



UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

Présentation du Laboratoire Parole et Langage

UMR 7309 • CNRS • AMU
PAROLE ET
LANGAGE

Le Laboratoire Parole et Langage (LPL) est une unité de recherche du CNRS associée à l'Université de Provence (Aix-Marseille I) et à l'Université de la Méditerranée (Aix-Marseille II). Il a été créé en 1972 et accueille des phonéticiens, des linguistes, des informaticiens, des psychologues, des neuroscientifiques, des physiciens et des médecins. Les activités du LPL portent sur les mécanismes impliqués dans la **production**, la **perception** et la **compréhension** du **langage oral** et du **langage écrit** dans leurs **contextes naturels** d'occurrence.

Les recherches au LPL s'inscrivent dans une volonté d'interdisciplinarité qui permet de créer de fortes synergies entre les domaines scientifiques qui y sont représentés : de la linguistique expérimentale à la linguistique du terrain, en passant par la phonétique, la psycholinguistique, la neurolinguistique, le traitement automatique des langues, la sociolinguistique, la didactique et l'étude des dysfonctionnements de la parole et du langage.

Le LPL se distingue par ses méthodes de recherche reposant à la fois sur l'expérimentation, l'investigation instrumentale et la formalisation. Approche originale dans ce champ scientifique, qui émerge à la fois aux domaines des sciences humaines, des sciences du vivant et des sciences pour l'ingénieur. Cette particularité explique, au-delà d'une forte activité de recherche fondamentale, l'importance des applications développées à partir des travaux menés dans les domaines du traitement de l'écrit, de l'intelligibilité du message parlé, de la conversion texte-parole de qualité, ou encore de l'évaluation et de la rééducation des troubles de la voix ou du langage. Cette dynamique entre recherche fondamentale et recherche appliquée permet au LPL d'aller jusqu'à une valorisation industrielle, tirant parti de son expérience en la matière. Le LPL a ainsi dirigé plus de 70 contrats nationaux et internationaux et a été à l'origine de la création de deux start-up. Ces caractéristiques font du Laboratoire Parole et Langage une unité de recherche adaptée aux défis scientifiques des sciences du langage, tout en étant impliquée dans leurs enjeux technologiques.

Le laboratoire regroupe actuellement plus de 80 personnes statutaires (chercheurs, enseignants-chercheurs, ingénieurs, techniciens, administratifs), auxquelles s'ajoutent 70 doctorants dont 20 boursiers. C'est le **laboratoire français le plus important dans ce domaine scientifique et l'un des premiers en Europe**.

Le LPL dispose désormais d'une plateforme technique regroupant un ensemble d'instruments pour l'investigation de la production et la perception de la parole : électro-encéphalographie, tracking oculaire, articulographie, électro-palatographie, évaluation articulatoire, etc. Cette ressource unique en Europe est mutualisée au sein du [Centre d'Expérimentation sur la Parole](#) (CEP), sous la responsabilité du LPL et en ses murs.

Le laboratoire assure également la responsabilité scientifique et technique du [SLDR](#) (Speech and Language Data Repository; Banque de données parole et langage) : il s'agit d'un service chargé de la mutualisation et de l'archivage de données orales, basé sur un modèle ouvert de système d'information pour l'archivage (OAI) garantissant un accès pérenne. Dans le cadre de l'[Equipex ORTOLANG](#), le SLDR pour les données orales et le CNRTL pour les données écrites sont les centres de ressources à partir desquels un sous-réseau de CLARIN est en construction en France.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

Depuis 2012, le LPL est l'un des acteurs principaux du [Labex BLRI](#) (Brain and Language Research Institute). Le BLRI implique six laboratoires de la région PACA et s'appuie sur un ensemble exceptionnel de plateformes expérimentales ([CREX](#)). Il regroupe des compétences pluridisciplinaires en linguistique, informatique, psychologie, neurosciences et médecine couvrant l'ensemble des compétences nécessaires à l'étude du traitement du langage et de ses bases cérébrales. L'objectif est d'élaborer un modèle générique et intégré du langage qui puisse être simulé numériquement.

Enfin le LPL est doté d'un [Centre de documentation](#) riche de près de 30 000 ouvrages, toutes catégories confondues. Ce fonds documentaire constitue un outil de recherche et de formation essentiel, largement ouvert à la communauté et aux étudiants.

