available on French. Some oral corpora are indeed transcribed, but not annotated. More generally speaking, one is forced to admit that there exist very few corpora of French which annotations contain different information types (and these corpora are particularly poor at the phonetic and morpho-syntactic levels). Although the development of spoken language corpora in France is an old undertaking as testified by the work led at Aix (see [Blanche-Benveniste90], or more recently several projects in the field of interaction analysis (led in particular in the laboratories ICAR and LPL).

Multimodal corpora are much more recent than text and speech ones although video has been used for a long time for observing and manually annotating non-verbal behaviour it is only in the last ten years that several studies using digital video and computer-based annotations have been conducted in a variety of contexts and data type in the past in human-computer interaction (using a real system or a simulated one in a Wizard of Oz protocol (Dahlbäck et al. 1993)) and human-human interaction (laboratory recordings, meetings, television, movies, field studies). Compared to the social sciences experiments described in the above sections, the digital corpus-based studies we summarized in this section aim at producing computational models of multimodal behaviour including details of individual behaviour which are required for the design of human-computer interfaces. A multimodal corpus features several communication modalities of video recordings such as speech, but also hand gesture, facial expression, head, and body movements, whatever the type of interaction – human-human or human-computer communication (Wegener Knudsen et al. 2002; Martin et al. 2004). A multimodal corpus does not only contain video recordings, but is an organized set of files including meta-data files (recording conditions, context), media (video / audio) files, trackers/sensors data, logged files from recognition modules, documentation (annotation guide, protocol description and annotation examples, subject's answers to pre and post-experimental questionnaires).

The elaboration of such corpora requires several time-consuming steps such as (Bakeman et Gottman 1997; Mondada 2005): identifying the goals of the corpus, studying the related work in social sciences, defining the collection protocol, defining a coding scheme, writing an annotation guide, testing the annotation with a few

samples, recruiting annotators, setting the software and hardware architecture (cameras, lights, microphones, network connections), recording the sessions, digitizing the videos, segmenting and formatting these videos, annotating at multiple levels both manually and automatically, validating these annotations (intra- and inter-coder agreement, combination of both), automatic or manual processing of the annotations, computing statistical measures and models. Repetitions of some procedures are required at multiple levels during this process which doubles or triples the amount of time required for the treatment of the initial resource.

Concretely, several computer-based annotation tools have been designed for the annotation of non-verbal behaviour[1] anchored in time and organized multiple layers. to mention some of these tools, one may consider Anvil (Kipp 2004), the Observer (Cadée et al. 2002, [NOLDUS06), Vista (Shi et al. 2004), ELAN (Wittenburg et al. 2006), Nite [NITE03] or EUDICO Linguistic Annotator (DASH) [EUDICO05]. Several annotation tools have been tested and described in (Dybkjær et al. 2001; Dipper et al. 2004; Dybkjær et Bernsen 2004; Dybkjær et Bernsen 2004; Kipp 2004; Loehr et al. 2005). The tools allow the annotating of the video signal at different temporal levels: time-based, segments annotations or more global lapses of time. Different annotation models have been defined for such annotation tools among which the annotation graphs (Bird et Liberman 2001) and the Abstract Corpus Model (Wittenburg et al. 2002), both based on a three level architecture which includes the physical, logical, application levels (Bird et Liberman 2001).

Tools for the management of metadata have also been elaborated (Wittenburg et al. 2002; Wittenburg et al. 2006). They ensure a certain consistency between annotations done by different coders (inter-coder agreement) in order to validate the coding scheme and the annotation protocol at least on a subset of the data. Several automatic image processing software are designed for the tracking of body parts and the automatic processing of multimodal behavior. In terms of the schemes used at an international level for the coding of gesture behaviour, McNeill (McNeill 2005) contributed much to the description of hand gestures coding the gesture form (handedness, shape of the hand using ASL standards, palm and finger orientation, place in gesture space), motion (shape of the motion, place in space, direction), phonological interpretation of the hand shape and motion. Methodological steps are described by S. Duncan2 and (Kipp 2004): segmentation of a gesture unit into different phases, or how to make use of a pre-analysis of video data to define a gesture lexicon within a categorical transcription approach. Lexicon entries are defined via formation features (Kipp 2004) (e.g. a necessary condition to identify the gesture): handedness, hand shape, location, orientation, movement.

Isabella Poggi defined a musical score coding scheme grounded on her model of meanings and goals (Magno Caldognetto et al. 2004). It enables the transcription and classification of signals in five modalities: verbal modality (words and sentences uttered), prosodic-intonational modality (speech rhythm, pauses, intensity, stress, intonation), gesture modality (movements of the hands, arms and shoulders), facial modality (head and eye movements, gaze, smile and other facial expressions), body modality (trunk and leg movements, body posture, orientation and movements in space). Five levels of analysis are annotated: signal description (e.g. hand shape and hand movement), signal type (e.g. gesture type such as deictic), a verbal formulation of meaning description, meaning type (according to the meaning classification, e.g. speaker's belief), function (e.g. repetitive). Meaning description, meaning type and function involve two layers of annotation each: literal and indirect. The multimodal score has been implemented in Anvil (Kipp 2004) and applied to video samples of different types: interviews, TV news, Italian Sign Language, laboratory experiments, commercial spots.

Sequential analysis is applied by (Bakeman et Gottman 1997) for the observation of interaction and social processes. It includes several steps: definition of a coding scheme, recording sequences of behaviour, assessing inter-coder agreement, analyzing the results. When annotating behaviours, a distinction is made between physically versus socially based coding schemes defined as a continuum in (Bakeman et Gottman 1997). Different representation schemes are defined at different levels of abstraction for representing multimodal behavior. Transition matrices (Van Hooff 1982) are for example used by (Kipp 2004) for computing gesture profiles.

More generally, multimodal corpora have been the object of several conferences such as: LREC (workshops on multimodal corpora in 2000, 2002, 2004, 2006), Measuring Behaviour, Intersensory forum, Interacting Bodies 2005, and Gesture Workshops. Current research study the definition of appropriate coding schemes and methodology, the integration of manual and automatic annotations, the definition of computational models of multimodal behavior, the sharing and re-use of existing corpora. The IST-ISLE3 project elaborated a state of the art in terms of resources (Dybkjær et Bernsen 2002; Wegener Knudsen et al. 2002) and coding schemes (Wegener Knudsen et al. 2002; Dybkjær et al. 2003; Dybkjær et Bernsen 2004).

[1] http://www.ldc.upenn.edu/annotation/gesture/
[2] http://mcneilllab.uchicago.edu/pdfs/Coding_Manual.pdf
[3] http://isle.nis.sdu.dk/

1.2.3. Formats and standards

Several international projects proposed standards for linguistic information encoding of annotation forms. Since the end of the Eighties, the Text Encoding Initiative proposed an exhaustive set of markers (regularly updated) to label all kinds of information being able to enrich a text. Most of the standards suggested respect the directives provided by the TEI; in particular the project THESE (Corpus Encoding Standard), or the group EAGLES (Expert Advisory Group on Language Engineering Standards) and its XML evolution called XCES. To be more precise this project provides some encoding specifications for linguistic annotation as well as a data structure for corpora linguistics. The XCES tagging format will be re-used and completed if necessary in our project.

Within the framework of the network of excellence Humaine (2004-2008), automatic and manual annotations of emotions and their form in various modalities will be collected from broadcast video news ([Martin & Al 05, 06]). The MUMIN network ([MUMIN06]) federates research in multimodality in the north countries and proposes an annotation of non-verbal behaviour in broadcast news oriented towards the study of speech turns. [Bird99] also presents a general specification body for multi-level annotations. However, in the field of multimodal corpus analysis, there has not yet been any real standardization initiative. Each coding scheme has been decided on individually in each modality. The most used coding schemes are for example FACS for facial expressions (see [Ekman78]), and structural and functional descriptions of hand gestures (see [Efron41], [McNeill92], [Kipp04]). These schemes are not computerized and are often adapted solely to the needs of the researchers and the annotation tools they use. They are studied in relation to higher annotation level like communicative functions or emotions which belong to separate annotation schemes (see [Pelachaud04]). A state of the art of annotation schemes was proposed in (Knudsen and Al, 2002).

1.2.4. Platforms, projects and tools

From a multimodal point of view, data-processing tools for the collecting and managing of multimodal corpora have been used more recently than for text and speech corpora. Several international initiatives have been born such as the working group “Natural Interaction and Multimodality” (NIMM) a workgroup under the aegis of the project International Standard for Language Engineering (ISLE) started in January 2000 after EAGLES (Expert Advisory Group on Language Engineering Standards) which made a state of the art (existing corpora, coding schemes, assisting tools in the annotation process) and which proposed recommendations. Several other projects have recently been proposed which aimed at accounting for large quantities of linguistic data. They may be grouped into several categories. The first of these categories contains projects which main goals are to determine annotation standards based on the inventory of existing resources and to propose helping tools for the production and query of multi-level corpora. Among present projects or former ones, one may quote some of them for their importance in the field:

- GATE (http://gate.ac.uk/): This project has been designed for the (manual or automatic) annotation of corpora. It is mainly oriented towards the written language and does not really allows natural speech processing (even if orthographic transcriptions of speech are used). The resulting annotations follow the TIPSTER recommendations. They are saved separately from resources, much alike natural speech annotations. Thus they refer to external resources and are consequently not structured in connection to each other.

- MATE (http://mate.nis.sdu.dk/, see [Dybkjaer98]): This project also proposes the specification of a coding scheme as well as a whole tool bench for the annotation of corpora. It is more particularly concerned with natural speech corpora and allows the encoding of information in various linguistic fields such as prosody, syntax, dialogue acts, etc. Our project is from this point of view very similar in its objectives. Several encoding levels are thus proposed separately: transcription, prosody, morphosyntax, coreference, discourse acts. Quite traditionally in this type of annotation, every information type is located in the speech signal through the means of a temporal cue. A recent evolution of the project called NITE (http://www.ltg.ed.ac.uk/NITE/) proposes a certain number of tools in the form of a library as well as a query language.

- ATLAS (http://www.nist.goc/speech/atlas/overview.html, see [Laprun02]): This project proposes a certain number of tools for the production, the edition and the processing of data encoded in the annotation graph format. The underlying idea of this approach is to interpose between the physical level and the implementation a logical level which contains the formal information on the speech signal independently from the encoding and stocking software. At this level structured data is associated with regions predefined in the speech signal. The software then takes as an entry the abstract representation instead of using the data directly, which allows a certain independence from encoding and saving modalities (as well as from the annotation scheme of a given level). Here again, the anchoring process, even if it is based on a supplementary abstraction level, is achieved through a reference to the speech signal.

The second category of projects contains the ones which principal aim is the development of annotated resources. At the European level the most long standing project is certainly the Map Task project which was started at the University of Edinburgh as early as 1992. Its aim is the development of a task oriented corpus. The corpus is organized around several files in which the orthographic, the syntactic, and the dialogical level (such as dialogue moves for instance) are annotated, as well as gaze directions (see [Anderson91]). The NITE project (Natural Interactivity Tools Engineering) aims at the development of tools for the annotation and analysis of human interactions ([Traum99] and [Carletta06]). DAMSL (Dialog Act Markup in Several Layers, see [Core97], sub-project of MATE) has seen the development of a consistent dialogue resource from an annotation scheme which aims at finely capturing the functions of propositions (forward move functions in dialogue for instance). Lastly, the Vermobil project consists in the annotation of task oriented corpora between humans. The scheme annotates dialogue acts as well as the different dialogue phases ([Alexandersson98]).

1.3 Objectifs et caractère ambitieux/novateur du projet/*Specific aims, highlight of the originality and novelty of the project*

The major differences between these projects and the one we are presenting today are the following: firstly, our project aims at describing all the speech levels, from the phonological to the gesture one while including interpretative levels. In this way, it is both progressive and ambitious. Moreover the corpora we work on stand in an intermediate position between conversational corpora and task oriented ones. Indeed we chose to record video films of dialogues in experimental conditions while trying to ensure the natural dimension of interaction. This ensuring of naturalness allows us to obtain recordings of speech as spontaneous as possible and to observe phenomena specific to spontaneous speech such as overlaps for instance.

The annotation of verbal and non-verbal communication modalities in an exploitable corpus means a federation of interdisciplinary teams with a competence in linguistic and non-verbal modalities. Such a reunification does not yet exist at a national level although several project teams have worked in isolation on corpora adapted to their personal usage.

The difficulties and weaknesses consequently concern the annotation and processing of corpora. There certainly is a discrepancy between the general level of sophistication of processing, synthesis, edition and observation of the audio and video signal, and the relative weakness of their annotation modules: most of these tools (and/or at least the most common ones) require specific formats for their annotation files. Paradoxically, the situation has become worse these last years, with the increase in number and diversity of available tools, and the improvement in the distribution of tools derived from automatic speech recognition (phonemics aligners), at least for other languages than English.

Even though some tools propose a more “advanced” annotation format (Transcriber or Anvil for instance) it does not bring any solution to the user’s problem (be it a phonetician, a clinician, a linguist, and so forth) who is eventually facing the following dilemma: he must choose between the richness and adequation of possible processing and observations (on the language or on speech), and the effective and comprehensive processing that ensues: while reinforcing certain features one obtains richer and more numerous annotations which are yet largely incompatible with one another and extremely heterogeneous. Their processing necessitates in turn a consequent editing work which is most of the time out of the user’s competence (and time). At the opposite extreme lies a simple – or even simplistic – annotation easily processed.

Some studies attempted to find a solution to these problems (for instance in the EMU framework, see [Cassidy01], as well as the LDC, in relation to Wavesurfer, see http://acl.ldc.upenn.edu/acl2001/STR/7-bird-et-al.pdf) but to the cost of a certain dependence towards the speech processing tools used (such as for instance the strong link with Xwaves). However, we are faced with the following problems:

- There is no resource for French,
- Annotation standards do not fully answer our needs,
- Tools are not well adapted.

Of course, the first of the problems listed above is linked to the quantity of work necessary for the annotation process. But there are also other more fundamental problems linked with the flexibility and reusability of annotations on the one hand, and the need to annotate information from different linguistic fields (phonetics, prosody, syntax, pragmatics and so forth) on the other.

The annotation of this sort of corpora is faced with several difficulties, of which we may define three main types: non-isomorphism of structures, juxtaposition of different levels, and representation need of paradigmatic information.

- Non-isomorphism: the main problem of the annotation of information at different linguistic levels lies in the fact that structures are not isomorphic. This is illustrated by syntactic and prosodic structures. The underlying basic unit here cannot be the word (since a prosodic unit can be composed of truncated words), nor the syllable (since a syllable may be composed of several words). These two structures can consequently not be superimposed. More precisely, we need to represent hierarchical information in which the basic units are different. The solution lies in a lower level of representation such as the segment.
- Different levels: another problem lies in the need to represent at the same time morphological information for example together with information on discourse structure or even metalinguistic annotations. Such annotations were used for instance in the Switchboard corpus in which some pieces of information on "editing" were created.
- Paradigmatic structures: this information type is necessary for the annotation of natural speech. It reveals an organization type which matches the traditional phrasal axis. Such an organization appears under several phenomena: hesitation, repetition, truncation, etc.

Among other needs lies the necessity of defining coding schemes which can be extended to different modalities and which render possible the anchoring of multimodal information (Anvil's temporal information model allows for instance the annotation of the links existing between the different modalities or with object categories which do not change in time), the defining of query languages or validation procedure programs especially designed for multimodal annotations. One of the particularities of our project is precisely to integrate fine grained annotations of the linguistic, extra-linguistic and multimodal levels.

Consequently, the originality of this project first lies in the quantity and quality of the assembled data. Such a resource for the study of French would be unique to this day considering both the size and quality of the data as well as the processing tools used conjointly. Besides, the coding scheme we propose is consistent with international standards and will lead to considerable improvements in the field of multimodal information processing, especially through the proposal of a complex anchoring system. Lastly, unlike most projects born in this field, we suggest to use (or even adapt) existing tools, including data handling. We will not attempt for example to develop a new query language which would be specific to the format proposed, but will instead re-use existing tools.

To conclude we insist on the necessity to mutualize the results obtained. The resources created, the tools as well as the corpora produced within this project will be distributed freely through the means of the Center for Numerical Resources of Oral Data (CRDO – Centre de Ressources Numériques des Données Orales, http://www.crdo.fr/).

1.4 Description des travaux : programme scientifique/*For each specific aim: a proposed work plan should be described (including preliminary data, work packages and deliverables)*

Our project aims at developing a platform integrating annotation and query tools for multimodal corpora. In a first step, we specify a coding scheme for multimodal annotations. The platform itself consists in process sequences leading from raw data to enriched annotations relative to each linguistic domain. The platform will also include editing tools as well as a query engine. Lastly, this platform described in the last part of the project will host several multimodal corpora covering different linguistic phenomena. The expected results are:

- Specification of a coding scheme for multimodal data
- Development of an annotation platform and query tools for multimodal corpora
- Creation and full annotation of new multimodal resources

The three parts of the project are presented in separate workpackages. Each workpackage is organized around different tasks which will lead to specific deliverables.

1. WP1 : Specification of the Multimodal Coding Scheme (MCS)

The present workpackage aims at defining a coding scheme for multimodal information. The objective is herein to propose a coding scheme allowing to encode precisely the different levels of information which describe the

linguistic domains: phonetics and phonology, prosody, morphology and syntax, discourse analysis, gesture study, etc. The first task will consist in surveying the existing schemes in each linguistic field and to carry out an analysis on specific needs for multimodal annotation. The Multimodal Coding Scheme (hereafter MCS) will be defined in a second step.

Task 1.1. Survey of existing schemes and needs analysis

The preliminary step will consist in an exhaustive survey of the different annotation schemes proposed in each linguistic domain. An analysis of the different needs and conventions in use among the teams participating to the project will first be performed. These results will be supplemented by some requirements appearing in other international projects, having in mind that the various coding representations are often tied to a formalism or theoretical options. The identification of the needs will benefit from the experience in the field of each team. In particular, the LPL (corpus CID, 8h) and RFC (conversational corpus, pedagogical corpus, 7h) have produced multimodal corpora, annotated manually or semi-automatically. The LIMSI also makes use of several existing interaction corpora. This preliminary step will provide us with a list of information (segmental unit, category, subcategorization, etc.) and their structures for every linguistic domain.