Site du LPL : www.lpl-aix.fr

Site du CEP : www.lpl-aix.fr/~cep

Site du SLDR : www.sldr.org

Site de l'Equipex Ortolang : www.ortolang.fr

Site du Labex BLRI : www.blri.fr

Actualisé le 01/09/2014





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

Quelques résultats du LPL



1. Production et perception du langage et de la parole

Apprentissage



Grâce à l'enregistrement de potentiels corticaux évoqués (électroencéphalographie), nous avons mis en évidence des changements dans la perception de la parole liés au contact quotidien avec une langue, soit avec une variante régionale d'une langue maternelle ou avec une langue étrangère. En effet, le contact avec des variations régionales entraîne une légère perte dans la

capacité de distinguer les contrastes qui ne sont plus pratiqués dans ces régions, même si le locuteur/auditeur les produit lui-même. Par ailleurs, nos résultats chez l'apprenant adulte mettent en évidence la mise en place de nouveaux contrastes, spécifiques à une langue seconde mais les capacités de perception de cette langue sont intermédiaires entre celles d'un locuteur natif et une personne ne la parlant pas.

La production des sons de la parole

Alors que les travaux portant sur la description des sons de la parole se sont longtemps concentrés sur la description de parole très contrôlée (mots ou phrases lus), des travaux récents, dont les nôtres, montrent que les sons produits en parole non contrôlée (conversations par exemple) sont extrêmement variables. Loin d'être aléatoire, cette variabilité s'inscrit dans des interactions hautement complexes. La production des sons est ainsi le fruit d'un équilibre entre des contraintes physiologiques, d'une part, et la structure linguistique, d'autre part. Nos travaux ont pu mettre en évidence, par exemple, un lien étroit entre le type de mots produits et le relâchement dans l'articulation des voyelles.

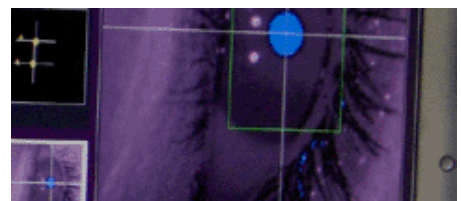
La perception des durées dans la parole

Nous avons montré que la perception des durées segmentales dans la parole était placée sous le contrôle d'une horloge mentale utilisée plus généralement dans le traitement du temps, et dont les ganglions de la base formeraient le substrat cérébral. Ces résultats nous permettent de mieux comprendre la manière dont s'opère la perception de la parole chez le sujet normal et dans certaines pathologies telles que la maladie de Parkinson.

2. Modèles pathologiques

Troubles visuo-attentionnels, lecture, dyslexie

Nous avons pu montrer l'existence d'une relation étroite entre les aptitudes de lecture et l'efficacité des traitements visuo-attentionnels. Cette relation apparaît indépendante des traitements linguistiques et justifie un dépistage spécifique des troubles dans ce domaine. Un logiciel de diagnostic des processus visuels et





UMR 7309

Laboratoire Parole & Langage

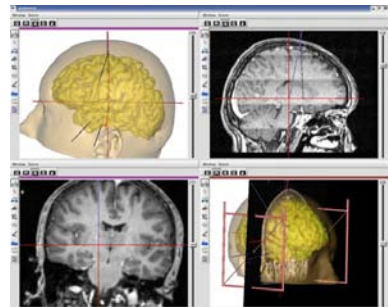
Aix-en-Provence, France

attentionnels chez l'enfant (ADIVA, *Aide au Diagnostic Visuo-Attentionnel*) a été mis au point. Il aide les praticiens à établir un diagnostic précis de la nature des déficits visuels et attentionnels lors de troubles acquis (traumatismes crâniens, pathologies génétiques) ou développementaux (dyslexie, dyspraxie, précocité intellectuelle, etc.).

Des recherches plus spécifiques sur la lecture à l'aide d'un dispositif d'enregistrement des mouvements oculaires montrent une saisie d'information visuelle différente chez les enfants dyslexiques par rapport à des apprenants lecteurs en général, qui suggèrent un déficit visuo-attentionnel chez les dyslexiques.

Maladie de Parkinson

Nos recherches sur les mécanismes neurophysiologiques de la production de la parole dans la maladie de Parkinson ont montré une altération de la boucle cérébello-corticale (liaison cerveau/cervelet) dans le contrôle moteur de la parole, *a contrario* de l'augmentation d'activation du cervelet observée pour compenser le déficit moteur des membres au niveau cortical. Plus généralement, nous portons notre effort sur l'analyse des troubles de la parole chez ces patients, notamment pour évaluer les effets de traitements pharmaceutiques ou d'une stimulation cérébrale profonde. Les moyens expérimentaux utilisés sont multiples : enregistrements acoustiques, aérodynamiques, imagerie cérébrale fonctionnelle.



Forçage vocal

Nous pensons que les troubles de la voix chez l'enseignant et le professionnel de la voix correspondent à un dysfonctionnement global de la communication parlée, depuis les phénomènes de serrage au niveau des cordes vocales jusqu'au comportement global de l'individu en interaction avec son environnement et avec ses interlocuteurs. L'accent est donc mis sur le rôle phonologique du larynx. Cette conception rapproche ainsi le phénomène médical de voix forcée du phénomène phonétique d'accentuation. Sur le plan de la modélisation physique, nous avons proposé un modèle de vibration inspiré de l'effet on-off (modèle "à relaxation") complémentaire du modèle sinusoïdal classique.



Surdit 

Les implants cochl aires (type de proth se auditive implant e) permettent aux enfants sourds de r cup rer une information auditive et ainsi d'am liorer l'acquisition du langage. Nous avons contribu    la compr hension et l'optimisation des r sultats de l'implantation cochl aire qui, chez les sourds profonds, permet de retrouver une boucle audio-phonatoire efficace. Dans ce domaine, les travaux portant sur l'effet Mc Gurk permettent d'analyser l'int gration multimodale de la parole chez ces patients particuliers. Nos travaux visent   optimiser l'analyse des r sultats et constituent un mod le de plasticit  neuronale.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

3. Organisation et fonctionnement de la parole

Nous expliquons le fonctionnement des mécanismes fondamentaux de la parole (compréhension du discours et organisation de la parole conversationnelle en particulier) par le développement de méthodes expérimentales et computationnelles. Nous avons pour cela développé des algorithmes et des logiciels spécialisés qui nous ont par exemple permis d'expliquer le fonctionnement et le rôle de l'intonation (en proposant un nouveau modèle) ou encore de distinguer automatiquement les structures linguistiques simples et complexes pour expliquer le fonctionnement cognitif du traitement linguistique. Ces travaux s'appuient sur l'idée nouvelle que chacun des composants : prosodie ([prononciation régulière des mots](#)), phonologie (étude des sons), syntaxe (règles de construction d'une langue), etc. doit être décrit en prenant en compte ses interactions avec les autres.



Prosodie

Nous avons proposé un nouveau modèle prosodique décrivant les unités prosodiques et leurs fonctions, en prenant en compte les aspects phonétiques, phonologiques et fonctionnels. Nous avons en particulier, grâce à une modélisation de l'intonation reprise aujourd'hui par de nombreux chercheurs dans le monde, pu décrire les caractéristiques d'un grand nombre de langues.

Applications : synthèse de parole de haut niveau, évaluation des troubles de la prosodie, langage pathologique, constitution de bases de données de référence pour la langue parlée (français, anglais), enseignement assisté par ordinateur de langue seconde basé sur la prosodie, etc.

Phonologie et articulation

Nous avons décrit précisément le fonctionnement des organes articulatoires et montré la corrélation entre les commandes du système laryngien et celles du conduit vocal. Ces résultats mettent en évidence l'importance des mécanismes articulatoires par rapport aux phénomènes acoustiques. Nous avons pour cela mis en place une étude sur plusieurs langues de la cinématique de 10 capteurs, suivie dans le temps, à l'aide d'un électromagnétomètre. Ces résultats permettent d'expliquer certains phénomènes comme la perception des contrastes lexicaux.



Syntaxe

Nous avons proposé une nouvelle théorie syntaxique permettant de décrire tous les usages de la langue, aussi bien écrite que parlée. Cette approche permet d'expliquer, au-delà de la syntaxe, comment les différents composants interagissent entre eux. Nous avons grâce à elle décrit les phénomènes de perception de la structure syntaxique par des sujets humains. Cette théorie permet par ailleurs de proposer un modèle de la complexité linguistique : les structures moins complexes sont plus faciles à traiter (et à apprendre) que les autres. Nous avons pour la première fois proposé un modèle computationnel permettant de quantifier cette complexité.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

Plusieurs algorithmes et logiciels ont été développés à l'aide de cette théorie. Par ailleurs, dans la même perspective, nos résultats en électrophysiologie mettent en évidence la mise en place graduelle du traitement syntaxique en langue seconde.