Several projects have developed coding schemes for multimodal annotation, such as MATE (and NITE), NIMM, EMMA, XCES, TUSNELDA, etc. What comes out of the various coding schemes is mainly that they are very precise in one or possibly two modalities. However, they generally don't cover the entire multimodal domain nor the very precise level of annotation required in every modality. We know by experience that critical problems may arise (both at theoretical and knowledge representation levels) when the different modalities are simultaneously considered together. We will carry out a synthesis of existing solutions and supplement them according to our peculiar requirements.

Deliverables Task 1.1

- Report
    - State of the art: the coding schemes in multimodality studies

Task 1.2. Multimodal Coding Scheme specification

This task is devoted to the building of a generic coding scheme on the basis of the recommendations specified by Task 1.1. The goal is herein to reach a level of genericity allowing the representation of the encoded information independently from the theoretical options chosen. It implies the use of models with a sufficient expressive power. The redundancy of information across several linguistic fields will be possibly accepted.

We will analyze, among the different formalisms proposed in the XML standard, which of them are well suited to take into account the heterogeneous characteristics of annotated multimodal corpora. In a first stage, we will propose a track-by-track representation, i.e. a representation for which each linguistic annotation level is independent from the others and is annotated following a hierarchical model (as proposed by standard XML tools). This coding scheme is intended to allow a global exploitation of corpora annotated at different levels: each level is independent in the representation; the global view comes from the possibility to interface them for edition and manipulation thanks to the knowledge of the structure. The objective will be achieved by extending our existing works on multistructured textual documents (cf. [Bruno & al. 07]). The formal model already experimented by our teams will be adapted (creation of a format and its associated grammar extending the XML language).

We propose to combine the existing schemes and to extend them so as to obtain an XML coding scheme that would be as complete as possible in all the following domains:

- Corpus metadata: we will use a TUSNELDA-like coding scheme (see [Tusnelda05]) in which all the information such as speaker name, sex, region, etc is noted.
- Morphology and Syntax: we propose to adapt the Maptask coding scheme to the French language in the morphological dimension, completed with syntactic relations and properties.
- Phonetics and prosody: some annotations have been inspired by MATE [Mengel00] and completed. The phonetic representation is coded in SAMPA4 and we use the INTSINT and MOMEL algorithms for the phonological representation of intonation.
- Gesture analysis: we adapt the MUMIN coding scheme (cf. [Allwood05], [Allwood06]) by coding separately gestures and discourse tags.

[4] The *Speech Assessment Methods Phonetic Alphabet* is a machine-readable phonetic alphabet. http://www.phon.ucl.ac.uk/home/sampa/

- Pragmatics and discourse analysis: we use the Maptask [Isard01] and DAMSL coding schemes, extended to other discourse types such as narration, description, etc.

This coding scheme will be organized by modality and cover a much broader range of domains than what has been done so far. This organization lies on a system of complex anchorage (see [Blache03]) so as to combine different unit types both in nature and size and thus allow a multimodal processing of the corpus. We propose in the following example an illustration of the coding scheme which could be adopted for encoding morphosyntaxtic informations:

```
<token spelling=”mange”>
  <tag id=“625“ rank=”1” prob=”43”>
    <form id=”mange” lemma =“manger“ freq=”786876” phon=”m~aG” ref=“54“/>
    <cat id= “V”>
       <gender id=”masc”/>
       <number id=”sing”/>
       <pers id=”1”/>
       <mood id=”ind”/>
       <tense id=”pres”/>
       <conjug id=”23”/>
       <defection id=”12”/>
       <verbal_type id=”main”/>
       <subcat id=”td”/>
       <pronominalization id=”o”/>
       <personal id=”p”/>
       <reform id=”n”/>
    </cat >
  </tag >
</token>
```

In practice, the coding scheme concerning facial expressions and head’s motion will be based on the FACS standards. A collaboration with some partners external to the project will be initiated in order to obtain the FACS certification for our coding scheme. As for gestures, the coding scheme will be derived from existing propositions (Kendon 2004; Kipp 2004; McNeill 2005). Gestures typology will be encoded following the scheme proposed in (McNeill 2005). In this scheme, a gesture may inherit more than one category, e.g. some gestures can share iconic and deictic properties. A gestures lexicon will be compiled from the existing descriptions found in the litterature (Kipp 2004; Krenn et Pirker 2004) and on the basis of our own experience. In particular, the scheme will describe relevant aspects of emotional and indivual profiles (handedness) in terms of motion’s quality (Martin 2006). Gesture expressiveness (Pelachaud 2005) and perceptual motion’s quality (Wallbott 1998) will be encoded by adapting the LIMSI manual annotation protocol.

Deliverables Task 1.2

- Resource
    - Specifications of MCS (Multimodal Coding Scheme)

2. WP2: Annotation environment

Corpus annotation consists in adding pieces of information to raw data. This enrichment process can be made manually, semi-automatically or automatically, the result having to be standardized. Until now, corpus annotation concerned essentially written corpora and consisted in adding morpho-syntactic information to a text as well as, in some cases such as the Susanne corpus, or the Penn treebank, syntactic bracketing. The annotation of oral corpora remains limited. Several transcribed oral corpora exist, but they rarely contain other information domains such as phonetics and prosody. We can say that large corpora with precise annotations in so many domains simply do not exist. The following figure describes the full process that makes it possible to produce such corpora. We specify the different steps as well as the methods (automatic or manual annotation) that we propose to use in this project.

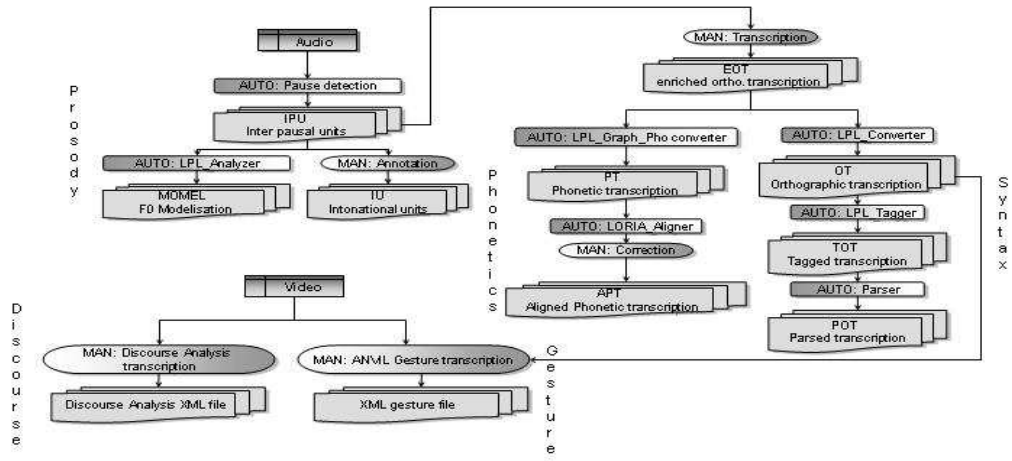

Figure 1. *Full annotation process of a multimodal corpus*

The very first step consists in finding manageable units on which to start the transcription of the speech audio tracks (each track lasting an hour, see “Audio” at the top left corner of figure 1). Speech is then transcribed orthographically with Praat by two experts, these transcriptions being checked by a third expert. This first orthographic transcription follows the guidelines of the GARS (Blanche-Benveniste & Jeanjean, 1987). It is an enriched transcription insofar as some phonetic details and regional pronunciations are noted in square brackets next to the standard spelling. This initial transcription is then converted by a phonetizer which produces a suite of phonemes. This correspondence is based on grapho-phonematic rules for French (see [Di Cristo & al. 01]), the tokens are coded in Sampa to facilitate exchanges between different file formats and computer platforms. The last step in the transcription process consists in aligning the phonemic tokens with the sound signal in an automatic way. This can be done for example by using the system Ants5 (see [Fohr & al. 04]). This alignment is checked manually and provides a precise phonemic transcription of the speech signal. It constitutes the transcription basis for other annotation levels.

We retail in what follows the tasks corresponding to the annotation of the different domains. For each task, we sum up the expected results in terms of recommendations and tools which constitute the deliverables of the task.

Task 2.1. The phonetic-prosodic domain

Transcription, phonetization and alignment steps are achieved with independent tools. A set of specific scripts relates each tool with the other.

Segmentation and automatic transcription

The toolbox developed by the LIA will be the basis of the development of a platform for assisted transcription. The system permits to segment the signal in sound events (speech, non speech, music, etc.), to characterize these events ("studio" speech, telephonic speech), to realize a segmentation and a tracking of the speakers' turns as well as to propose an automatic transcription of speech. In addition, for an improved transcription and alignment with words and phonemes, the system provides a list of words from which the annotator decides on a "good" candidate in case of mismatching. During the correction of the transcriptions, the annotator can re-launch the automatic transcription which will take into account his own corrections.

Segmentation in Interpausal-Units (IPU)

The transcriptions are done from an automatic pre-segmentation of the speech signal in Interpausal-Units (IPU). The IPU are blocks of speech bounded by silent pauses of at least 200 ms (the length may vary according to the language). The IPU segmentation is often used on large corpora. Due to its formal and objective nature, it is distinct from other "prosodic" units such as intonative units for example, which labelling requires the manual intervention of experts, who may not agree in their choices (see [Koiso and al. 98]). This

[5] Automatic News Transcription System.

IPU segmentation facilitates the transcription. It is also needed for the phonetization step as well as the alignment with the audio signal.

Transcription

Each dialog is transcribed by two experts by means of the PRAAT system (see [Boersma & Weenink 05]). The transcription conventions are inspired by those of the GARS [Blanche-Benveniste & Jeanjean 87]; this enriched orthographic transcription (hereafter TOE, for "Transcription Orthographique Enrichie") specifies elisions, particular phonetic realizations (for example "`je sais`" pronounced "`ché`", to account for schwa elision), etc. The transcription also includes all typical speech phenomena such as filled pauses (`euh, mhm, humph`, etc.), false starts, repetitions, truncated words, etc These phenomena must be described precisely in a first transcription since the phonetization and phoneme alignment quality depends on that precision of the initial transcription. The aligner is based on a model of standard French phonemes and works in adverse conditions: devoiced vowels, atypical realization of phonemes, etc. The TOE facilitates the task of the automatic aligner (via the phonetizer) by describing, or even by simulating as much as possible, these particularities. Two other transcriptions are derived from the TOE: one is phonological (needed by NLP modules), the other is "pseudo-phonetic" (for the phonetizer). The text below provides an example:

- TOE: `et c(e) qu(i) était encore plus le choc c'est que en [fait, faiteu]`
- Phonologic version: `et ce qui était encore plus le choc c'est que en fait`
- Pseudo-phonetic version: `et c' qu était encore plus le choc c'est que en faiteu`

The "pseudo-phonetic" transcription is a complex work which may have consequences on the phonetization and automatic alignment steps. Experts often transcribe segments that are not physically present in the speech signal. More rarely, some segments may be present in the speech signal but not transcribed by the experts. This phenomenon does not result from the quality of expert transcription. Speech perception processes lead to an unconscious partial reconstruction of absent or altered phonetic information on the part of the listener. It is also quite normal that transcribers are deaf to some hesitation phenomena. Another manual correction of the alignment is necessary in some cases to re-establish the phonetic annotation in order to match the real speech production.

Phonetization

This step produces the list of phonemes needed by the aligner. After a tokenization, the symbolic phonetizer we use (see [DiCristo & al. 01]) provides a list of word tokens and their phonetization coded in SAMPA. The TOE may sometimes be difficult to use, and a direct phonetic transcription can be, in some cases, simpler for the transcriber; the phonetizer therefore accepts mixed orthographic and SAMPA symbols as an input; SAMPA symbols come out unchanged in the output. The TOE is consequently more precise and standard.

Alignment

The aligner (see [Brown & al. 04]) takes as input the list of phonemes and the audio signal of each IPU. The list of phonemes may contain different phonological variants. The aligner chooses the more appropriate one. The aligner provides in output the time localization of each phoneme on the signal. A script gives the time localization of tokens, both from the phonetisation of the tokens and the time value of the phonemes. The two levels of annotation are connected to the other levels, which will allow the expression of their relation, particularly in the time domain. Figure 2 presents the result of the alignment, visualized with Praat:

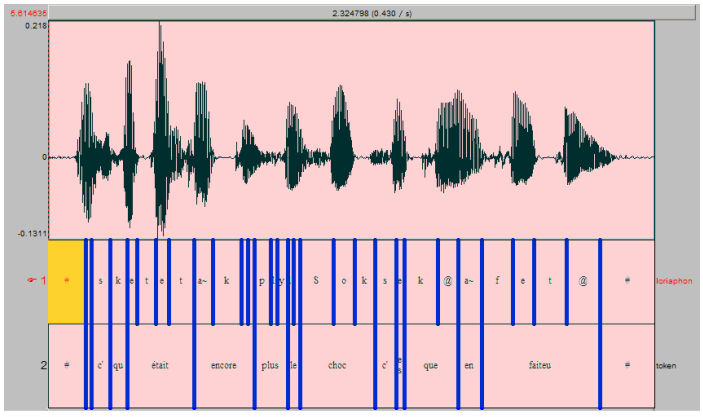

Figure 2. *Time alignment in Praat*

The independence of the transcription, phonetization and alignment procedures facilitates the maintenance, the evolution of the tools (toolkit approach), and their optimal combination. But to be efficient, this approach needs

adequate specification and processing of the data flow; in particular an information present at step n-1 should be retained at step n, even though not used. For example, the aligner should accept some tags (tokenisation tags for ex.) among the input flow of phonemes, and copy them in the output flow of aligned phonemes. Thus time event information can be retained in the aligned phonemes flow, which is not currently the case.

Prosody

Prosodic annotations essentially encode the prosodic categories (intonation and accentual units) and the intonation patterns associated to them. Such annotations are done by experts exclusively. We also use the INTSINT system (see [Hirst and al. 00]) which does not suppose any a priori knowledge of the phonological system of the language. The interest to have both manual annotations and automatic INTSINT annotations is to improve INTSINT itself, but also the knowledge, which is still very fragmentary, of the prosodic domains in French. INTSINT is a symbolic coding of prosody which consists of an alphabet of 8 symbols distributed in two groups :

- symbols coding target points according to the speaker's intonation range
- symbols coding target points in relation to the previous target points

This coding is achieved automatically and is naturally integrated in our system. Basic data are temporal series (audio recording, physiological, aerodynamic parameters, etc.). Until these last years, most of these signal corpora were annotated by tabulated label files: a label is a time value associated to a list of attributes and values. These annotations are made by hand (most frequently) and/or automatically. The use of such annotated corpora can be schematized in the following steps:

- Selection (S): selection of subsets of labels according to given criteria (ex: time sorting by means of regular expressions on the attributes)
- Calculation (C): from the previous selections, parameters and measures are computed / extracted from the data, often with toolboxes of signal processing (Xwaves / Esps, Praat, SFS)
- Analysis (A): these results are analyzed statistically (R, SAS, SPSS, Matlab)

This approach is facilitated by the use of script languages (Awk, Perl, various shells, etc..) as well as specific tools (R, Praat): the S, C and A steps may be grouped together, partially or entirely. At the moment, the toolboxes available fulfil most of the needs of the C and A steps. The present difficulty concerns management of annotations. Tabulated label files are not suited to the orthographic transcription; the cross management of several "levels" of labels is limited. Every annotation/observation tool defines its own format and needs the development of ad hoc tools for the S phase, considering the data solely from the initial survey. This in turn influences the design of the corpus, which is itself most of the time restricted to a unique field of analysis. Consequently, there are few reusable annotations and/or corpora.

Deliverables Task 2.1

- Recommendations
  - Transcription conventions in TOE format
- Original tools
  - Segmenter in IPU
  - Converter TOE → phonological form
  - Converter TOE → pseudo-phonetic form
- Adaptation of existing tools
  - Help to the automatic transcription (Speeral)
  - Phoneme-signal aligner
  - INTSINT annotator

Task 2.2. The morphosyntactic and syntactic domain

For the morphosyntactic annotation, transcribed words are matched with their associated morpho-syntactic categories. There exist several systems (POS–taggers) making it possible to automate this task, with some degree of success: around 95% for written material, a bit less (still not clearly evaluated) for oral ones. Available systems of this kind for French include WinBrill, Cordial, as well as the tagger developed by the LPL (integrated with the LPLsuite platform, see [VanRullen05]). This technology is well known and relatively simple. As such, developing or adapting a tool to augment a corpus with morphosyntactic annotations in a given format (in our case, as we will see later, as arcs represented by XML elements) is rather trivial. However, the tagger must still be adapted to spoken language analysis. Therefore, the system will have to be trained with appropriate data, and custom correcting code heuristics will have to be developed.

As for syntax, annotation remains hard to automatize. Still, a few parsers can be used as a basis for producing annotations. Depending on the desired level of description, two kinds of approaches are possible. The simplest annotation (bracketing) may be based solely on shallow parsing techniques. More precise annotations (for

example relations, functions, roles, etc.) have to be generated by means of deep parsing techniques. The annotation and the formalism used have to be independent from the theoretical frame chosen for the syntactic description, which is true at all levels of the analysis presented here. In other words, whatever the linguistic theory (e.g. HPSG, GP, Dependency Grammars, etc.), there must be a way to generate a model using the MCS scheme proposed in this project.

Shallow parsers, originating from the work of Abney (cf. [Abney91]) only provide information on boundaries between items. They have proved their efficiency for large corpora analyses. They are especially robust and can provide estimations of syntactic units (which serve as a basis for a manual annotation). This kind of analysis is mainly used as an embedded component for applications with a larger scope such as data mining, dialog systems and speech synthesis systems. Several systems are available including those developed at the LPL (see [Blache05]). In terms of symbolic information processing, shallow parsers are based on a set of rules which identify the left and right frontiers of the components, depending on the current item and the various properties of the treated entry. The system takes as an input a tagged text and makes use of functional categories as opening and closing borders for the units. The adaptation of this kind of parser to the treatment of spoken language requires the development of a custom grammar, and the introduction of correcting code heuristics.

Deliverables Task 2.2

- Resource
  - Phonetized lexicon
  - Grammar adapted to the spoken language
- Original tools
  - LPL tagger for speech data, in MCS format
  - LPL parser for speech data, in MCS format

Task 2.3. Multimodal level

Communicative behaviour expressed by other modalities than speech will be annotated: hand gestures, head movements, gaze, posture/position. Hand gestures will be annotated with :

- traditional structural descriptions (preparation, stroke, retraction, ...),
- traditional functional descriptions (deictics, iconics, metaphorics, beats, adaptators),
- hands shape and position

Several tools exist for gestural annotation, some of them have actually been tested by Rohlfing et al. [2006] and the LPL. The tested systems are Anvil, Elan, Exmeralda, Tasx and MacVissta. Each of these tools present advantages and drawbacks. We chose to use Anvil for specific reasons:

- The first point we wanted to make sure of is that we should be able to import Praat tiers into the annotation of gesture phenomena: some transcriptions on the corpus have already been done under Praat (e.g. word transcription) and such a transcription is useful when annotating gestures so it was necessary that the initial transcription would not have to be done again. For this reason, we did not choose Elan which does not support the Praat format.
- We also eliminated McVissta since this particular tool is not multiplatform and only runs on Mac.
- As far as Tasx and Exmeralda are concerned, while the former doesn't allow hierarchically based annotations, the latter is rather designed for verbal transcription to which one can add "comments" about gesture phenomena occurring at the same time. In our annotations, both a hierarchical structure and a very precise transcription of gestures and movements aligned with the sound signal are required so the tools were also ruled out.

Anvil indeed presents all the characteristics needed for the type of linguistic study we want to develop: first, it is a multiplatform tool which guarantees that the annotations made will be usable by the rest of the linguistic community. The fact that it is possible to import tiers from Praat or from other tools provided that the annotations are time-aligned saves a considerable amount of time in the annotation of gestures. Besides, if gestures are generally not aligned with speech since they can occur during a silent pause, some fine annotation (decomposition of the gesture into several phases such as "preparation", "stroke", "hold" and "retraction") requires the possibility of a hierarchical structure. The left boundary of the preparation phase is aligned with the left boundary of the gesture phrase (gesture type in the McNeill typology) and the right boundary of the retraction phase should be aligned with the right boundary of the gesture phrase with extreme precision. Such alignment is essential in a multimodal analysis since some studies have shown that only part of the gesture (mostly the stroke phase) is in correlation with other linguistic phenomena, mostly prosodic ones (Loehr, 2004). To the best of our knowledge, no systematic study involving other multimodal dimensions has yet been done but we may expect from a pilot study developed at the LPL on intensive gestures (gestures which reinforce mostly verbal information, annotated on a short video of 15 minutes) that there will be a strong

interaction between this kind of gestures and phenomena at the morphosyntactic level. The pilot study already showed that intensive gestures occur mainly in correlation with adverbs and connectors (words or phrases linking Turn Constructional Units) and this induces us to predict that there will be a correlation between intensive gestures and the NP in utterance initial position.

Annotation in Anvil is made in two steps: first, we have to establish a specification scheme which will be inserted in a "specification file". This specification file in XML is necessary for the tool to run and contains all the labels needed for the annotation as well as the hierarchies between the different tracks which must be pre-specified. We already constructed several "test" specification files in order to make sure that the tool was adapted to the kind of annotation required for our linguistic needs. This is also the reason why the whole coding scheme must be established in advance: any change in the specification file may result in the loss of information. The second step consists in the annotation proper. This process is quite long and in order to guarantee the quality of the annotations, it must be done by several annotators and the files then cross-examined for inter-annotator agreement.

Deliverables Task 2.3

- Recommendations
    - Conventions of gesture annotations using Anvil
    - Edition of technical recommendations

Task 2.4. Pragmatic level

The needs in terms of labelling, annotating and querying the pragmatic domain concern three different levels:

- meta-data labelling: the corpus and speakers description must be specifically categorized
- identification of three data types:
    - forms or units with variable spans corresponding to the segments temporally delimited
    - non temporal links between annotated categories (anaphoric links or reformulation links for example)
    - discontinuous units which belong to the same discursive activity (humor phenomena for example)
- query needs:
    - listing of all annotated units belonging to the same category
    - distribution of a specific category
    - categories co-occurrences

Some tools exist for the annotation of the pragmatic information level, in particular the RST tool (Michael O'Donnell, WagSoft Linguistic Software), making it possible to indicate the rhetorical structure, following Mann and Thompson's proposal. This system enables to segment texts, mark the structural relations and indicate the set of discursive relations. In the same perspective, we can quote the Nb tool developed by Giovanni Flammia (Spoken Language Systems, MIT Laboratory for Computer Science) which enables the annotation of written finalized dialogue corpora.

This level makes use of recent evolutions in the task-oriented-task domain. In spite of its limitations, this dialogue conception nevertheless constitutes an interesting investigation procedure. A track marks the discursive relations according to (Wolf & Gibson, 2005) and (Hobbs, 1979). A specific track indicates events (accept marks and assertions) which affect the speakers' beliefs. This is a way to note the amount of information in the dialogue context. This proposition is borrowed from the project frames DAMSL (Core & Allen, 1997) and TRINDI [Traum et al., 1999].

At the pragmatic level, other kind of information, in the perspective of Conversational Analysis, concerns elaboration and management of turn-taking (Sacks et al. 1974; Ford and Thompson, 1996, Selting, 2000). The Conversational Analysis framework proposes a mechanism relying on two components: a turn-constructional component (unit construction) and a turn-allocation component (regulation and negotiation of floor control). The first annotations, exclusively manual, refer to turn constructional units (TCUs) which are defined as "the smallest interactionally relevant complete linguistic unit" (Selting: 1998: 40).

Deliverables Task 2.4

- Recommendations
    - Annotations Conventions
- Adaptation of existing tools
    - RST-Tool

Task 2.5 Conversion, edition and validation tools

The ANVIL system will be used as editing, annotating and correcting tool in our project.. Consequently, all the annotations, whether manually or automatically produced (and represented in MCS format), will have to be editable with ANVIL, in order to make it possible to display and correct them.

The number of annotation tracks is not limited in ANVIL. Furthermore, it supports file importation from Praat (indeed, phonetic and prosodic levels of annotation are done with Praat, because of its powerful functionalities, and then imported in ANVIL to display the linguistic categories). The ANVIL coding scheme proposes a dependency process linking the tracks. In the primary tracks, the borders of annotated elements indicate a temporal anchoring. When a secondary track is depending on a primary track, the element's borders correspond to the index of the primary track borders. Moreover, ANVIL makes it possible to create secondary tracks depending on other secondary tracks. Using this mechanism, the hierarchical structure existing in MCS can be converted into ANVIL format, thus creating as many interdependent secondary tracks as levels in the initial structure. Reciprocally, this mechanism makes it possible to convert ANVIL generated annotations to MCS format.

We will develop these XML generic converters MCS/ANVIL and ANVIL/MCS. The goal is double: (1) the possibility of editing and correcting, with ANVIL, the annotations from the different linguistic domains encoded in MCS format; (2) the possibility of recovering the ANVIL manual corrections in MCS format. The XML converter ANVIL/MCS will be integrated as a plug-in into the ANVIL software.

Delivarables Task 2.5

- Original tools
    - XML converter ANVIL_to_MCS
    - XML converter MCS_to_ANVIL

3. WP3 Queries and manipulation

One of the basic hypotheses is that gesture, intonation, syntactic structure or contexts are parts of language if we consider the “direct” oral functioning mode. Interpretation is only possible in relationship to a particular linguistic production. To check the hypothesis of stable functions for an association between parameters from different layers in a language, it is necessary to proceed to exhaustive case studies. The objective is to identify regularities; some of them are starting to appear through the analysis of links between indications from different layers. We first propose in this work package to precisely identify the needs about multimodal corpus manipulation. We shall analyse practices in order to define generic query types. The second step which is the core of this package consists in the development of tools for querying and manipulating this kind of data. Lastly, we shall propose to validate our tools on a test corpus in a specific study with the corpus produced in WP2.

Task 3.1 Needs Identification

It is possible to detect some interaction regularities between modalities linked to the normal functioning of an oral dialogue. For instance, analysis of video recordings has shown that movements of the head or eyes can be interpreted in conjunction with intonational variations and morphosyntactic properties, in the context of the enunciation theory (see [Bouvet & al. 02]). The co-enunciative attitude of the speaker is made clear by his gaze direction: he systematically avoids to look at the listener when he is going to provide modal and referential data that he wants to share with the listener. But, his gaze comes back to the listener before the end of the production of his differentiated personal position (rheme). It enables the speaker to check the validity of his anticipations compared to the possible reactions of the listener (consensus, disagreement or incomprehension). On the contrary, when the speaker faces a formulation difficulty, he systematically looks away from the listener, either at the ground, or up or at a particular point of the dialogue space. His hand gestures are mainly interpreted in considering their anticipative help to the formulation and to the organization of discourse. They enable the deictic location of a referent in the real space of the oral exchange and they mark the search of the good formulation of what the speaker has to say; therefore they help managing the smooth continuation of the dialogue.

Some query tools for this kind of data have been recently proposed in the context of the construction of French syntactically annotated corpora. Here are only mentioned advanced systems which are more expressive than simple regular expressions. [Christ94] developed a query tool named XKWIC which is able to extract fragments matching regular expressions on shapes and categories from a corpus. [Kallmeyer00] defined a query tool

dedicated to syntactically annotated corpora which takes into account the relations of dominance and precedence.
This kind of functionalities will be extended in order to be able account for every interaction type and the queries necessary to process the data. Our first use of queries will be to check the validity of annotations by means of the calculus of inter-annotator agreements. Moreover, measures like the following ones will be calculated:

- A descriptive analysis of each modality (for instance gesture lemmas)
- A matrix of transition between behaviour types in the same modality.
- Search of the modalities in synchronic or rhythmic relations.
- Characteristics of a modality (for instance the gesture profile)
- Balance of the repartition of linguistic information between linguistic modules (Blache & Meunier, 2004)
- Relations between phonetic variability and lexical components (Meunier, 2007)

Once the measurements obtained, we will be able to calculate statistic models and to explore new hypotheses on multimodal behaviour.

Delivrables Task 3.1

- Report
  - Definition of the different kinds of interaction
- Resource
  - Test set of queries (expressed in natural language)

Task 3.2 Definition of a query tool

Our objective is to let each annotator use his own tool bench and to propose a common underlying data model, an architecture, and a set of operators to enable the multimodal exploitation of the data. Our theoretical standpoint being to share data and resources, we will use open standards (for formal models and tools) from the XML universe as much as possible.

The main difficulty in defining a data model comes from the heterogeneity and the distribution of the resources. We will first choose a common formal model and will then define automatic procedures to convert legacy data to the common formalism.

A set of operators specific to our data model will be proposed. Some of these operators will be designed to enable the independent querying of each modality, while others will be dedicated to the concurrent querying, that is to say the querying of annotations belonging to two or more modalities or even to query the relationships between modalities. For instance, we want to be able to query the description of a gesture (Q1 : « When did the speaker look at the listener ? ») and the link between gesture and intonation contours (Q2 : « What type of intonational contour did the speaker use when he looked at the listener ? »). We will provide the user with: a query language, a tool to navigate within the data, and mechanisms to elaborate different perspectives of the global data or even to build new data (using the query language).

More precisely, the manipulation will consist in querying that is to say: find patterns, navigate in the structure and test parts of the corpus using classical criteria (on values and annotations) and existing relations (temporal or structural ones). The study of multimodal data when each modality is described using a rich hierarchical structure opens new kinds of manipulations; for example discovering (strict or lax) inclusions or overlaps between annotations. The results of those queries will be useful to obtain statistical results but also to help in constructing new annotations or to extend existing ones.

Our proposal is to define these operators as an extension of a standard query language for XML. This approach will enable us to combine expressive operators in the same query (selection, join, group by, nested queries ...). This will enable the user to express fine grained queries and even to build new data. The localisation of fragments in the corpus will be done using a simple and powerful language using path models (XPath). To do this, we will extend XQuery (http:www.x3.org/xquery) the recommended query language for XML by the needed operators. We will rely as much as possible on open and standard models and tools. We will use our own implementation of XQuery (developed in the context of multistructured XML documents: http://lsis.univ-tln.fr/msxd/).

To validate our model and query language a set of representative multimodal queries will be proposed conjointly by all the partners (see http://www.w3.org/TR/xquery-use-cases/).

Notice that a wide range of hypotheses (and consequently of queries) can be done on this type of corpus. We will mainly focus on the exact querying of the annotations and content of the data. In particular, we may use

data directly based on the raw signal but only represented in terms of our model (and not directly from the binary data). A set of specific analyses (for exemple statistical ones) will be done as post treatment using existing tools, that is why we will define simple XML output format, which will be easily converted.

Deliverables Task 3.2

- Resources
  - A set of operators extending XQuery
- Original tools
  - A query language
  - Tools to describe and manipulate the data
  - A prototype of data storage system.

Task 3.3 Exploitation of the query tools

We propose to set up an experimental study to validate the data model and the query tools developed in this workpackage. This study will focus on listening signals such as backchannels (BC) that is to say the set of signals (speech, sound, gesture) emitted by the participant of a dialogue to show his interest, his understanding, his acknowledgement, etc. of the discourse. Even if the state of the art about these phenomena is important, no study presents a complete analysis or show the role of these signals in the interaction between the speaker and the listener. Following Fox Tree (1999), we are making the hypothesis that BCs give information on listening processes but also on the way the speaker produces his discourse in highlighting the important steps of its elaboration. Moreover, and even if they can be expressed both verbally and/or in gestures no systematic multimodal study on BCs has already been conducted and very few studies have focused on the way the two kinds of BCs interact. We explain this gap in the analysis of BCs by the lack of dedicated corpora. It also explains why only few linguistic levels are taken into account in linguistic analyses. Explaining the apparition context of BCs in the interaction gives out a better rendering of speech turns during the exchange. Our hypothesis, relying on several studies concerning speech turns, states that the speaker and listener use various resources to keep or give up floor control. In the same way, BCs do not appear randomly and occur after cues of different types (a given intonation contour, a given gesture, a combination of both, ...). Lastly, the cues change according to the function of the modality of the BC produced. A gesture BC and a verbal one may be induced by different factors.

This kind of study involves the concurrent querying of several levels of annotations and the existence of tools to assist in the generation of the queries. The following list of queries illustrates the role and the importance of multimodal cues occurring in the apparition of BCs:

- Among the different intonation contours, is one of them more specifically linked to BCs?
- Does a given intonation contour have an impact on the BCs' modality? For instance, would a continuation rising contour rather induce a gesture BC whereas a terminal falling contour would rather induce a verbal BC?
- Are BCs rather induced by conversational units which end would be marked by specific phonetic correlates such as 'creaky voice'?
- Are the BCs in overlap with speech produced more frequently at the end of a syntactic unit, an intonational unit or a conversational unit?
- Concerning the cumulative weight of cues, is a verbal and gesture BC induced by a multimodal combination such as a particular type of conversational unit associated both with a certain intonational contour and a particular morphosyntactic category?

This query processing step will be followed by a statistical analysis of the data leading to a formal and functional typology of BCs. This typology will be based on the whole set of elements noticed as important for the apparition of BCs and not only on one criterion. This kind of study constitutes one of the hypotheses we want to explore in order to explain the mechanisms underlying turn taking in interactions.

Deliverables Task 3.3

- Reports
  - Study of backchannels in the test corpus
  - Report of the validation of the query language tool

4. WP4 Validation

The project will deliver several multimodal corpora annotated with the platform designed in the framework of the project. We will include different recording settings (among which recordings in anechoic room and TV setting)

and different conversational types (monologues, dialogues, interviews, talk-shows, semi-directed conversations such as Map-Task interactions (Anderson & al.1991)).

We will manually annotate conversational gestures during face to face interactions using the dimensions of gesture phases & phrases, entries in a gesture lexicon which will be defined along with a pre-analysis of the collected data, as well as gesture expressivity for the study of emotional gestures (McNeill 2005, Kendon 2004, Kipp 2004, Pelachaud 05, Martin et al. 05). We will also study gestures produced by the speaker during telephone conversations, much in the way of the « Switchboard » corpus (Godfrey & al. 1992). The corpus will also include pathological gestures (produced by aphasic patients) recorded during speech therapy sessions. Facial expressions will be annotated using a coding scheme based on FACS (Ekman et al. 2002).

We will produce 1) additional annotations of existing multimodal corpora, and 2) newly collected and annotated multimodal corpora. These corpora will be distributed to the research community via the CRDO resource centre.

Task 4.1. Additional annotations of existing multimodal corpora

The LPL has recorded several corpora of dyadic spontaneous spoken conversations. The CID corpus is one of them (Bertrand & al 07). If features 8 hours of data recorded in anechoic room and containing 110 000 words. Each speaker of the dialogue is equipped with a headset microphone enabling the recording of the two speakers' voices on two different tracks. This enables the study of speech at a phonemic and prosodic level. It also enables the study of overlapping speech which is frequent in spontaneous interactions but seldom analyzed because of the difficulty to separate the intertwined voices of the two speakers a posteriori. Yet, overlapping speech plays a main role in conversation and requires experimental investigations.

In this corpus, we aimed at collecting useful data for the investigation of all the levels under study: the audio quality is optimum and they have been videotaped with a high quality digital camera.

The LPL also recorded 6 hours of Map-Task interactions. This is the first French version of such a corpus which has been collected for several other languages such as English, Swedish, Italian. The RFC team is also the depositor of several corpora collected in a variety of interactive situations:

- Spontaneous speech between two participants who know each other (2h30)
- Asymmetrical conversations between a Japanese learner of French and a native French speaker (1 h)
- Training sessions of job interviews in an engineering school (2h30)
- Pedagogical interactions during a seminar (1 h)

The RFC corpora are partly annotated at the prosodic level (especially the intonation contour using Praat), but annotations also consist of word by word transcriptions, discourse annotations, annotations of gaze directions, head, torso and hand movements of the participants in each interaction.