Multimodalité de la communication

La communication humaine repose sur l'utilisation simultanée de différentes modalités, en particulier la parole et les gestes. Nous avons élaboré un modèle complet permettant une description précise des informations de ces modalités dans des conversations. Une description des interactions entre les modalités, mais également entre les composants linguistiques, est ainsi rendue possible. Nous avons pour cela mis au point une plateforme permettant la constitution de base de données à grande échelle. Celle-ci a été expérimentée pour un corpus (*Corpus of Interactional Data*) qui constitue la première expérimentation à grande échelle réalisée sur une langue et comportant un niveau d'information précis pour toutes les modalités. Cet outil nous a déjà permis de mettre en lumière un grand nombre d'éléments nouveaux pour l'explication de la communication multimodale.

4. Le langage dans l'espace géographique et social

Les parlers des jeunes en milieu urbain : diffusion et imitation

Des travaux portant sur l'accent dit "quartiers Nord" à Marseille révèlent que malgré la vive ségrégation dont il fait l'objet de la part d'une partie de la population locale, cet accent joue un rôle croissant à Marseille, au point de devenir une variante de prestige auprès de certains jeunes qui en adoptent les traits. Les tensions observées autour de cet accent rendent compte du changement en cours à l'œuvre dans les pratiques d'une partie de la jeunesse française et de la façon dont il est vécu localement.



Créoles

Nous proposons une définition des créoles en tant que langues à part entière, avec un processus d'évolution qui n'a rien d'exceptionnel par rapport à d'autres situations de contact entre plusieurs langues. Ainsi, la créolisation peut servir de modèle pour décrire les mécanismes à l'œuvre lors de l'émergence de nouvelles variétés, surtout lorsque celles-ci sont issues du contact linguistique, cas de plus en plus fréquent dans un monde globalisé. Par ailleurs, le LPL héberge un fonds documentaire riche de près de 15 000 titres dans ce domaine.

L'humour dans la conversation quotidienne

Nous proposons une analyse du fonctionnement de l'humour en situation conversationnelle, reposant sur des données naturelles. La description de corpus a permis l'identification de plusieurs paramètres (cible, incidence dans le déroulement de l'échange, attitude du locuteur, etc.) permettant de décrire le fonctionnement de l'humour dans la conversation (analyse qualitative et quantitative des occurrences, fonction des rires, etc.). Cette approche repose sur la prise en compte de domaines variés et permet une étude comparative multilingue.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

5. Instruments et logiciels développés par le LPL

- **EVA** : L'*Evaluation Vocale Assistée* est un système d'aide au diagnostic des pathologies vocales. Il s'appuie sur l'étude de la plupart des paramètres de la production de parole: son, hauteur, intensité de la voix, débits d'air, pressions, etc. Doté de nombreux capteurs permettant ces prises de mesure, il permet d'affiner le diagnostic du praticien, de faire le suivi d'une intervention chirurgicale, d'un traitement pharmaceutique ou d'une rééducation. Les travaux actuels portent sur des situations cliniques complexes notamment dans la maladie de Parkinson et le forçage vocal qui sont des modèles "difficiles" pour les méthodes de mesure et contribuent à affiner et à valider les procédures (voir également <http://www.sqlab.fr/evaRootFR.htm>).
- **PCA** : La *Plateforme de Communication Alternative* est un logiciel d'aide à la communication destiné aux personnes dont l'élocution est difficile, voire impossible, et dont la motricité est réduite. Quelques interactions suffisent pour composer un message, quel qu'il soit. Deux types de communication sont proposés : la communication verbale, c'est-à-dire par écrit et la communication non-verbale à base d'icônes.
- **PERCEVAL** (*PERCEption EVALuation Auditive & Visuelle*) est une station automatisée de tests de perception auditive et visuelle développée au LPL. Elle fournit un environnement complet pour la préparation, la configuration, la passation et le recueil des données de l'expérience. PERCEVAL permet de construire des expériences où le sujet est exposé à une série de stimuli visuels et/ou audio. Elle est donc spécialement adaptée pour l'étude de la perception de la parole et du langage.
- **PHONEDIT** est un éditeur de signal développé pour des recherches sur la voix et la parole. Il est utilisé pour visualiser, segmenter, marquer, mesurer et traiter des paramètres acoustiques, aérodynamiques, palatographiques et kinésiographiques. PHONEDIT est capable de lire la majorité des fichiers des différents systèmes informatiques utilisés par les phonéticiens. Une nouvelle version intégrant les sources visuelles et permettant une synchronisation des différentes sources de données est en cours de développement.
- **SPPAS** (*SPeech Phonetization Alignment and Syllabification*) est un logiciel qui produit automatiquement ou semi-automatiquement des annotations phonétiques qui incluent la segmentation en phrases, mots, syllabes et phonèmes à partir de fichiers sons et de leur transcription.
- **EMATOOLS** est une boîte à outils sous Matlab pour le traitement de données d'électro-magnétoarticulographie et d'électro-palatographie.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