Deliverables Task 4.1

- Completion of existing resources
    - Full annotation of the CID corpus
    - Full annotation of the MapTask corpus
    - Additional annotation of the RFC corpus

Task 4.2. Collection and annotation of new multimodal corpora

In order to validate the platform built within the framework of the project, three new multimodal corpora will be collected and annotated.

- A corpus of narration: we will define a French version of the protocols adopted by McNeill for the study of conversational gestures (e.g. cartoon narration and house description).
- Corpus of telephone conversations: this will enable the annotation of facial expressions and gestures when the listener and the speakers do not see each other.
- Corpus of pathological communication: two recordings of aphasic patients have already been collected. New recordings will be collected and fully annotated.

Deliverables Task 4.2

- New resources

- Collection and full annotation of the corpus of narration
- Collection and full annotation of the corpus of telephone conversations
- Additional recordings and full annotation of the corpus of pathological communication

1.5 Résultats escomptés et retombées attendues/*Expected results and potential impact*

The main expected results of this project are:

- Multimodal coding scheme
- Conventions and standards for multimodal annotation
- Multimodal annotation tools
- Multimodal data query tools

As explained above, these results are of great importance, both theoretically and for applications. Multimodal interaction still has to be described precisely: at the moment, no global theory explains the different interaction between modalities. The first step towards this goal consists in bringing together, into a unique resource, all the different information, coming from the different domains (phonetic, pragmatic, gestures, semantics, etc.). Building such resources needs first the specification of conventions and the definition of a coding scheme. This will be the first main result of the project, no global multimodal coding scheme being available at this moment.

Starting from these formal aspects, the annotation task itself is addressed by the project which aims at creating or adapting different tools helping in the task of building multimodal annotated corpora. The goal is to create automatically or semi-automatically the different enrichments or annotations, starting from raw data. The interest of the project lies in the possibility of helping in the creation of all the different enrichments in a homogeneous way.

The last important result consists in developing different tools making it possible to manipulate data. The goal, instead of developing new query language, is to reuse existing techniques, adapting them to the specific needs of multi-structured documents.

The result of the project is then threefold: standards and conventions, tools and resources.

Finally, here is the recap of the different results, by work package:

WP1 : Specification of the Multimodal Coding Scheme (MCS)

- Task 1.1. Survey of existing schemes and needs analysis
  Report: State of the art: the coding schemes in multimodality studies
- Task 1.2. Multimodal Coding Scheme specification
  Resource: Specifications of MCS (Multimodal Coding Scheme)

WP2: Annotation environment

- Task 2.1. The phonetic-prosodic domain
  Conventions: Transcription conventions in TOE format
  Tools: IPU segmenter, Converters TOE → phonological form, TOE → pseudo-phonetic form, Help to the automatic transcription Phoneme-signal aligner, INTSINT annotator
- Task 2.2. The morphosyntactic and syntactic domain
  Resource: Phonetized lexicon, Grammar adapted to the spoken language
  Tools: tagger for speech data (in MCS format), parser for speech data, in MCS format
- Task 2.3. Multimodal level
  Conventions: gesture annotations, edition of technical recommendations
- Task 2.4. Pragmatic level
  Conventions: Annotations
  Tools: adaptation of RST-Tool
- Task 2.5 Conversion, edition and validation tools
  Tools: XML converter ANVIL_to_MCS, XML converter MCS_to_ANVIL

WP3 Queries and manipulation

- Task 3.1 Needs Identification
  Report: Definition of the different kinds of interaction
  Resource: Test set of queries (expressed in natural language)
- Task 3.2 Definition of a query tool
  Resources: A set of operators extending XQuery
  Tools: query language, data description and manipulation, data storage system.
- Task 3.3 Exploitation of the query tools
  Reports: Study of backchannels in the test corpus, report of the validation of the query language tool

WP4 Validation

- Task 4.1. Additional annotations of existing multimodal corpora
  Resources: Full annotation of the CID, the MapTask, the RFC corpora
- Task 4.2. Collection and annotation of new multimodal corpora
  Resources: Annotated multimodal corpus (narration, phone conversations, pathological communication)

1.6 Organisation du projet/*Project flow*

Le projet est organisé autour de trois partenaires, chacun étant responsable d'un type d'activité particulier. Le LPL assurera la coordination de l'ensemble. Il s'agira d'une part de vérifier le déroulement des tâches assignées et d'autre part de synchroniser les travaux collectifs. Ces derniers sont de trois types :

- Etat de l'art
- Specification du schéma d'encodage
- Validation

Par ailleurs, la coordination du projet reposera sur des rencontres régulières de tout ou partie des partenaires, en fonction des tâches. Le développement des outils devra reposer sur les spécifications établies par la description des besoins : cette partie concerne plutôt le LPL et le LIMSI. Le système de manipulation de données reposera plus particulièrement sur la mise en relation des formats, des outils et du langage et concernera donc plus particulièrement le LSIS et le LPL. La validation enfin exploitera des données, des besoins et des outils et reposera pour cela essentiellement sur le LIMSI et le LSIS. Enfin, la tâche de coordination consistera à s'assurer de la bonne exécution du budget et de la gestion des ressources humaines.

1. WP1 : Définition du schéma d'encodage
- Task 1.1. Besoins et bilan des schémas existants
- Task 1.2. Spécification du MCS

2. WP2 : Environnement d'annotation
- Task 2.1. Le domaine phonétique-prosodique
- Task 2.2. Le domaine morpho-syntaxique et syntaxique
- Task 2.3. Le domaine des comportements multimodaux
- Task 2.4. Le domaine pragmatique
- Task 2.5. Outils de conversion, édition, validation

3. WP3 Requêtes et exploitation
- Task 3.1. Analyse des besoins
- Task 3.2. Développement d'un outil de requête
- Task 3.3. Exploitation de l'outil sur corpus test

4. WP4 Validation
- Task 4.1. Annotation de corpus multimodaux existants
- Task 4.2. Création et annotation de nouveaux corpus multimodaux

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
WP1												
Task1.1												
Task1.2												
WP2												
Task2.1												
Task2.2												
Task2.3												
Task2.4												
Task2.5												
WP3												
Task3.1												
Task3.2												
Task3.3												
WP4												
Task4.1												
Task4.2												

Tab. 1 : Planning par tâche

La situation générale du projet et son état de l'art sont d'ores et déjà établis. Le projet pourra donc débuter directement sur des tâches spécifiques conditionnant le développement de la plateforme. En particulier, un bilan précis sur les schémas d'encodage existants d'une part et la spécification des besoins d'autre part constitueront une étape préliminaire. Les deux parties les plus importantes en termes de développement concernent d'une part les outils d'aide à l'annotation et à l'édition et d'autre part les outils de requête et de manipulation des données. Ces développements, relativement indépendants en termes techniques, pourront être conduits en parallèle dès lors que le schéma d'encodage sera fixé. La dernière partie du projet impliquera l'ensemble des participants ; Elle permettra, en produisant de nouvelles ressources, de valider l'ensemble des propositions. Elle consistera plus particulièrement à annoter des corpus existants ou de nouveaux corpus selon le protocole établi. Cette étape nécessitera donc sans doute l'adaptation des outils développés dans le second workpackage, ce qui apparaît en grisé dans le calendrier ci-dessus.

	T1.1	T1.2	T2.1	T2.2	T2.3	T2.4	T2.5	T2.6	T3.1	T3.2	T3.3	T4.1	T4.2
LPL													
LIMSI													
LSIS													

Tab. 2 : Intervention des équipes par tâche

La coordination générale du projet reposera sur le LPL qui intervient à ce titre sur toutes les tâches (à l'exception du développement des outils de requête). La première phase de spécification concerne toutes les équipes. Les travaux de développement et d'intégration de la plateforme reposent sur le LPL (en particulier annotation prosodique, mortphosyntaxique, syntaxique, sémantique, pragmatique, conversion de formats), le LIMSI (outils de conversion, alignement et le développement de l'aide à la transcription automatique) et le LSIS (manipulation de données, langage de requête). Les travaux de description et d'annotation linguistique reposent sur le LPL et le LIMSI. En particulier, les ressources humaines seront concentrées sur le dernier workpackage, destiné à valider la plateforme par la constitution de nouvelles ressources.

Tâches	**LPL**	**LSIS**	**LIMSI**	**Total / tâche**
T1.1	3	2	6	11
T1.2	3	2	6	11
T2.1	14		16	30
T2.2	14			14
T2.3	14		14	28
T2.4	14			14
T2.5	12		12	24
T3.1	2	4	2	8
T3.2		32	2	34
T3.3	12	6	14	32
T4.1	16	2	14	32
T4.2	14	2	16	32
TOTAL	118	50	102	270

Tab. 3 : Répartition personnel (homme-mois) par tâche et par équipe

Le consortium proposé repose sur une grande complémentarité des équipes. Il convient tout d'abord de souligner l'équilibre thématique, déterminant pour un projet interdisciplinaire : une équipes se situe principalement dans le domaine de la linguistique, deux dans l'informatique. Du point de vue du nombre de participants, le consortium est équilibré entre les deux disciplines.

Par ailleurs, il convient de remarquer que chaque tâche fait intervenir plusieurs équipes, à l'exception de la tâche portant sur le développement du langage de requête, portée par le LSIS car reposant sur une technicité spécifique. Chaque équipe, parmi ses interventions, interviendra de façon plus sensible sur certains aspects :

- LPL : Intégration de la plateforme, validation de l'outil de requête, annotation de corpus existants et nouveaux corpus, coordination de l'ensemble.

- LIMSI : Etat de l'art, développement du schéma d'encodage, développement d'outils de validation, validation de la chaîne (annotation de corpus existants), développement d'outils d'édition, développement de l'aligneur, de l'aide à la transcription

- LSIS : Développement du schéma d'encodage, développement du langage de requête et manipulation des données, validation des outils.

TABLEAU des LIVRABLES et des JALONS (le cas échéant)/ *Deliverables and milestones*			
Tâche/ *Task*	Intitulé et nature des livrables et des jalons/ *Title and substance of the deliverables and milestones*	Date de fourniture nombre de mois à compter de T0/ *Delivery date, in months starting from T0*	Partenaire responsable du livrable/jalon/ *Partner in charge of the deliverable/ milestone*
1. Specification of the Multimodal Coding Scheme			

	Report: State of the art: the coding schemes in multimodality studies		LPL
	Resource: Specifications of MCS (Multimodal Coding Scheme)		LIMSI
2. Annotation environment			
	Conventions: Transcription conventions in TOE format		LPL
	Tools: IPU segmenter		LPL
	Tools: Converters TOE → phonological form ; TOE → pseudo-phonetic form		LPL
	Tools: Help to the automatic transcription (Speeral),		LIMSI
	Tools: Phoneme-signal aligner		LPL
	Tools: INTSINT annotator		LPL
	Tools: tagger for speech data (in MCS format)		LPL
	Tools: parser for speech data (in MCS format)		LPL
	Conventions: gesture annotations		LIMSI
	Conventions: edition of technical recommendations		LIMSI
	Tools: XML converter ANVIL_to_MCS, XML converter MCS_to_ANVIL		LPL
	Resource: Phonetized lexicon, Grammar adapted to the spoken language		LPL
3. Queries and manipulation			
	Report: Definition of the different kinds of interaction		LIMSI
	Resource: Test set of queries		LSIS
	Tools: query language, data description and manipulation, data storage system		LSIS
	Reports: Study of backchannels in the test corpus, report of the validation of the query language tool		LSIS
	Resources: A set of operators extending XQuery		LSIS
4. Validation			
	Resources: Full annotation of the CID, the MapTask, the RFC corpora		LIMSI
	Resources: Annotated multimodal corpus (narration, phone conversations, pathological communication)		LIMSI

1.7 Organisation du partenariat/*Consortium organisation*

Le projet OTIM est le résultat d'une collaboration engagée dans le cadre du pré-projet PRAX, initié voici 2 ans par la fédération ILF et réunissant les partenaires actuels. Les différentes équipes du consortium entretiennent donc déjà des collaborations ayant notamment conduit au développement d'outils et de ressources multimodales.

1.7.1 Complémentarité et synergie des partenaires/*Added value of the consortium*

Le projet OTIM est fortement interdisciplinaire : il rassemble des spécialistes de la description et du traitement de chacune des modalités impliquées dans l'interaction mutlimodale. Il s'agit donc de linguistes, de pragmaticiens et de spécialistes de l'interaction homme/machine. De plus, les traitements visés étant, au moins partiellement automatisés, le projet s'appuie sur des informaticiens, spécialisés dans le traitement automatique des langues naturelles. Enfin, le type de données visé étant spécifique et les données elles-mêmes très volumineuses, ce projet comporte également des spécialistes du traitement des données.

Chacun des trois partenaires est plutôt spécialisé dans l'un de ces secteurs : linguistique pour le LPL, interaction homme-machine pour le LIMSI et traitement des données pour le LSIS. Mais il est important de souligner le fait que les différents partenaires ont d'ores et déjà une longue pratique de l'interdisciplinarité dans ce domaine : les collaborations entre linguistes et informaticiens font partie du projet scientifiques des laboratoires LPL et LIMSI. Le LSIS quant à lui est engagé depuis de nombreuses années dans des collaborations avec différentes équipes (et notamment le LPL) pour le traitement de données linguistiques.

Au total, le projet est donc très interdicplinaire et repose sur une forte interaction entre les partenaires, rendue possible par leur complémentarité et leur expérience de l'interdisciplinarité.

1.7.2 Qualification du coordinateur du projet et des partenaires/*Principal investigator and partners : résumé and CV*

Philippe Blache, Directeur de Recherche, CNRS

✓ **Academic degrees**

- Habilitation à diriger des recherches (1999), Université Paris 7
- PhD in Computer Science (1990), Université d'Aix-Marseille II

✓ **Professional experience**

- [2002--] Research Director, CNRS, Laboratoire LPL (Aix-en-Provence)
- [1998-02] Researcher, CNRS, Laboratoire LPL (Aix-en-Provence)
- [1992-97] Researcher , CNRS, Laboratoire 2LC (Sophia Antipolis)
- [1990-92] Assistant Professor, Université de Neuchâtel (Switzerland)

✓ **Responsibilities**

- Head of the LPL (UMR 6057 CNRS & Université de Provence)
- Board Member: EACL (European Association for Computational Linguistics), ESSLLI (European Summer School in Logic, Language and Information), TALN (Traitement automatique des langues naturelles)
- Editorial board member : "Traitement Automatique des Langues", "Information, Interaction, Intelligence"

✓ **Head of research projects**

- [1988-92] « Prolog et TALN », Industries de la Langue, Université de Montréal (Canada).
- [1994-96] « Intégration Prosodie/TALN », GDR-PRC CHM
- [1995-97] Groupe de pilotage GRACE (évaluation des systèmes de compréhension de textes), AUPELF
- [1994-98] « Théories linguistiques et traduction automatique » LTT-AUPELF
- [1999-01] « AC-TALG : Analyse par contraintes pour un traitement automatique des langues générique » APN (CNRS)
- [2000-02] « Plateforme de communication alternative pour l'aide aux handicapés», programme Cognitique, Ministère de la Recherche
- [2000-02] « Multimodalité et communication assistée», programme Société de l'Information, CNRS
- [2003-05] « Autisme et communication», CNRS

✓ **Publications**

- 2 Books, 12 Reviews, 14 Chapters, 82 Articles

✓ **Invited conferences**

- [1998] Workshop on Constraint Programming and NLP, Tübingen, May 1998
- [2000] International Conference NLP-2000, Patras (Greece), June 2000
- [2001] Workshop " NLP Environments", Paris, December 2001
- [2004] International Conference on Categorial Grammar, Montpellier, June 2004
- [2004] International Workshop on Constraint Solving and Language Processing, Copenhagen, September 2004
- [2006] Workshop "Geometry of Calculation", Marseille, February 2006

✓ **Conference Chair**

- [1994-96] TALN'94, TALN'95, TALN'96, Conférence «Le Traitement Automatique du Langage Naturel», Marseille
- [2004] ACL'04 Poster and Demo session, Barcelona, July.
- [2005] LACL'05 (Logical Aspects of Computational Linguistics, Bordeaux, April.
- [2006] CSLP-06, ACL-COLING workshop, Sydney, July

✓ **Organizing committee chair**

- [1992] Workshop «Les correcteurs grammaticaux», SGAICO'92, Université de Neuchâtel
- [1996] HPSG'96, Conférence Internationale sur HPSG, Marseille
- [1996] ATALA Workshop «Prosodie et Syntaxe», Paris
- [1997] ESSLLI'97, European Summer School on Language Logic and Information, Aix
- [1998] Workshop on "Constraints in Programming and Linguistics", Tübingen
- [2000] ATALA Workshop «Représentation et traitement de l'ambiguïté», Paris
- [2000] Workshop on «New Methods and Formalisms for Corpus Linguistics», Aix
- [2001] Assises de Cognisud, Marseille
- [2002] Rencontres « Art, sciences et cognition », Marseille, 8 Mars 2002.
- [2004] JEP-TALN Conference, Fès (Maroc)
- [2004] ICCG-3 Conference, Marseille

1.8 Stratégie de valorisation et de protection des résultats/*Data management, data sharing, intellectual property strategy, and exploitation of project results*

Parmi les principaux résultats attendus du projet se trouvent d'une part la mise au point d'outils, et d'autre part création de ressources. Dans les deux cas, la question de leur mise à disposition de la communauté est centrale. Nous nous appuierons pour cela sur le centre de ressources numérique CRDO (Centre de Ressources Numériques des Données Orales, http://www.crdo.fr/), dont la vocation est la diffusion de données et ressources pour le traitement de l'oral.

Le CRDO s'appuie sur un conseil scientifique qui sera associé à la spécification des besoins et jugera du type de mise à disposition. Le principe du CRDO repose sur la spécification d'une licence de distribution, élaborée conjointement par les auteurs des ressources et outils et le conseil scientifique. Notre objectif dans le cadre du projet est de mettre à disposition librement de l'ensemble de la communauté académique tous les outils de traitement ainsi que les ressources de premier niveau (lexiques, grammaires). Les ressources à forte valeur ajoutée (i.e. les corpus annotés) seront distribués de façon plus contrôlée : des échantillons seront mis à disposition de façon libre, mais l'accès à la totalité des ressources se fera sur demande spécifique, dans uneperspective de recherche collaborative avec les créateur des données.

Collez ici le tableau de récapitulatif des données financières de la fiche budgétaire.

FICHES BUDGÉTAIRES - Blanc

Fiche Partenaire 1

Nom Complet du partenaire	Catégorie de partenaire	Base de calcul pour l'assiette de l'aide
BLACHE Philippe	Organismes de recherche+Fondation de recherche	Coût marginal

Données financières (montant HT en € incluant la TVA non récupérable) 1

EQUIPEMENTS (€)	Personnels						Prestations de service externe (€)	Missions (€)	Autres dépenses (€)	Dépenses justifiées sur facturation interne (€)	Totaux (€)
	permanents		non permanents à financer par l'ANR		Autres non permanents						
	personne. mois	Coût (€)	*personne. mois*	Coût (€)	*personne. mois*	Coût (€)					
	88,00	367 224	*30,00*	109 170				13 500	12 500		502 394
Montant maximum des frais de gestion/ frais de structure (€)				5 407			<--Frais de gestion / frais de structure demandés (€)-->				5 407
Uniquement pour laboratoire d'organisme public de recherche ou fondation financé au coût marginal, indiquer le taux d'environnement								80,0%	Frais d'environnement (€)		381 115
										Coût complet (€)	**888 916**
										Coût éligible pour le calcul de l'aide : Assiette (€)	**135 170**
1							Taux d'aide demandée)-->	100%	***Aide demandée (€)***		***135 170***

Fiche Partenaire 2

Nom Complet du partenaire	Catégorie de partenaire	Base de calcul pour l'assiette de l'aide
MARTIN Jean-Claude	Organismes de recherche+Fondation de recherche	Coût marginal

Données financières (montant HT en € incluant la TVA non récupérable) 1

EQUIPEMENTS (€)	Personnels						Prestations de service externe (€)	Missions (€)	Autres dépenses (€)	Dépenses justifiées sur facturation interne (€)	Totaux (€)
	permanents		non permanents à financer par l'ANR		Autres non permanents						
	personne. mois	Coût (€)	*personne. mois*	Coût (€)	*personne. mois*	Coût (€)					
	72,00	300 456	*30,00*	109 170				8 500	9 500		427 626

Montant maximum des frais de gestion/ frais de structure (€)	5 087		<--Frais de gestion / frais de structure demandés (€)-->		5 087
Uniquement pour laboratoire d'organisme public de recherche ou fondation financé au coût marginal, indiquer le taux d'environnement		80,0%	Frais d'environnement (€)		327 701
			Coût complet (€)		**760 414**
			Coût éligible pour le calcul de l'aide : Assiette (€)		**127 170**
1		Taux d'aide demandée)-->	100%	***Aide demandée (€)***	***127 170***

Fiche Partenaire 3

Nom Complet du partenaire	Catégorie de partenaire	Base de calcul pour l'assiette de l'aide
MURISASCO Elisabeth	Organismes de recherche+Fondation de recherche	Coût marginal

Données financières (montant HT en € incluant la TVA non récupérable) 1

EQUIPEMENTS (€)	Personnels						Prestations de service externe (€)	Missions (€)	Autres dépenses (€)	Dépenses justifiées sur facturation interne (€)	Totaux (€)
	permanents		non permanents à financer par l'ANR		Autres non permanents						
	personne. mois	Coût (€)	*personne. mois*	Coût (€)	*personne. mois*	Coût (€)					
	34,00	141 882	*20,00*	58 224				4 500	6 500		211 106

Montant maximum des frais de gestion/ frais de structure (€)	2 769		<--Frais de gestion / frais de structure demandés (€)-->		2 769
Uniquement pour laboratoire d'organisme public de recherche ou fondation financé au coût marginal, indiquer le taux d'environnement		80,0%	Frais d'environnement (€)		160 085
			Coût complet (€)		**373 960**
			Coût éligible pour le calcul de l'aide : Assiette (€)		**69 224**
1		Taux d'aide demandée)-->	100%	***Aide demandée (€)***	***69 224***

Fiche Partenaire 4

Nom Complet du partenaire	Catégorie de partenaire	Base de calcul pour l'assiette de l'aide

Veuillez préciser la catégorie de partenaire

Données financières (montant HT en € incluant la TVA non récupérable) 2

EQUIPEMENTS (€)	Personnels						Prestations de service externe (€)	Missions (€)	Autres dépenses (€)	Dépenses justifiées sur facturation interne (€)	Totaux (€)
	permanents		non permanents à financer par l'ANR		Autres non permanents						
	personne. mois	Coût (€)	*personne. mois*	Coût (€)	*personne. mois*	Coût (€)					
											-

Montant maximum des frais de gestion/ frais de structure (€)	-		<--Frais de gestion / frais de structure demandés (€)-->	-
Uniquement pour laboratoire d'organisme public de recherche ou fondation financé au coût marginal, indiquer le taux d'environnement			Frais d'environnement (€)	-
			Coût complet (€)	**-**
			Coût éligible pour le calcul de l'aide : Assiette (€)	**-**
1		Taux d'aide demandée)-->	***Aide demandée (€)***	*-*

Récapitulatif des données financières

	EQUIPEMENTS (€)	Personnels						Prestations de service externe (€)	Missions (€)	Autres dépenses (€)	Dépenses justifiées sur facturation interne (€)	Totaux (€)
		permanents		non permanents à financer par l'ANR		Autres non permanents						
		personne. mois	Coût (€)	*personne. mois*	Coût (€)	*personne. mois*	Coût (€)					
Partenaire 1	-	88	367 224	30	109 170	-	-	-	13 500	12 500	-	502 394
Partenaire 2	-	72	300 456	30	109 170	-	-	-	8 500	9 500	-	427 626
Partenaire 3	-	34	141 882	20	58 224	-	-	-	4 500	6 500	-	211 106
Partenaire 4	-	-	-	-	-	-	-	-	-	-	-	-
	-	**194,00**	**809 562**	***80,00***	**276 564**	**-**	**-**	**-**	**26 500**	**28 500**	**-**	**1 141 126**

Frais de gestion / frais de structure demandés (€)-->	13 263
Frais d'environnement (€)	868 901
Coût complet (€)	**2 023 289**
Coût éligible pour le calcul de l'aide : Assiette (€)	**331 564**
Aide demandée (€)	***331 564***

2 Justification scientifique des moyens demandés/*Requested budget : detailed financial plan*

2.1 Partenaire 1/*Partner 1*

Le LPL assurera la coordination globale du projet. Il interviendra dans trois directions : développement informatique, spécifications et descriptions linguistiques, création et annotation de corpus. Pour ce qui concerne la partie informatique, le LPL conduira le développement de la plateforme en adaptant et développant un certain nombre d'outils. Il aura en particulier à sa charge le développement d'outils d'alignement, d'étiquetage prosodique, morpho-syntaxique, syntaxique, sémantique ainsi que les outils de conversion. Par ailleurs, le LPL assurera l'intégration de la plateforme et conduira ses tests. Pour ces interventions, un renfort sous la forme d'un ingénieur informaticien pour 24 mois est demandé.

La partie de validation de la plateforme consistera à créer des corpus et les annoter en totalité. Il s'agit d'un travail considérable, auquel participeront plusieurs membres du LPL, et qui devront être secondés par 4 vacataires, pour une durée totale de 19 mois de vacation. Ces vacataires seront recrutés parmi les étudiants de master ou de doctorat du LPL.

2.1.2 Equipement/*Large equipment*

Le matériel demandé concerne d'une part l'acquisition de données (banc audio-vidéo), leur traitement et leur stockage. Il s'agit de matériel standard, complétant celui déjà à disposition dans l'unité.

Matériel	Destination	Quantité	Coût
Station de traitement vidéo	Annotation	1	3 000
Stockage de données	Stockage	1	3 000
Station de travail	Traitement	2	4 000
Acquisition audio-vidéo	Acquisition de données	1	2 500
Total			12 500

2.1.3 Personnel/*Manpower*

Le personnel nécessaire à la réalisation du projet concerne d'une part le développement des outils de traitement et d'autre part la création de ressources.

Pour ce qui concerne le traitement, un ingénieur informaticien, ayant la connaissance du traitement de données XML et une expérience de développement en Java sera recherché. Nous avons choisi Java d'une part pour assurer l'interopérabilité des plateformes et d'autre part pour assurer la continuité avec les outils déjà développés dans ce domaine. De plus, leur distribution en sera facilitée.

Par ailleurs, le travail de validation reposera essentiellement sur l'utilisation des outils développés dans le cadre de la création de nouvelles ressources. Nous demandons pour cela des vacations dont lé bénéficiaires seront des étudiants de linguistique.

Type	Profil	Durée
Ingénieur	Informaticien	12 mois
Vacataires	Master linguistique	18 mois

2.1.4 Missions/*Travels*

Concernant les missions, le LPL assurera l'organisation de 2 réunions par an du consortium. Régulièrement, et à l'occasion de ces réunions, un collègue étranger sera invité à y participer et effectuer un séjour de recherche. Par ailleurs, des déplacements pour la présentation des résultats lors de conférences internationales sont également prévues.

Type		Montant
Réunions du consortium	6 réunions (2 par an)	6 000
Conférences internationales	6 (2 par an)	5 000
Chercheurs extérieurs invités	3 (1 par an)	2 500
Total		13 500

2.2 Partenaire 2/*Partner 2*

Le LIMSI interviendra en particulier sur la spécification des formats, l'annotation des corpus, le développement d'outils de conversion et la validation des données produites. Il mettra à profit son expérience dans ce domaine et ses contacts internationaux (notamment pour l'annotation gestuelle) sur ces tâches. Par ailleurs, le LIMSI coordonnera plus particulièrement les travaux de validation. Il demande pour cela un post-doc pour les phases de validation et d'exploitation des ressources expérimentales produites ainsi que pour l'adaptation des outils développés.

Les missions envisagées par le LIMSI portent sur l'accueil de partenaires internationaux et la participation à des conférences sur le domaine. En termes de fonctionnement, le LIMSI demande l'acquisition d'un ordinateur portable destiné à l'annotation vidéo et facilitant la coopération entre les partenaires.

2.2.1 Equipement/*Large equipment*

Le matériel demandé concerne l'acquisition de données, ainsi que leur traitement. Ce matériel sera nécessaire aux phases de développement et de validation de la plateforme.

Matériel	Destination	Quantité	Coût
Station de traitement vidéo	Annotation	1	3 000
Station de travail	Traitement	2	4 000
Acquisition audio-vidéo	Acquisition de données	1	2 500
Total			9 500

2.2.2 Personnel/*Manpower*

La participation du LIMSI sera complétée par un post-doc et des vacations. Le sujet du post-doc, en informatique, portera plus particulièrement sur les questions d'interaction multimodale. Il s'agira, sur la base des données recueillies, d'examiner précisément les types d'interaction de façon à préciser le mode de synchronisation utilisé. Cette question est essentielle y compris du point de vue de la représentation des connaissances. Le type d'interaction est en effet variable d'un domaine à un autre. Mais il est indispensable d'en donner une représentation homogène permettant de prendre en compte des ancrages de niveau variable.

Les vacations seront par ailleurs destinées à effectuer l'annotation de nouveaux corpus à l'aide de la plateforme développée.

Type	Profil	Durée
Post Doc	Informatique	12 mois
Vacataires	Linguistique informatique	18 mois

2.2.3 Missions/*Travels*

Les missions prévues dans le cadre de ce projet portent d'une part sur la présentation des résultats lors de conférences internationales et d'autre part sur des séjours effectués dans des équipes de recherche dont les travaux portent sur ce domaine (notamment en Allemagne).

Type		Montant
Conférences internationales	6 (2 par an)	5 000
Collaborations extérieures	3 (1 par an)	3 500
Total		8 500

2.3 Partenaire 3/*Partner 3*

Le LSIS interviendra en amont sur la spécification des formats d'encodage et l'élaboration du schéma MCS, il participera à la réflexion et l'état de l'art des systèmes de requête sur les données multimodales. Son intervention principale portera sur le développement des outils de manipulation de données et de requête. Les trois participants permanents du LSIS ont déjà l'expérience de ce type de développement. Ils seront renforcés dans cette tâche par deux ingénieurs pour une durée de 8 mois. La partie validation (et adaptation des outils créés) sera assurée par les permanents.

Le LSIS demande par ailleurs un soutien pour participer à des missions lui permettant de présenter le résultat de ses travaux.

2.3.1 Equipement/*Large equipment*

L'intervention du LSIS portera essentiellement sur le développement d'outils de manipulation de données. Nous avons pour cela besoin de stations de travail complémentaires, dotées de capacité de stockage adéquate compte tenu du type de données à traiter.

Matériel	Destination	Quantité	Coût
Station de travail	Développement	2	3 500
Unités de stockage complémentaire	Stockage	2	1 000
Total			4 500

2.3.2 Personnel/*Manpower*

Le LSIS se propose d'une part d'expérimenter des outils existants dans le cadre du développement d'un environnement de manipulation de données multistructurées. Il s'agit d'un travail à la fois de recherche et d'expérimentation qui sera conduit sur la base des recherches en cours au sein du laboratoire. Cette activité sera appuyée par un post-doc informaticien, spécialisé dans les langages de requêtes et le traitement de données. Par ailleurs, le développement nécessitera le complément d'intervention d'un ingénieur pendant une période de 8 mois.

Type	Profil	Durée
Post Doc	Informatique	12 mois
Ingénieur	Informatique	8 mois

2.3.3 Missions/*Travels*

Nous demandons dans le cadre de ce projet un soutien nous permettant de participer à de conférences internationales de façon régulière.

Type		Montant
Conférences internationales	6 (2 par an)	4 500
Total		4 500

Annexes

Description des partenaires/*Partners informations*

Laboratoire Parole et Langage
Joint Research Unit : Université de Provence/CNRS
Aix-en-Provence

The « Laboratoire Parole et Langage » (Speech and Language Laboratory) in Aix-en-Provence is one of the largest CNRS-University research Centres in France for the study of speech and language. It brings together an international, multidisciplinary community of researchers conducting basic research on spoken and written language. Research areas cover a variety of domains ranging from the cognitive aspects of language to the instrumental investigation of production and perception processes of normal and disordered speech.

The staff of about 80 persons includes engineers and technicians (20), computer scientists (15), phoneticians (20), psychologists (10), linguists (10) and clinicians (5). In addition, 60 PhD students are currently involved in various doctoral projects.

The laboratory is equipped with state-of-the-art equipment for speech recording, extensive software for signal processing and speech analysis. The laboratory at Aix has in particular developed specific tools for digital multisensor multichannel acoustic, physiological and aerodynamic data recording. This workstation is equally used in a number of European clinics as an evaluation and monitoring tool of speech therapy programme. A long tradition of research in the field of speech prosody resulted in the creation in 2000 of an international special interest group (SProSIG) affiliated to ISCA and which launched a biennial series of International conferences (2002: Aix-en-Provence, France; 2004 Nara, Japan; 2006 Dresden, Germany; 2008 Campinas, Brazil)

The laboratory has a long-standing tradition of cooperation with industry and has been involved in the last ten years in a large number of collaborative actions, including European ICT projects. It is also presently host to Marie Curie incoming international fellowships.

LIMSI
UPR CNRS
Orsay

Le LIMSI couvre les différentes compétences nécessaires à la collecte, l'annotation et la modélisation de comportements multimodaux et émotionnels. Le LIMSI a organisé les workshops internationaux LREC sur les Corpus Multimodaux en 02, 04, 06, 08. Deux workshops sur les corpus émotionnels sont également organisés en 06 et 08. Un numéro spécial de la revue Language Resource and Evaluation sur les Corpus Multimodaux va sortir en 2008.

LSIS
UMR CNRS
Toulon

Les travaux de l'équipe INCOD concernent l'ingénierie des systèmes d'information et de connaissance en s'intéressant plus particulièrement à leurs aspects distribués, coopératifs et de sûreté. L'objectif d'un système d'intégration est de fournir un accès transparent et efficace à une multitude de sources d'informations généralement hétérogènes et distribuées sur un réseau. Cet accès est fourni à travers une vue uniforme globale (schéma ou ontologie) sur les sources (schémas locaux), par l'intermédiaire d'un langage de requêtes (global). Ceci implique principalement la prise en compte des problèmes suivants :

- le choix d'une architecture d'intégration virtuelle – de médiation – ou matérialisée (entrepôt),
- la construction d'un couple « modèle d'intégration, modèle de requêtes » adapté : des solutions classiques d'intégration ont été proposées dans un cadre « relationnel, SQL » ou « logique, Datalog »,
- la définition et la gestion d'un ensemble de règles de transformations ou mappings qui permettent la récriture d'une requête globale en termes des requêtes locales,
- la gestion d'une évolutivité due à l'ajout de nouvelles sources et/ou de nouvelles ressources (outils par exemple), dans un contexte Web notamment.

La recherche développée dans ce thème s'appuie sur trois éléments : la culture BDSI (bases de données et systèmes d'information) de l'équipe, nos travaux antérieurs dans le domaine de l'interopérabilité des SI, et les résultats de projets similaires ou de référence (TSIMMIS, Infosleuth, PICSEL, etc.).