- **PROSODIA** est un logiciel d'aide à l'apprentissage de la prosodie de l'anglais (extension possible à d'autres langues ou à la rééducation).
- **EMAA** (*Eye Movement Analysis and Acquisition*) : ce logiciel permet de piloter en temps réel des expériences de mouvements oculaires chez l'enfant à partir du système EYELINK 2 et de traiter les données recueillies.
- **ADIVA** (*Aide au Diagnostic Visuo-Attentionnel*) : Ce logiciel aide les praticiens à établir un diagnostic précis de la nature des déficits visuels et attentionnels lors de troubles acquis (traumatismes crâniens, pathologies génétiques) ou développementaux (dyslexie, dyspraxie, THADA, précocité intellectuelle, etc.). Il n'existe à ce jour pas d'équivalent en tests cliniques informatisés.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

6. Valorisation

Le LPL a été à l'origine de la création de deux start-up :

AEGYS

La société Aegys est une jeune entreprise innovante intervenant dans le secteur des technologies pour le handicap. Créée fin 2003, les activités de la société Aegys s'orientent vers quatre directions principales :

- Faire évoluer, maintenir, distribuer et commercialiser le produit phare de la société : le logiciel d'aide à la communication "Plateforme de Communication Alternative" (PCA) ;
- Proposer, concevoir et développer des solutions logicielles alternatives innovantes pour améliorer la vie au quotidien des personnes handicapées (Aegys a récemment mis au point un capteur générique sensible à l'émission d'un son, d'un souffle, à une légère pression de contact ou à la contraction du muscle orbiculaire de l'œil) ;
- Intervenir en tant que prestataire de services dans des programmes de recherche portant sur la problématique de l'aide à la communication et, plus généralement, sur l'aide au handicap (Aegys est actuellement partenaire d'un projet de recherche sur l'autisme) ;
- Conseiller et former les professionnels du secteur social et médico-social sur le thème de l'aide à la communication pour les personnes handicapées.

SQLab

La société SQLab a débuté son activité en 1997 au sein de la Technopole de l'Arbois, dans le Centre Européen des Entreprises Innovantes (C.E.E.I. Provence), à Aix-en-Provence. Toujours implantée à Aix-en-Provence, elle se situe à proximité des grands centres industriels, universitaires et hospitaliers de la région d'Aix-Marseille.



SQLab est liée contractuellement avec le CNRS pour la fabrication et la commercialisation du dispositif EVA. Les produits SQLab sont utilisés dans d'importants centres scientifiques à travers le monde.

Le produit phare de la société est le dispositif d'Evaluation Vocale Assistée EVA2™ conçu pour l'étude de la plupart des paramètres de la production de parole: son, hauteur, intensité de la voix, débits d'air, pressions... Doté de nombreux capteurs permettant ces prises de mesure, il permet, dans son utilisation clinique issue d'un transfert de technologie, d'affiner le diagnostic du praticien, de faire le suivi d'une intervention chirurgicale, d'un traitement pharmaceutique ou d'une rééducation. Le dispositif est constitué par un micro-ordinateur PC auquel sont connectés des capteurs acoustiques, aérodynamiques et électrophysiologiques. Le tableau de commande des instruments de mesure et l'affichage des résultats sont gérés par l'ordinateur au moyen de logiciels spécifiques à chaque protocole d'investigation clinique.



Une cinquantaine d'équipements a été diffusée à travers le monde (France, Belgique, Allemagne, Italie, Turquie, USA, Canada, Brésil) dans des centres cliniques et universitaires. Cette opération de valorisation a eu un effet fédérateur au niveau national dans la constitution d'un réseau d'étude des troubles de la parole par des acteurs multidisciplinaires cliniciens et phonéticiens dans les grands centres hospitalo-universitaires (Aix-en-Provence, Marseille, Montpellier, Grenoble, Toulouse, Paris, Lille, Strasbourg).





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

Centre d'Expérimentation sur la Parole

Plateforme expérimentale mutualisée



Le Centre d'Expérimentation sur la Parole (CEP) est une plateforme expérimentale mutualisée destinée au recueil et au traitement de données pour l'étude de la production et de la perception de la parole. Elle permet de recueillir un grand nombre d'informations, audio et vidéo, mais également physiologiques, comportementales et cérébrales.

Le CEP est composé de différents plateaux techniques permettant de conduire ces expérimentations: station d'électro-encéphalographie, articulographe, chambre anéchoïque, système de suivi des mouvements oculaires, stations de test de perception, banc de palatographie statique et dynamique etc.

Il constitue un outil exceptionnel permettant non seulement de conduire des expérimentations classiques, mais également d'imaginer des expériences plus complexes conduisant à la production de données uniques, en combinant les informations produites par différents instruments scientifiques rassemblés sur un même site.

De plus, ce centre permet d'élaborer des expérimentations à grande échelle et de réaliser des tests de perception ou de production écrite/orale réunissant 10 sujets simultanément.

Le CEP constitue une plateforme unique en son genre en Europe de par la diversité des instruments d'investigation proposée. Financé avec le soutien du CNRS (140k€) et du CPER (120k€) le Centre d'Expérimentation sur la Parole est déployé et accueille déjà des équipes de recherche externes au LPL, nationales et internationales.

Site du CEP : www.lpl-aix.fr/~cep

Composition de la plateforme

Le Centre d'Expérimentation sur la Parole (CEP) met au service de la communauté de recherche sur la parole et le langage un ensemble d'instruments et une infrastructure permettant de conduire des expériences exceptionnelles. Dans cette section sont présentés les principaux instruments qui constituent la plateforme.





UMR 7309

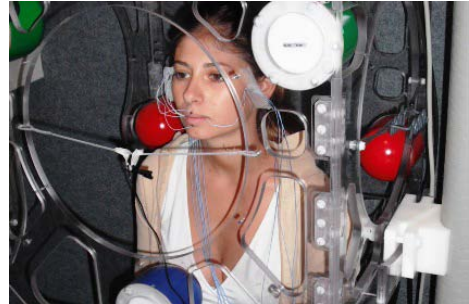
Laboratoire Parole & Langage

Aix-en-Provence, France

- **L'Articulographe**

La production de parole requiert une coordination très précise de plusieurs articulateurs (langue, lèvres, mâchoire, voile du palais) qui présentent des natures et des masses très différentes. Pour étudier leurs mouvements relatifs on utilise l'articulographe électromagnétique. Cet instrument permet de mesurer avec une très bonne résolution spatiale (moins d'1 mm) et temporelle (toutes les 5 ms) la position de douze capteurs placés sur la langue, les dents, les lèvres, la mâchoire....

Le principe de fonctionnement est très simple, il s'agit pour chaque capteur, d'analyser la composante enregistrée de chacun des trois champs magnétiques de référence générés par l'instrument. Une fois la position des capteurs connue, on peut en fonction du sujet d'étude s'intéresser à la position, à la vitesse ou à l'accélération des articulateurs impliqués dans la production de parole.



- **L'Electropalatographe**



L'electropalatographie ou EPG permet d'enregistrer dynamiquement les contacts entre la langue et le palais pendant la production de parole.

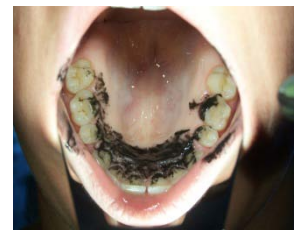
La technique nécessite le port d'un palais artificiel moulé (pseudo-palais) muni d'électrodes. Les électrodes sont reliées via un multiplexeur à l'unité EPG qui enregistre les signaux et permet la visualisation des contacts langue/palais en temps réel.

- **Le banc de palatographie statique**

La palatographie statique consiste à enduire la langue d'une pâte colorée, à demander au locuteur la production étudiée (par exemple /na/), et à analyser la forme et la position de l'empreinte laissée par la langue sur le palais.

Cette méthode ancienne prend un nouvel essor avec les photos numériques auxquelles sont appliqués des traitements d'images permettant une extraction automatique multiparamétrique de l'empreinte.

Cette technique rend possible l'observation de phénomènes très fins (excellente résolution spatiale) et permet une approche métrologique et statistique pour des études intra et inter locuteurs. Notons que cette technique simple, peu coûteuse est compatible avec les contraintes liées aux enquêtes de terrain.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

- **L'Electroglottographe**



L'électroglottographe permet d'étudier les mouvements d'ouverture et de fermeture de la glotte lors de la phonation (électroglottographie phonatoire). Cette méthode est également utilisée pour l'étude des mouvements lents des plis vocaux lors de la respiration (électroglottographie respiratoire).