ANNEXE II : Biographies/*Résumés and CV*

PARTENAIRE 1

CURRICULUM VITAE

Mr BLACHE Philippe

Né le 6 Octobre 1959

Téléphone : 04 42 95 36 25
Fax : 04 42 95 37 44
E-mail : pb@lpl.univ-aix.fr
LPL, Université de Provence
29 av. Robert Schuman,
13621 Aix en Provence Cedex 1

SITUATION ACTUELLE

Directeur de Recherche au CNRS
Directeur du LPL

Affectation : Laboratoire Parole et Langage, UMR 6057, CNRS et Université de Provence

PARCOURS PROFESSIONNEL

Cursus

- Habilitation à diriger des recherches (1999), Université Paris 7
- Doctorat d'Informatique (1990), Université Aix-Marseille 2
- DEA de Mathématiques et Informatique (1987), Université Aix-Marseille 2
- DESS d'Informatique (1986), Université Aix-Marseille 2
- DEA de Linguistique (1985), Université Aix-Marseille 2

Expérience Professionnelle

- 2002 – Directeur de recherche au CNRS (LPL)
- 1992 – 2002 Chargé de Recherche, 2LC (Sophia) puis LPL (Aix)
- 1990 – 1992 Maître de Conférences (Université de Neuchâtel, Suisse)

Principales Responsabilités actuelles ou passées

- Directeur du LPL (UMR 6057)
- Directeur du 2LC (URA1235)
- Président du réseau Cognisud
- Membre du bureau d'organismes internationaux (EACL, ESSLLI)
- Rédacteur en chef de la revue TAL
- Membre du comité de direction de : ILF, I3
- Organisateur de 7 conférences nationales, 6 conférences internationales (dont 3 à l'étranger), 1 école d'été internationale (500 participants).
- Responsable de 15 projets de recherche

PUBLICATIONS

Les cinq publications les plus récentes :

[1] Philippe Blache (2006), "A Robust and Efficient Parser for Non-Canonical Inputs", in proceedings of ROMAND-06

[2] Philippe Blache, Stéphane Rauzy (2006), "Mécanismes de contrôle pour l'analyse en Grammaires de Propriétés", in actes de TALN-06

[3] Philippe Blache, Barbara Hemforth & Stéphane Rauzy (2006), "Acceptability Prediction by Means of Grammaticality Quantification" , in proceedings of COLING-ACL 06

[4] Tristan Vanrullen, Philippe Blache, Jean-Marie Balfourier (2006), "Constraint-Based Parsing as an Efficient Solution: Results from the Parsing Evaluation Campaign EASy", in proceedings of LREC-06

[5] Omar Nouali, Philippe Blache (2005), "Filtrage automatique de courriels : une approche adaptative et multi niveaux", in Annales des Télécommunications, Hermès, vol. 60, no. 11-12.

Les cinq publications les plus significatives :

[6] P. Blache (2001) *Les Grammaires de Propriétés : des contraintes pour le traitement automatique des langues naturelles*, Hermès Sciences

[7] P. Blache (2003) "Analyse des dépendances à distance à l'aide de graphes de contraintes", in revue *TAL*, 43:3, 63-83

[8] P. Blache (2005), "Property Grammars: A Fully Constraint-Based Theory" , in *Constraint Solving and Language Processing,* H. Christiansen & al. (eds), LNAI 3438, Springer, 1-16

[9] P. Blache & D. Hirst (2001) "Aligning Prosody and Syntax in Property Grammars", in proceedings of Eurospeech-2001.

[10] P. Blache & A. Di Cristo (2002) "Variabilité et dépendances des composants linguistiques", in actes de TALN-2002.

CURRICULUM VITAE

Me BERTRAND Roxane

Née le 11 février 1970

Téléphone : 04 42 95 36 29
Fax : 04 42 95 37 88
e-mail : roxane.bertrand@lpl.univ-aix.fr
adresse : LPL, Université de Provence
29 av. Robert Schuman,
13621 Aix en Provence Cedex 1

SITUATION ACTUELLE

Chargée de Recherche au CNRS, classe 1

Affectation : Laboratoire Parole et Langage, UMR 6057, CNRS et Université de Provence

PARCOURS PROFESSIONNEL

2006- :	Chargée de recherche, 1ère Classe, CNRS, Laboratoire Parole et Langage
depuis 1999 :	Chargée de recherche, 2ième classe CNRS, Laboratoire Parole et Langage
1998-1999 :	Attachée Temporaire de Recherche, Université de Provence, Sciences du Langage
1999 :	Thèse de Doctorat, Sciences du Langage, Université de Provence, Aix-en-Provence: *De l'hétérogénéité de la parole: analyse énonciative de phénomènes prosodiques et kinésiques dans l'interaction interindividuelle*
1996-1998 :	Chargée de cours, Département Sciences du Langage, Université de Provence,
1995 :	DEA de Linguistique Générale, Université de Provence, Aix-en-Provence
1994 :	DEA de Phonétique Expérimentale, fonctionnelle et appliquée, Université de Provence, Aix-en-Provence

PUBLICATIONS

Les cinq publications les plus récentes :

[1] Bertrand R.; Ferré, G.; Blache, P.; Espesser, R.; Rauzy, S. (2007): "Backchannels revisited from a multimodal perspective". *Proceedings of Auditory-visual Speech Processing* Hilvarenbeek, The Netherlands, Cederom, 2007, 5 pages.
[2] Portes C. & Bertrand R., Espesser R., "Contribution to a grammar of intonation in French. Form and function of three rising patterns". *Nouveaux Cahiers de Linguistique Française*, 28, 2007,155-162.
[3] Ferré G., Bertrand R., Blache P., Espesser R., Rauzy S., "Intensive Gestures in French and their Multimodal Correlates", *Proceedings of Interspeech*, Antwerp, Belgium, 2007, 690-693.
[4] D'Imperio M., Bertrand R., Di Cristo A. & Portes C., "Investigating phrasing levels in French: Is there a difference between nuclear and prenuclear accents?" Selected Papers from LSRL, Rutgers, [A paraître]
[5] Bertrand R.; Blache P.; Espesser R.; Ferré G.; Meunier C.; Priego-Valverde B.; Rauzy S., "Le CID - Corpus of Interactional Data -: protocoles, conventions, annotations", TIPA, 25, 2007, 25-55.

Les cinq publications les plus significatives :

[1] D'Imperio M., Bertrand R., Di Cristo A. & Portes C., "Investigating phrasing levels in French: Is there a difference between nuclear and prenuclear accents?" Selected Papers from LSRL, Rutgers, [A paraître]
[2] Portes C. & Bertrand R., 2006, "Some cues about the interactional value of the " continuation " contour in French", *Proceedings of Discourse-Prosody Interface Symposium*, 2006, Cederom, 14 pages.
[3] Bertrand R. & Chanet C., , "Fonctions pragmatiques et prosodie de *enfin* en français spontané ", *Revue de Sémantique et Pragmatique* 17, 2005, 41-68.
[4] Di Cristo A., Auran C., Bertrand R., Chanet C., Portes C., "Outils prosodiques et analyse du discours" , *Cahiers de l'Institut de Linguistique de Louvain* 28, 2004, 27-84
[5] Bertrand, R.; Espesser, R., "Voice diversity in conversation : a case study", In B. Bel; I. Marlien (eds.) *Proceedings of Speech Prosody 2002*. Aix-en-Provence: Laboratoire Parole et Langage, 2002, 171-174. + Cederom

CURRICULUM VITAE

Mr RAUZY Stéphane

Né le 11 février 1966

Téléphone : 04 42 95 36 36
Fax : 04 42 95 37 88
e-mail : stephane.rauzy@lpl.uni-aix.fr
adresse : LPL, Université de Provence
29 av. Robert Schuman,
13621 Aix en Provence Cedex 1

SITUATION ACTUELLE

Ingénieur de Recherche au CNRS
Chef de projet en développement d'applications

Affectation : Laboratoire Parole et Langage, UMR 6057, CNRS et Université de Provence

PARCOURS PROFESSIONNEL

2006- :	Ingénieur de Recherche au CNRS, Laboratoire Parole et Langage
2004-2005 :	Chef de projet R&D et responsable administratif, société AEGYS
2002-2003 :	Ingénieur contractuel au CNRS en Informatique (Laboratoire Parole et Langage, université de Provence)
2001-2002 :	Ingénieur R&D en informatique (société Sémantia, détaché au Laboratoire Parole et Langage, université de Provence)
1999-2000 :	Research Assistant (Department of Physics and Astronomy, University of Glasgow, Scotland, UK)
1995-1998 :	Séjour postdoctoral (Centre de Physique Théorique, Marseille)
1993-1995 :	Attaché Temporaire d'Enseignement et de Recherche en Mathématiques (Université de Provence, Marseille)
1992-1993 :	Research Associate (Department of Physics, Queen's University, Kingston, Ontario, Canada)
1989-1992 :	Thèse de Doctorat (Service d'Astrophysique, CEA Saclay)
1984-1988 :	Cursus de formation universitaire en Science Physique (DEA de Physique Théorique)

PUBLICATIONS

Les cinq publications les plus récentes :

[45] HIRST, D.; ALI, S.; CHENTIR, A.; CHO, H.; NESTERENKO, I.; RAUZY, S. The AANVIS project: towards the automatic multilingual analysis of non-verbal information in speech. Proceedings of ICPhS 2007, août 6-10 : Saarbrücken, GERMANY). 2007, p. 1-2.
[44] BLACHE, P.; FERRÉ, G.; RAUZY, S. An XML Coding Scheme for Multimodal Corpus Annotation. Proceedings of Corpus Linguistics (4 : 2007 juillet 27-30 : Birmingham, UNITED KINGDOM). 2007, p. 1-17.
[43] ADER, M.; BLACHE, P.; RAUZY, S. Souris et claviers virtuels pour le contrôle alternatif de l'environnement informatique. Actes, ASSISTH 2007, novembre 19-21 : Toulouse, FRANCE, p. 1-8.
[42] NESTERENKO, I.; RAUZY, S. On the use of probabilistic grammars in speech annotation and segmentation tasks. Proceedings of Speech and Computer (12 : 2007 octobre 15-18 : Moscow, RUSSIA). 2007, p. 1-7.
[41] FERRÉ, G.; BERTRAND, R.; BLACHE, P.; ESPESSER, R.; RAUZY, S. Intensive Gestures in French and their Multimodal Correlates. Proceedings of Interspeech (2007 août 27-31 : Antwerp, BELGIUM). . Interspeech, Antwerp, Belgium: Interspeech. 2007, p. 690-693.

Les cinq publications les plus significatives :

[39] BLACHE, P.; RAUZY, S. Le module de reformulation iconique de la Plateforme de Communication Alternative. TALN 2007, juin 5-8 : Toulouse, FRANCE, p. 519-528.
[35] Blache P., Hemforth B., Rauzy S., 2006 ; ACL 2006, 17-21 july 2006, Sydney, Australia : « Acceptability Prediction by Means of Grammaticality Quantification »
[33] Blache P., Rauzy S., 2006 ; TALN 2006, 10-13 avril 2006, Leuven, Belgique : « Mécanismes de contrôle pour l'analyse en Grammaires de Propriétés »
[22] Rauzy S., 2001; Monthly Notices of the Royal Astonomical Society, 324, 51 : "A simple tool for assessing the completeness in apparent magnitude of magnitude redshift sample."
[2] Rauzy S., Lachièze-Rey M., Henriksen R.N., 1992; Astonomy&Astrophysics, 256, 1 : "Detecting non-Hubble velocity fields in the universe."

CURRICULUM VITAE

Mme MEUNIER, Christine

Né le 18 juillet 1963

Téléphone : 04 42 95 36 39
Fax : 04 42 95 37 88
e-mail : christine.meunier@lpl.univ-aix.fr
adresse : LPL, Université de Provence
29 av. Robert Schuman,
13621 Aix en Provence Cedex 1

SITUATION ACTUELLE

Chargée de Recherche au CNRS depuis 1998

Affectation : Laboratoire Parole et Langage, UMR 6057, CNRS et Université de Provence

PARCOURS PROFESSIONNEL

2000-2002 : ACI Cognitique (Ministère de la recherche) *"Densité des systèmes vocaliques et traitement cognitif des unités sonores dans différentes langues"* - Responsable scientifique: Christine MEUNIER, Laboratoire Parole et Langage, CNRS UMR 6057, Université de Provence, Aix

1994-1998 : Projet financé par Fond National de la Recherche Scientifique (Suisse): *"La perception du langage parlé: la reconnaissance des mots"* - Responsable scientifique: U.H. FRAUENFELDER, Laboratoire de Psycholinguistique, Faculté de Psychologie, Université de Genève, Suisse

1992-1993 : Contrat entre Thomson-C.S.F. et le C.N.R.S.: *"Evaluation de l'intelligibilité de 9 types de vocoders"* (matériel de transmission) - Laboratoire "Parole et Langage", URA 261 CNRS, Université de Provence, France

1991-1993 : Contrat entre l'Institut National de Plongée Professionnelle (I.N.P.P.) et le C.N.R.S: *"Evaluation de l'intelligibilité de la parole en milieu hyperbare"* - Laboratoire "Parole et Langage", URA 261 CNRS, Université de Provence, France

1989-1991 : Projet européen ESPRIT-SAM (2589), section Linguistic Tools: *"Segmentation et étiquetage du corpus français de la Base de Données Européenne"* - Laboratoire "Parole et Langage", URA 261 CNRS, Université de Provence, France

PUBLICATIONS

Les cinq publications les plus récentes :

[11] Meunier, C.; Espesser, R.; Frenck-Mestre, C. (2006) "Aspects phonologique et dynamique de la distinctivité au sein des systèmes vocaliques: une étude inter-langue". *JEP*, 2006 juin 12-16, Dinard, FRANCE, p. 333-336.

[12] Frenck-Mestre, C.; Meunier, C.; Espesser, R.; Holcomb, P.; Daffner, K. (2005) "Perceiving Nonnative Vowels: The Effect of Context on Perception as evidenced by Event-Related Brain Potentials". *Journal of Speech Language and Hearing Research*, vol. 48, p. 1496-1510.

[13] Meunier C. (2005) "Invariants et Variabilité en Phonétique", *chap. 13, in Phonologie et phonétique (Traité IC2, série Cognition et Traitement de l'Information)*, N. Nguyen, S. Wauquier-Gravelines, J. Durand, Hermès, 350-374.

[14] Meunier C., Frenck-Mestre C., Lelekov-Boissard T., Le Besnerais M. (2004) "La perception des systèmes vocaliques étrangers: une étude inter-langues", *Journées d'Etudes sur la parole*, Avril 2004, Fès, Maroc.

[15] Blache, P. & Meunier, C. (2004) "Domaines et propriétés : une description de la répartition de l'information linguistique", *Journées d'Etudes Linguistiques*, Mai 2004, Nantes.

Les cinq publications les plus significatives :

[16] Frenck-Mestre, C.; Meunier, C.; Espesser, R.; Holcomb, P.; Daffner, K. (2005) "Perceiving Nonnative Vowels: The Effect of Context on Perception as evidenced by Event-Related Brain Potentials". *Journal of Speech Language and Hearing Research*, vol. 48, p. 1496-1510.

[17] Meunier C. (2005) "Invariants et Variabilité en Phonétique", *chap. 13, in Phonologie et phonétique (Traité IC2, série Cognition et Traitement de l'Information)*, N. Nguyen, S. Wauquier-Gravelines, J. Durand, Hermès, 350-374.

[18] Meunier C., Frenck-Mestre C., Lelekov-Boissard T., Le Besnerais M. (2003) "Production and perception of foreign vowels: does the density of the system play a role?", *Proceedings of the 15th ICPhS*, Barcelona, Spain.

[19] Hallé P., Segui J., Frauenfelder U., Meunier C. (1998) "The Processing of Illegal Consonant Clusters: A Case of Perceptual Assimilation", *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 24, No. 2, 592-608.

[20] Marchal A., Meunier C. (1993) "A Database of Subaquatic and Hyperbaric Speech: PSH/DISPE", *Journal of the Acoustical Society of America*, pp 2990-2993.

CURRICULUM VITAE

Mlle ADER Morgane

Née le 12 novembre 1982

Téléphone : 04 42 95 36 36
Fax : 04 42 95 37 88
e-mail : morgane.ader@gmail.com
adresse : LPL, Université de Provence
29 av. Robert Schuman,
13621 Aix en Provence Cedex 1

SITUATION ACTUELLE

Ingénieur de Recherche CDD à l'université Aix-Marseille I
Informatique

Affectation : Laboratoire Parole et Langage, UMR 6057, CNRS et Université de Provence

PARCOURS PROFESSIONNEL

2007-2008 :	Ingénieur contractuel à l'université Aix-Marseille I en informatique, Laboratoire Parole et Langage.
2006-2007 :	Ingénieur contractuel au CNRS en informatique, Laboratoire Parole et Langage.
2006 :	Stage de fin d'études de 6 mois en informatique, société AEGYS.
2004-2006 :	Cycle ingénieur, spécialité Logiciel et Système Informatique (ENSSAT - École Nationale Supérieure des Sciences Appliquées et de Technologie, Lannion).
2000-2004 :	Maîtrise, Licence et DEUG MASS (Mathématiques Appliquées et Sciences Sociales - Université Toulouse II).

PUBLICATIONS

[1] Bertrand R.; Ader M.; Blache P.; Ferré G.; Espesser R.; Rauzy S., 2007 ; Ecole thématique CONTACI, 4-8 juin 2007, Lyon, France, « Représentation, édition et exploitation de données multimodales : le cas des backchannels dans le corpus CID »

[2]Ader M.; Blache P.; Rauzy S.,2007 ; Conférence internationale ASSISTH', 19-21 novembre 2007, Toulouse, France, « Souris et claviers virtuels pour le contrôle alternatif de l'environnement informatique ».

CURRICULUM VITAE

Mr ESPESSER Robert

Né le 22 Septembre 1951

Téléphone : 04 42 95 36 26
Fax : 04 42 95 37 88
e-mail : Robert.Espesser@lpl.univ-aix.fr
adresse : LPL, Université de Provence
29 av. Robert Schuman,
13621 Aix en Provence Cedex 1

SITUATION ACTUELLE

Ingénieur de Recherche au CNRS
Affectation : Laboratoire Parole et Langage, UMR 6057, CNRS et Université de Provence

PARCOURS PROFESSIONNEL

2002- : Ingénieur de Recherche au CNRS, Laboratoire Parole et Langage
1977-2002 : Ingenieur d'étude au CNRS, Laboratoire Parole et Langage

PUBLICATIONS

Les cinq publications les plus récentes :

[Bertrand R.; Blache P.; Espesser R.; Ferré G.; Meunier C.; Priego-Valverde B.; Rauzy S., 2007, TIPA, vol. 25. 2007, p. 25-55, « Le CID - Corpus of Interactional Data -: protocoles, conventions, annotations »

GHIO, A.; MEYNADIER, Y.; TESTON, B.; LOCCO, J.; CLAIRET, S.; ESPESSER, R.; MEUNIER, C., MARLIEN, I., DENIAUD, C. Peut-on parler sous l'eau avec un embout de détendeur ? Etude articulatoire et perceptive. Actes des Journées d'Etudes sur la Parole (JEP) (26 : 2006 juin 12-16 : Dinard) Rennes: Irisa, Afcp, Isca. 2006, p. 379-382.

MEUNIER, C.; ESPESSER, R.; FRENCK-MESTRE, C. Aspects phonologique et dynamique de la distinctivité au sein des systèmes vocaliques: une étude inter-langue. Journées d'Etude sur la Parole (JEP) (26 : 2006 juin 12-16 : Dinard, FRANCE). . Rennes: Irisa, Afcp, Isca. 2006, p. 333-336.

MEUNIER, C.; ESPESSER, R.; FRENCK-MESTRE, C. Phonetic variability as a static/dynamic process in speech communication: a cross linguistic study. Laboratory Phonology (LabPhon) (10 : 2006 juin 29-juillet 1 : Paris, FRANCE).

PASDELOUP, V.; ESPESSER, R.; FARAJ, M. Rate sensitivity of syllable in French: a perceptual illusion? Proceedings of Speech Prosody (3 : 2006 mai 2-5 : Dresden, GERMANY). 2006. Cederom.

Les cinq publications les plus significatives depuis 5 ans

[D'IMPERIO, M.; ESPESSER, R.; LŒVENBRUCK, H.; MENEZES, C.; NGUYEN, N.; WELBY, P. Are tones aligned with articulatory events? Evidence from Italian and French. In Cole, Jennifer; Hualde, José I. (eds.) Papers in Laboratory Phonology 9.
Berlin: Mouton de Gruyter. 2007, p. 577-608.

PASDELOUP, V.; ESPESSER, R.; FARAJ, M. Rate sensitivity of syllable in French: a perceptual illusion? Proceedings of Speech Prosody (3 : 2006 mai 2-5 : Dresden, GERMANY). 2006

FRENCK-MESTRE, C.; MEUNIER, C.; ESPESSER, R.; HOLCOMB, P.; DAFFNER, K. Perceiving Nonnative Vowels: The Effect of Context on Perception as evidenced by Event-Related Brain Potentials. Journal of Speech Language and Hearing Research, vol. 48. 2005, p. 1496-1510.

NGUYEN, N.; ESPESSER, R.; MEUNIER, C. Etude comparative de la structure acoustique des systèmes de voyelles: premiers résultats. Colloque international Phonologie et phonétique du français: données et théories (2003 décembre 11-13 : Paris, FRANCE). 2003, p. 1-32.

REY, V.; DE MARTINO, S.; ESPESSER, L.; HABIB, M. Temporal Processing and Phonological Impairment in Dyslexia : Effect of Phoneme Lenghtening on Order Judgement of Two Consonants. Brain and Language, no. 80. 2002, p. 576-591.

CURRICULUM VITAE

Mr HIRST Daniel

Né le 29 septembre 1946

Téléphone : 04 42 95 36 28
Fax : 04 42 95 37 88
e-mail : daniel.hirst@lpl.univ-aix.fr
adresse : LPL, Université de Provence
29 av. Robert Schuman,
13621 Aix en Provence Cedex 1

SITUATION ACTUELLE

Directeur de Recherche première classe au CNRS
Responsable de l'équipe Prosodie et Représentation Formelle du Langage

Affectation : Laboratoire Parole et Langage, UMR 6057, CNRS et Université de Provence

PARCOURS PROFESSIONNEL

2005-2007	Directeur de Recherche 1e classe CNRS
1990-2005	Directeur de Recherche 2e classe CNRS
2002-2004	Chargé de Cours, UFR DEMA (Anglais) Université de Provence.
1982-1990	Chargé de Recherche CNRS
1977-1982	Attaché de Recherche CNRS
1971-1973	Chargé de cours, IUT, Aix-en-Provence
1970-1977	Lecteur de langues étrangères, Université de Provence

PUBLICATIONS

Les cinq publications les plus récentes :

[21] Hirst, D.J. 2006. The prosody of speech and language. in K. Brown (ed.) *Encyclopaedia of Language and Linguistics*. 2nd edition. Oxford, Oxford University Press., vol. 10. 167-178.

[22] Hirst, D.J. 2005. Form and function in the representation of speech prosody. in K.Hirose, D.J.Hirst & Y.Sagisaka (eds) *Quantitative prosody modeling for natural speech description and generation* (=*Speech Communication* 46 (3-4)), 334-347.

[23] Hirst, D.J.; Auran, C. 2005. Analysis by synthesis of speech prosody: the ProZed environment. *Interspeech/Eurospeech 05. 9th European Conference on Speech Communication and Technology*, September 2005, Lisbon. 3225-3228.

[24] Hirst, D.J.; Bouzon, C. 2005. The effect of stress and boundaries on segmental duration in a corpus of authentic speech (British English). *Proceedings of Interspeech/Eurospeech 05. 9th European Conference on Speech Communication and Technology*, September 2005, Lisbon. 29-32.

[25] Auran, C.; Bouzon, C.; Hirst, D.J. 2004. The Aix-MARSEC project: an evolutive database of spoken British English. in *Proceedings of Second International Conference on Speech Prosody*, Nara March 2004. 561-564

Les cinq publications les plus significatives :

[26] Hirst, D.J. 2004. Speech prosody: from acoustics to interpretation. in G.Fant, H.Fujisaki, J.Cao & Y.Xu (eds) *From Traditional Phonology to Modern Speech Processing (Festschrift for Professor Wu Zongji's 95th Birthday)* Foreign Language Teaching and Research Press, Beijing, 177-188

[27] Hirst, D.J. 2001. Automatic analysis of prosody for multi-lingual speech corpora. in E. Keller, G. Bailly, A. Monaghan, J. Terken & M.Huckvale (eds.) *Improvements in Speech Synthesis*. (London, John Wiley). 320-327.

[28] Campione, E.; Hirst, D.J.; Véronis, J. 2000. Automatic stylisation and symbolic coding of FO: implementations of the INTSINT system. in Botinis, A. (ed.) *Intonation: Research and Applications.* (Dordrecht, Kluwer Academic Publishers.), 185-208

[29] Hirst, D.J., Di Cristo, A. & Espesser, R. 2000. Levels of representation and levels of analysis for intonation. in M. Horne (ed) *Prosody : Theory and Experiment*. Kluwer Academic Publishers, Dordrecht. 51-87

[30] Hirst, D.J. & Di Cristo, A. 1998. A survey of intonation systems. in Hirst & Di Cristo (eds) 1998. *Intonation Systems. A survey of Twenty Languages.* (Cambridge, Cambridge University Press), 1-44.

PARTENAIRE 2

CURRICULUM VITAE

Mr MARTIN Jean-Claude

Né le 26 Octobre 1965

Téléphone : 06 84 21 62 05
Fax : 01 69 85 80 88
e-mail : martin@limsi.fr
adresse :
LIMSI, BP 133, 91403 Orsay Cedex

SITUATION ACTUELLE

Maître de Conférences en Informatique Université Paris 8

Affectation recherche : LIMSI-CNRS
Responsable du thème « Agents Conversationnels Animés », groupe AMI

PARCOURS PROFESSIONNEL

1999 -	:	Maître de Conférences en Informatique, LIMSI-CNRS
2006	:	Habilitation à Diriger des Recherches en Informatique soutenue le 6 décembre 2006 « Multimodal Human-Computer Interfaces and Individual Differences. Perception, representation and generation of situated multimodal behaviors »
1995	:	Thèse de Doctorat en Informatique, Télécom Paris

PUBLICATIONS

Les cinq publications les plus récentes :

Grynszpan, O., Martin, J.-C., Nadel, J. (to appear) Multimedia interfaces for users with high functioning autism: an empirical investigation. International Journal of Human - Computer Studies (IJHCS).

Pelachaud, C., Martin, J.-C., André, E., Chollet, G., Karpouzis, K., Pellé, D. (2007) in Procs of the 7th International Conference on "Intelligent Virtual Agents", Springer, LNCS 4722. Springer on-line proceedings

Martin, J.-C., d'Alessandro, C., Jacquemin, C., Katz, B., Max, A., Pointal, L. and Rilliard, A. (2007). 3D Audiovisual Rendering and Real-Time Interactive Control of Expressivity in a Talking Head. In procs of IVA'2007, LNAI 4722, Springer. C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, D. Pelé (Eds.). p 29-36.

Buisine, S., Aoussat, A. and Martin, J.-C. (2007). Embodied Creative Agents: A social-cognitive framework., In procs of IVA'2007, LNCS 4722, 304–316, Springer. C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, D. Pelé (Eds.).

Schröder, M., Devillers, L., Karpouzis, K., Martin, J.-C., Pelachaud, C., Peter, C., Pirker, H., Schuller, B., Tao, J. and Wilson, I. (2007). What should a generic emotion markup language be able to represent?, 2nd International Conference on Affective Computing and Intelligent Interaction (ACII'2007), Lisbon, Portugal, 12-14 September. Springer, LNCS, vol. 4738. A. Paiva, R. Prada and R. Picard (Eds.), 440-451.

Les cinq publications les plus significatives :

Grynszpan, O., Martin, J.-C., Nadel, J. (2007), Exploring the Influence of Task Assignment and Output Modalities on Computerized training for Autism., International Journal "Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems", John Benjamins Publishing, 8(2): 241-266

Buisine, S., Martin, J.-C. (2007), The effects of speech-gesture co-operation in animated agents' behaviour in multimedia presentations., International Journal "Interacting with Computers: The interdisciplinary journal of Human-Computer Interaction". Elsevier, 19: 484-493.

Martin, J.-C., Caridakis, G., Devillers, L., Karpouzis, K. and Abrilian, S. (2007), Manual Annotation and Automatic Image Processing of Multimodal Emotional Behaviours: Validating the Annotation of TV Interviews. International Journal "Personal and Ubiquitous Computing", Special issue on 'Emerging Multimodal Interfaces' following the special session of the AIAI 2006 Conference, Springer, May.
Martin, J.-C., Niewiadomski, R., Devillers, L., Buisine, S., Pelachaud, C. (2006). Multimodal Complex Emotions: Gesture Expressivity And Blended Facial Expressions. Special issue of the Journal of Humanoid Robotics on "Achieving Human-like Qualities in Interactive Virtual and Physical Humanoids". Eds: C. Pelachaud, L. Canamero. Vol. 3, No. 3,September 2006, 269-291. World Scientific Publishing Co. http://ejournals.wspc.com.sg/ijhr/ijhr.shtml

Martin, J.-C., Buisine, S., Pitel, G., Bernsen, N.O. (2006). Fusion of Children's Speech and 2D Gestures when Conversing with 3D Characters. Special Issue of the Journal Signal Processing on « Multimodal Human-computer Interfaces ». Eds.: Thierry Dutoit, Laurence Nigay and Michael Schnaider. Vol. 86, isssue 12, pp. 3596-3624. Elsevier. http://www.sciencedirect.com/science/journal/01651684

CURRICULUM VITAE

Melle FERRE Gaëlle

Née le 19 septembre 1970

Téléphone : 02 40 14 11 34
e-mail : Gaelle.Ferre@univ-nantes.fr
adresse : Université de Nantes, CIL
Département d'Etudes Anglaises
Chemin de la Censive du Tertre
BP 81227
44312 Nantes cedex 3

SITUATION ACTUELLE

Maître de Conférences – Université de Nantes

PARCOURS PROFESSIONNEL

2007… : Maître de Conférences, Université de Nantes
2006-2007 : Postdoc au CNRS, Laboratoire Parole et Langage
2004-2005 : Lectrice de français dans le *Department of Language and Linguistic Science*, University of York, Grande-Bretagne
2001-2004 : Vacations à l'Université de Nantes et à l'Université de Paris III (Département de Sciences du Langage). Vacations à l'Ecole d'Orthophonie de Nantes
2000-2004 : Thèse de doctorat en Sciences du Langage (Université de Paris 3)
1997-2000 : ATER à l'Université de Nantes
1995-1997 : Vacations à l'Université de Nantes et à l'Université de Paris III
1994-1995 : DEA de Phonétique (Université de Paris 3)
1991-1994 : Cursus de formation universitaire en Anglais (Maîtrise d'anglais, Université de Nantes)
1991-1992 : Assistante de français au Lycée de Pontypridd, Pays de Galles.

PUBLICATIONS

Les 5 publications les plus récentes :

[1] Ferré, G., Bertrand R., Blache P., Espesser R. & Rauzy S., 2007, Interspeech 2007, Anvers, Belgique., p. 690-693, « Intensive gestures in French and their Multimodal Correlates »

[2] Bertrand, R.; Ferré, G.; Blache, P.; Espesser, R.; Rauzy, S., « Backchannels revisited from a multimodal perspective ». *Proceedings of Auditory-visual Speech Processing* Hilvarenbeek, The Netherlands, 2007, Cederom, non paginé.
[3] Bertrand R.; Blache P.; Espesser R.; Ferré G.; Meunier C.; Priego-Valverde B.; Rauzy S., 2007, TIPA, vol. 25. p. 25-55, « Le CID - Corpus of Interactional Data -: protocoles, conventions, annotations »
[4] Ferré G., 2006, in Actes des XXVIèmes Journées d'Etude sur la Parole, Dinard, p. 541-544, « La répétition stylistique en anglais oral »
[5] Ferré G., 2005, , York Papers in Linguistics 2(3), University of York, p. 55-90, « Gesture, Intonation and the Pragmatic Structure of Narratives in British English Conversation »

CURRICULUM VITAE

Mme MOREL Mary-Annick

Née le 29 septembre 1944

Téléphone : 01 46 61 12 15
Fax : 01 46 61 12 15
e-mail : marym@ext.jussieu.fr
adresse : Paris 3 – Censier
Bureau 431
13 rue Santeuil
75005 Paris

SITUATION ACTUELLE

Professeur à l'Université de Paris 3 – Sorbonne Nouvelle

Affectation : Université de Paris 3 – EA1483 Recherche sur le Français Contemporain (RFC)

PARCOURS PROFESSIONNEL

2001- :	Professeur CE à Paris 3
1988-2006 :	Responsable de l'EA1483 Recherche sur le Français Contemporain (RFC)
1983-2001 :	Professeur (2ème puis 1ère classe) à Paris 3
1980 :	Doctorat d'Etat ès Lettres et sciences humaines à Paris 3
1978-1983 :	Maître de Conférences à Paris 3
1972-1978 :	Professeur agrégée détachée dans l'Enseignement Supérieur
1967-1972 :	Professeur agrégée au Lycée
1967 :	Agrégation de Grammaire
1962-1967 :	Formation universitaire en Lettres classiques et en Grammaire à la Sorbonne

PUBLICATIONS

Les cinq publications les plus récentes :

[10] Morel M.-A., 2007, « Stratégies intonatives et discursives du meneur de débat oral dans les médias français », in Actes du Coll. International, Stockholm, 10-12 juin 2005, M. Forsgren, M. Broth, F. Syllander (éds), *Le français parlé des médias*, Université de Stockholm, Romanica Stockholiensia 24 : 537-552.
[9] Morel M.-A., 2007, « Le postrhème dans le dialogue oral en français », *L'Information grammaticale* n°113, mars 2007 : 40-46.
[8] Morel M.-A., 2007, « Un thème, deux thèmes, un préambule ? Intonation, geste et morphosyntaxe dans le dialogue oral en français », *LINX* n°55 « Thème et thématisation », M.-L. Elalouf coor d. : 133-152.
[7] Morel M.-A., 2007, « La reformulation dans le dialogue finalisé en français. Propriétés intonatives et mimico-gestuelles », *Recherches Linguistiques* n°29, Mohammed Kara coord., « Usages et analyses de la reformulation » : 123-144.
[6] Meunier A. et **Morel M.-A.**, 2006, « Enonciation et Intonation. La phrase segmentée selon Charles Bally », in Actes du Colloque, Paris, 2004, J.-L. Chiss (éd.), *Charles Bally : historicité des débats linguistiques et didactiques*, Louvain, Peeters : 93-120.

Les cinq publications les plus significatives :

[5] Danon-Boileau L. et Morel M.-A, 2006, « L'intonation et la gestuelle d'un jeune adulte autiste au cours d'un entretien thérapeutique », in Grossen M. et Salazar Orvig A. (éds), *L'entretien clinique en pratiques. Analyse des interactions verbales d'un genre communicatif hétérogène,* Paris, Belin Sup : 233-250.

[4] Nakahara M. et Morel M.-A., 2006, « Intonation, mimique-gestuelle et morphosyntaxe dans un dialogue en français entre une Japonaise et une Française. Modifications après un an de séjour en France », in M. Faraco (éd.), *Regards croisés sur la classe de langue : Pratique, Méthodes, Théories*, P.U.P. : 285-306.

[3] Danon-Boileau L. et Morel M.-A., 2003, « Le locuteur vicariant », in *Le sujet*, J.-M. Merle coord., Paris-Gap, Ophrys, Bibliothèque de Faits de Langues : 235-246.
[2] Bouvet D. & Morel M.-A. (2002) *Le ballet et la musique de la parole. Le geste et l'intonation dans le dialogue en français*, Paris-Gap, Ophrys, Bibliothèque de Faits de Langues : 134 pages.
[1] Morel M.-A. & Danon Boileau (1998) *Grammaire de l'intonation. L'exemple du français*, Paris-Gap, Ophrys, Bibliothèque de Faits de Langues : 232 pages.