Le principe de la mesure consiste à placer de part et d'autre du cartilage thyroïdien deux électrodes dans lesquelles circule un courant alternatif de très faible intensité (quelques milliampères).

Lors du cycle glottique, la surface de contact entre les cordes évolue, ce qui modifie l'impédance (résistance au passage du courant) mesurée entre les deux électrodes.

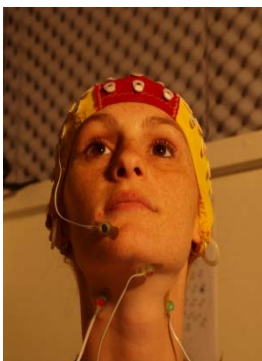
L'observation du signal recueilli permet de mesurer la fréquence d'oscillation des cordes vocales. L'analyse de ce signal permet aussi de déterminer avec précision les instants d'ouverture et de fermeture du plan glottique.

- **La station d'Evaluation Vocale Assistée (EVA)**



Le système d'Evaluation Vocale Assistée EVA est conçu pour l'étude de la plupart des paramètres de la production de parole: son, hauteur, intensité de la voix, débits d'air, pressions... Doté de nombreux capteurs permettant ces prises de mesure, il permet d'affiner le diagnostic du praticien, de faire le suivi d'une intervention chirurgicale, d'un traitement pharmaceutique ou d'une rééducation vocale. Le dispositif est constitué d'un micro-ordinateur PC auquel sont connectés des capteurs acoustiques, aérodynamiques mais aussi électrophysiologiques. Les tableaux de contrôle des paramètres et d'affichage des résultats sont gérés par l'ordinateur au moyen de logiciels spécifiques à chaque protocole d'investigation clinique. Les enregistrements sonores sont effectués à partir de microphones qui permettent une mesure calibrée de l'intensité.

- **La station d'Electro-encéphalographie (EEG)**



Cet instrument est destiné à l'analyse de l'activité électrique enregistrée en surface du crâne (scalp) induite par une activité cérébrale. Cette activité peut être consécutive à une stimulation audio ou vidéo du sujet. Les potentiels électriques sont recueillis par 64 électrodes actives permettant l'acquisition de signaux présentant un excellent rapport signal à bruit. Ce système permet d'enregistrer simultanément avec les signaux d'électroencéphalographie, le clignement des yeux (EOG), l'activité musculaire (8 voies EMG), ainsi que l'activité cardiaque (ECG). Ce système est placé dans un box isolé (cage de Faraday) qui jouxte la station de contrôle de stimulation et d'enregistrement.



UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

- **Le système d'électro encéphalographie EEG haute densité (256 voies)**



Ce dispositif ouvre un large spectre de recherches potentielles dans le domaine de l'imagerie cérébrale au sein du laboratoire. Conçu par la société Electrical Geodesics Inc. (Eugene, USA), ce système est couplé à un dôme de photogrammétrie équipé de 11 caméras permettant de connaître la position exacte du casque sur le participant. Le logiciel fourni permet l'analyse des potentiels électriques enregistrés en surface du scalp mais aussi de localiser le siège de l'activité cérébrale à l'origine du signal électrique.

- **Les systèmes de suivi des mouvements oculaires**

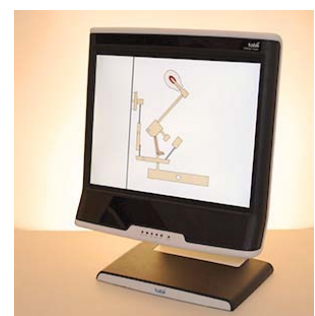
L'étude des mouvements oculaires est un moyen d'étude du traitement du langage écrit et parlé. Le **Centre d'expérimentation sur la parole (CEP)** dispose de trois systèmes complémentaires :

Eyelinek II : Cet instrument est tout à fait adapté aux études portant sur l'apprentissage et les troubles de l'acquisition chez l'enfant scolarisé. Il peut aussi être utilisé pour des études chez l'adulte.



Eyelinek 1000 : Ce système de suivi oculaire très précis est réservé aux études chez l'adulte. Il exige un positionnement contrôlé du sujet (mentonnière et repose tête).

TOBII T-120 : Cet instrument est sur le point d'être déployé. Il est tout à fait complémentaire aux deux autres systèmes. Sa résolution angulaire est moins grande mais il présente la particularité d'être très peu contraignant. Il permet de mener des études sur le développement du très jeune enfant ou encore chez les patients autistes par exemple. Très facilement transportable il est destiné à des études menées dans les murs du CEP mais aussi en écoles ou hôpitaux.