CURRICULUM VITAE

Mr NOCERA Pascal

Né le 23 aout 1963

Téléphone : 04 90 84 35 07
Fax : 04 90 84 35 01
e-mail : pascal.nocera@univ-avignon.fr
adresse : LIA,
Site Agroparc. BP 1228,
84911 Avignon Cedex 9

SITUATION ACTUELLE

Maitre de Conferences
Université d'Avignon et des Pays de Vaucluse

Affectation : Laboratoire d'Informatique d'Avignon

PARCOURS PROFESSIONNEL

1992- : Maître de Conférences (Université d'Avignon)
1989-1992 : Thèse de Doctorat (Université d'Avignon)
1984-1988 : Cursus de formation universitaire en Informatique (DEA de Mathématiques et Informatique,Marseille)

PUBLICATIONS

Les cinq publications les plus récentes :

[31] Christophe Servan, Christian Raymond, Frédéric Béchet et Pascal Nocera, Conceptual decoding from word lattices: application to the spoken dialogue corpus MEDIA, 2006 ISCLP'06, Pittsburgh, USA

[32] B. Lecouteux, G. Linarès, P. Nocera, J.F. Bonastre, Imperfect transcript driven speech recognition, 2006, ISCLP'06, Pittsburgh, USA

[33] Nimaan Abdillahi, Pascal Nocera, Jean-François Bonastre, Automatic transcription of Somali language, 2006, ISCLP'06, Pittsburgh, USA

[34] Benoît Favre, Frédéric Béchet, Pascal Nocera, Robust Named Entity Extraction from Spoken Archives, 2005, Proceedings of HLT-EMNLP'05, Vancouver, Canada

[35] Benoît Favre and Frédéric Béchet and Pascal Nocera, Mining Broadcast News data: Robust Information Extraction from Word Lattices, 2005, Lisboa, Portugal

Les cinq publications les plus significatives :

[36] C. Lévy, G. Linarès, P. Nocera and J.F. Bonastre, Embedded mobile phone digit-recognition, Chapter 7 in Digital Signal Processing for In-Vehicle and Mobile Systems 2, H. Abut, J.H.L. Hansen and K. Takeda (Eds.), ISBN 038733503X, Springer Science, New York, NY, Scheduled for Spring 2006

[37] D. Massonié, P. Nocéra, G. Linarès ; Scalable Language Model Look-Ahead for LVCSR, 2005, InterSpeech'05, Lisboa, Portugal

[38] Matrouf D., Bellot O., Nocera P., Linares G., Bonastre J.-F., Structural Linear Model-Space Transformations for Speaker Adaptation, Eurospeech 2003, Genève, Sept. 2003

[39] Nocera P, Linares G, Massonie D, Lefort L. Phoneme lattice based A* search algorithm for speech recognition, TSD2002, Brno, Rep. Tchèque 2002

[40] Nocera P., Bulot R. Rule-Driven Neural Networks for Speech Recognition ; IEEE, International Neural Networks Conference 90, Paris 1990

CURRICULUM VITAE

Mr **LINARES Georges**

Né le 29 décembre 1968

Téléphone : 04 90 84 35 20
Fax : 04 90 84 35 01
e-mail :linares.georges@univ-avignon.fr
adresse : LIA,
Site Agroparc. BP 1228,
84911 Avignon Cedex 9

SITUATION ACTUELLE

Maitre de Conferences
Université d'Avignon et des Pays de Vaucluses

Affectation : Laboratoire d'Informatique d'Avignon

PARCOURS PROFESSIONNEL

1998- : Maître de Conférences (Université d'Avignon)
1995-1998 : Thèse de Doctorat (Université d'Avignon)
1995 : DEA de Mathématiques et Informatique,Marseille
1994 : Maitrise de Mathématiques apliquées, Toulon, UTV

PUBLICATIONS

Les cinq publications les plus récentes :

[41] B. Lecouteux, G. Linarès, Y. Estève, J. Mauclair, System combination by driven decoding, 2007, ICASSP'07, Hawai, USA

[42] C. Lévy, G. Linarès, P. Nocera and J.F. Bonastre, Embedded mobile phone digit-recognition, Chapter 7 in Digital Signal Processing for In-Vehicle and Mobile Systems 2, H. Abut, J.H.L. Hansen and K. Takeda (Eds.), ISBN

[43] B. Lecouteux, G. Linarès, P. Nocera, J.F. Bonastre, Imperfect transcript driven speech recognition, 2007, ISCLP'06, Pittsburgh, USA

[44] C.Lévy, G. Linarès, J.-F. Bonastre, 2006, « GMM-based acoustic modeling for embedded speech recognition », InterSpeech'06 Pitttsburgh - USA

[45] G. Linarès, C. Lévy, J.C. Plagniol, Estimation rapide de modèles semi-continus discriminants, 2006 JEP'06, Dinard - France

Les cinq publications les plus significatives :

[47] B. Lecouteux, G. Linarès, Y. Estève, J. Mauclair, System combination by driven decoding, 2007, ICASSP'07, Hawai, USA

[48] C. Lévy, G. Linarès, P. Nocera and J.F. Bonastre, Embedded mobile phone digit-recognition, Chapter 7 in Digital Signal Processing for In-Vehicle and Mobile Systems 2, H. Abut, J.H.L. Hansen and K. Takeda (Eds.), ISBN 038733503X, Springer Science, New York, NY, Scheduled for Spring 2006

[49] D. Massonié, P. Nocéra, G. Linarès ; Scalable Language Model Look-Ahead for LVCSR, 2005, InterSpeech'05, Lisboa, Portugal

[50] Matrouf D., Bellot O., Nocera P., Linares G., Bonastre J.-F., Structural Linear Model-Space Transformations for Speaker Adaptation, Eurospeech 2003, Genève, Sept. 2003

[51] Nocera P, Linares G, Massonie D, Lefort L. Phoneme lattice based A* search algorithm for speech recognition, TSD2002, Brno, Rep. Tchèque 2002

PARTENAIRE 3

CURRICULUM VITAE

Murisasco Elisabeth

23/02/65

Téléphone : 04 94 14 26 21
Fax : 04 94 14 28 97
e-mail : murisasco@univ-tln.fr
adresse :
Université du Sud Toulon-Var (USTV)
Avenue de l'université
BP 20132
83957 La Garde cedex

SITUATION ACTUELLE

Professeur à l'Université du Sud Toulon Var (section 27)

Recherche : LSIS (Laboratoire des Sciences de l'information et des systèmes) - UMR 6168
Enseignement : UFR de Sciences et Techniques de l'USTV
Thèmes de recherche : Bases de données (BD), BD XML.
Mots-clés : Modèle de données et langage d'interrogation

PARCOURS PROFESSIONNEL

Depuis septembre 2007 : Professeur (USTV)
Octobre 1992 – septembre 2007 : MCF Informatique (USTV)

PUBLICATIONS

Les cinq publications les plus récentes :

- E. Bruno, E. Murisasco, "An XML environment for multistructured textual documents", in: Proceedings of the Second International Conference on Digital Information Management (ICDIM'07), Lyon, october 2007.
- E. Bruno, S. Calabretto, E. Murisasco (2007) ``Documents textuels multistructurés : un état de l'art'', in: Revue I3 (Information - Interaction - Intelligence), mars 07.
- E. Bruno, E. Murisasco (2006) , MultiStructured XML textual documents, GESTS International Transactions on Computer Science and Engineering, Vol 24, n°1 – novembre 2006.
- E. Bruno, E. Murisasco (2006) ``MSXD : a model and a schema for concurrent structures defined over the same textual data'', in: Proceedings of the DEXA Conference, Lecture Notes in Computer Sciences 4080, pp. 172-181, Krakow, Poland.
- E. Bruno, E. Murisasco (2006) ``Describing and Querying hierarchical structures defined over the same textual data'', in: Proceedings of the 2006 ACM Symposium on Document Engineering (DocEng 2006), pp. 147-154, Amsterdam, The Netherlands.

Les cinq publications les plus significatives :

- E. Bruno, E. Murisasco (2006) ``Describing and Querying hierarchical structures defined over the same textual data'', in: Proceedings of the 2006 ACM Symposium on Document Engineering (DocEng 2006), pp. 147-154, Amsterdam, The Netherlands.
- E. Bruno, J. Le Maitre and E. Murisasco (2003), Extending XQuery with Transformation Operators, Proceedings of the 2003 ACM Symposium on Document Engineering (DocEng 2003), ACM Press, Grenoble, France, November 20-22 2003, pp. 1-8.
- Y. Marcoux, J. Le Maitre, E. Murisasco, SgmlQL + XGQL = Powerful XML pattern-matching and data-manipulation in a single language, Actes de la 6ème conf. internationale Recherche d'Informations Assistée par Ordinateur (RIAO' 2000), vol. 2, pp.1346-1362, Avril 2000, Paris, France.
- P. Ramadour, E. Murisasco, J. Le Maitre, Optimization of distributed SgmlQL queries, Networking and Information Systems Journal (NIS), volume 2 n°2/1999, Hermès, pp 181-201 (sélection des meilleurs articles de INFORSID' 98).
- J. Le Maitre, E. Murisasco, M. Rolbert, From annotated corpora to Databases : the SgmlQL language, in Linguistic Databases, CSLI Lectures notes number 77 pp 37-58, ed. J. Nerbonne, janvier 1998.

CURRICULUM VITAE

Bruno Emmanuel

30/08/1975

Téléphone : 04 94 14 20 33
Fax : 04 94 14 28 97
e-mail : bruno@univ-tln.fr
adresse :
Université du Sud Toulon-Var (USTV)
Avenue de l'université
BP 20132
83957 La Garde cedex

SITUATION ACTUELLE

MCF

Recherche : LSIS (Laboratoire des Sciences de l'information et des systèmes) - UMR 6168
Enseignement : UFR de Sciences et Techniques de l'USTV
Thèmes de recherche : Bases de données (BD), BD XML.
Mots-clés : Modèle de données et langage d'interrogation

PARCOURS PROFESSIONNEL

Depuis octobre 2002 : MCF Informatique (USTV)

PUBLICATIONS

Les cinq publications les plus récentes :

- E. Bruno, N. Faessel, J. Le Maitre, "Indexation of Web Pages Based on their Visual Rendering", in: Proceedings of the IADIS International Conference WWW/Internet 2007, vol. 2, pp. 193-197, Vila Real, Portugal, October 2007.
- E. Bruno, E. Murisasco, "An XML environment for multistructured textual documents", in: Proceedings of the Second International Conference on Digital Information Management (ICDIM'07), Lyon, october 2007.
- E. Bruno, S. Calabretto, E. Murisasco (2007) ``Documents textuels multistructurés : un état de l'art'', in: Revue I3 (Information - Interaction - Intelligence), mars 07.
- E. Bruno, E. Murisasco (2006) , MultiStructured XML textual documents, GESTS International Transactions on Computer Science and Engineering, Vol 24, n°1 – novembre 2006.
- E. Bruno, E. Murisasco (2006) ``MSXD : a model and a schema for concurrent structures defined over the same textual data'', in: Proceedings of the DEXA Conference, Lecture Notes in Computer Sciences 4080, pp. 172-181, Krakow, Poland.

Les cinq publications les plus significatives :

- E. Bruno, E. Murisasco (2006) ``Describing and Querying hierarchical structures defined over the same textual data'', in: Proceedings of the 2006 ACM Symposium on Document Engineering (DocEng 2006), pp. 147-154, Amsterdam, The Netherlands.
- E. Bruno, E. Murisasco, "MSXD : a model and a schema for concurrent structures defined over the same textual data", in: Proceedings of the DEXA Conference, Lecture Notes in Computer Sciences 4080, pp. 172-181, Krakow, Poland, September 2006.
- E. Bruno, J. Le Maitre and E. Murisasco (2003), Extending XQuery with Transformation Operators, Proceedings of the 2003 ACM Symposium on Document Engineering (DocEng 2003), ACM Press, Grenoble, France, November 20-22 2003, pp. 1-8.
- A. Gabillon, E. Bruno, "Regulating Access to XML documents", in: Proceedings of the 15th Annual IFIP 11.3 Working Conference on Database Security, pp. 311-328, Niagara on the Lake, Canada, July 2001.
- E. Bruno, J. Le Maitre, E. Murisasco, "Controled Hypertextual Navigation in the SgmlQL Language", in: Proceedings of the 10th International Conference DEXA'99, Lecture Notes in Computer Sciences n°1677, pp. 83-99, Florence, I taly, August 1999.

CURRICULUM VITAE

Mr LE MAITRE Jacques

Né le 1er avril 1946

Téléphone : 04 91 41 26 40
Fax : 04 94 14 28 97
e-mail : lemaitre@univ-tln.fr
adresse : LSIS, Université du Sud Toulon-Var
BP 20132, 83957 La Garde

SITUATION ACTUELLE

Professeur à l'Université du Sud Toulon-Var

Affectation : LSIS (UMR 6168)

PARCOURS PROFESSIONNEL

depuis septembre 1995 : professeur à l'Université de Toulon et du Var
1991 à août 1995 : chargé de recherche détaché au CNRS
1974 à 1990 : ingénieur contractuel puis ingénieur d'études au CNRS
1972 à 1973 : service national comme assistant à l'Université de Rabat au Maroc
1968 à 1971 : ingénieur contractuel au Centre d'Etudes Nucléaire de Cadarache

PUBLICATIONS

Les cinq publications les plus récentes :

N. Faessel, J. Le Maitre, « Une extension de XQuery pour la recherche textuelle d'information dans des documents XML », *Actes des 7e Journées Francophones Extraction et Gestion des Connaissances*, Namur, Belgique, janvier 2007, Revue des Nouvelles Technologies de l'Information, RNTI-E-9, vol. 1, pp. 379-389.

J. Le Maitre, "Representing multistructured XML documents by means of delay nodes", *Proceedings of the 2006 ACM Symposium on Document Engineering (DocEng 2006)*, Amsterdam, The Netherlands, October 2006, pp. 155-164.

S. Tollari, H. Glotin and J. Le Maitre, "Enhancement of Textual Images Classification using Segmented Visual Contents for Image Search Engine", *Multimedia Tools and Application Processing*, Special Issue on *Metadata and Adaptability in Web-based Information Systems*, vol. 25, n°3, pp. 405-417, Springer, 2005.

J. Le Maitre, "Indexing and Querying Content and Structure of XML Documents According to the Vector Space Model", *Proceedings of the IADIS International Conference WWW/Internet* 2005, Lisbon, Portugal, October 2005, vol. II, pp. 353-358.

E. Bruno, E. Murisasco et J. Le Maitre, « Temporalisation d'un document multimédia », *Document Numérique*, numéro spécial *Temps et Documents*, vol. 8, n°4, pp. 125-141, Hermès, 2004.

Les cinq publications les plus significatives :

S. Tollari, H. Glotin and J. Le Maitre, "Enhancement of Textual Images Classification using Segmented Visual Contents for Image Search Engine", *Multimedia Tools and Application Processing*, Special Issue on *Metadata and Adaptability in Web-based Information Systems*, vol. 25, n°3, pp. 405-417, Springer, 2005.

E. Bruno, J. Le Maitre and E. Murisasco, "Extending XQuery with Transformation Operators", *Proceedings of the 2003 ACM Symposium on Document Engineering (DocEng 2003)*, Grenoble, France, November 2003, pp. 1-8.

J. Le Maitre, E. Murisasco and M. Rolbert, "From Annotated Corpora to Databases: the SgmlQL Language", in J. Nerbonne (Ed.), *Linguistic Databases*, *CSLI Lecture Notes n°77* , 1997, pp. 37-58.

N. Ide, J. Le Maitre and J. Véronis, "Outline of a model for lexical databases", *Information Processing & Management*, vol. 29, n°2, pp. 159-186, 1993.

J. Le Maitre et O. Boucelma, « LIFOO, un langage fonctionnel de requêtes pour bases de données orientées objets », *Technique et science informatiques (TSI)*, vol. 11, n°5, pp. 67-94, 1992.

ANNEXE III : Implication des personnes dans d'autres contrats

Partenaire Partner	Nom de la personne participant au projet Name of the person involved in the project	Personne. Mois Man.month	Intitulé de l'appel à projets Source de financement Montant attribué Name call for proposals Other fundings from different organisms Allocated budgets	Titre du projet Proposal title	Nom* du coordinateur Name Principal Inverstigator	Date début -Date fin Start-End of the project
N°1	Robert Espesser	4	ANR Blanc	PROSODY IN GRAMMAR (PRO-GRAM)	Jean-Marie Marandin	2006-2010
N°1	Roxane Bertrand	8	ANR Blanc	PROSODY IN GRAMMAR (PRO-GRAM)	Jean-Marie Marandin	2006-2010
N°1	Philippe Blache	4	ANR corpus	Rhapsodie : Corpus intono-syntaxique de référence du français parlé	Anne Lacheret	2007-2010

Partenaire Partner	Nom de la personne participant au projet Name of the person involved in the project	Personne. Mois Man.month	Intitulé de l'appel à projets Source de financement Montant attribué Name call for proposals Other fundings from different organisms Allocated budgets	Titre du projet Proposal title	Nom* du coordinateur Name Principal Inverstigator	Date début -Date fin Start-End of the project
N°2	Jean-Claude Martin	5	Programme Audiovisuel et Multimédia	CARE : Cultural Experience: Augmented Reality and Emotion	J.-B. Delarivière	2007-2010
N°2	Jean-Claude Martin	5	Technologies Logicielles	Affective Avatars	L. Devillers	2008-2010
N°2	Laurence Devilliers	12	Technologies Logicielles	Affective Avatars	L. Devillers	2008-2010
N°2	Ferré, Gaëlle	4	ANR corpus	Rhapsodie : Corpus i ntono-syntaxique de référence du français parlé	Anne Lacheret	2007-2010

Partenaire Partner	Nom de la personne participant au projet Name of the person involved in the project	Personne. mois Man.month	Intitulé de l'appel à projets Source de financement Montant attribué Name call for proposals Other fundings from different organisms Allocated budgets	Titre du projet Proposal title	Nom* du coordinateur Name Principal Inverstigator	Date début -Date fin Start-End of the project
3	Elisabeth Murisasco	1 (pers.an)	ANR-07- MDCO	ROSES (Really Really Open Simple Efficient Syndication)	Bernd Amann (PR LIP6)	Janvier 2008-janvier 2011
3	Emmanuel Bruno	1.2 (pers.an)	ANR-07- MDCO	ROSES (Really Really Open Simple Efficient Syndication)	Bernd Amann (PR LIP6)	Janvier 2008-janvier 2011
3	Jacques Le Maitre	0.6 (pers.an)	ANR-07- MDCO	ROSES (Really Really Open Simple Efficient Syndication)	Bernd Amann (PR LIP6)	Janvier 2008-janvier 2011
3	Jacques Le Maitre	1.5 (pers.an)	ANR-06- MDCA	AVEIR (Annotation automatique et extraction de concepts visuels pour la recherche d'images)	Patrick Gallinari (PR LIP6)	Janvier 2007-janvier 2010