UMR 7309

Laboratoire Parole & Langage

Aix-en-Provence, France

- **Stations de test de perception**



Le participant à l'expérience est soumis à une stimulation auditive ou visuelle. Il est invité à faire un choix par le biais d'un boîtier réponse, d'un clavier ou encore à répondre oralement à la sollicitation.

Deux salles sont dédiées à ce type de test : une cabine individuelle insonorisée et une salle de passation collective permettant d'accueillir jusqu'à 10 sujets simultanément.

- **Chambre sourde**



La chambre anéchoïque est une chambre disposant d'un revêtement spécial (mur et plafond) absorbant les ondes sonores. Il s'agit d'une chambre relativement grande (70m³). Cet outil sert à l'acquisition de données aussi bien qu'au test ou au calibrage de matériel.

A ce dispositif est associé un ensemble de microphones de grande qualité et enregistreurs qui constituent l'équivalent d'un studio professionnel.

Un ensemble de signaux audio et vidéo permettent à la régie de communiquer avec les sujets situés dans la chambre sourde. La régie est équipée d'une station ZOOM HD16 permettant simultanément l'enregistrement de 8 pistes audio stéréo.

- **Station de montage vidéo**



Un équipement d'enregistrement et de montage vidéo complet permet le tournage de séquences dans de très bonnes conditions.

Cet équipement concerne un grand nombre de thèmes de recherche tels que ceux qui portent sur :

- La communication multimodale,
- Les langues des signes,
- Les études posture/voix.

SLDR - Speech and Language Data Repository - www.sldr.org

Le LPL assure la responsabilité scientifique et technique du SLDR (*Banque de données parole et langage*), un service chargé de la mutualisation et de l'archivage de données orales. Le SLDR est issu de l'initiative CRDO-Aix, en réponse à un appel à projets de la Direction de l'Information Scientifique et du Département scientifique Homme et Société du CNRS en 2006. Au terme d'un projet pilote coordonné par le TGE-Adonis, le SLDR fonctionne aujourd'hui comme un dispositif interopérable et ouvert s'appuyant, pour l'archivage pérenne, sur le *Centre informatique de l'enseignement supérieur* (CINES) et, pour le partage des données, sur le *Centre de calcul de l'Institut national de physique nucléaire et de physique des particules* (CC-IN2P3).

L'objectif du SLDR est de regrouper l'ensemble des données portant sur du signal oral. Une telle centralisation permet tout d'abord d'éviter que les différents laboratoires travaillant dans ce domaine ne soient conduits à recréer sans cesse des données. La mutualisation des ressources en sciences du langage répond

par ailleurs au besoin stratégique de rapprocher les connaissances aujourd'hui dispersées dans des domaines variés comme la linguistique descriptive, formelle et computationnelle, la sociolinguistique, la littérature, la traductologie, les neurosciences, la psycholinguistique etc.

Les ressources partagées avec la communauté scientifique (et le public en général) comprennent des corpus de parole (spontanée ou lue), leurs enrichissements (transcriptions et annotations), des lexiques et autres bases de connaissances, ainsi que des outils adaptés au traitement de ces données. Les corpus proprement dits peuvent inclure des enregistrements audio/vidéo mais aussi des mesures de l'activité physiologique : endoscopie, potentiels évoqués, palatographie, etc.

Le partage des ressources est basé sur un modèle ouvert de système d'information pour l'archivage (OAIS) garantissant un accès pérenne aux objets par le biais d'un identifiant unique. Le dispositif mis en place au SLDR permet de reconstituer intégralement un objet archivé à partir de la version distribuée (au CC-IN2P3) et donc de libérer l'espace correspondant sur le serveur. En réponse aux besoins les plus diversifiés, l'équipe a privilégié le développement d'un système générique qui minimise les contraintes techniques sur la structure et la volumétrie des objets archivés : segmentation automatique des objets de grande taille, pas de limite sur les noms de fichiers, prise en compte de l'évolution des objets (grâce à l'archivage intermédiaire) et gestion des droits d'accès dans les phases successives d'un projet.

Le téléchargement d'un objet peut être réservé à certaines catégories d'utilisateurs. Les restrictions d'accès aux documents déposés en archive publique s'appuient sur les dérogations au principe de libre communicabilité prévues par le Code du patrimoine (art. L213-2, loi du 15 juillet 2008).

Le serveur du SLDR conserve la trace des transactions sur certaines ressources et les utilisateurs s'engagent, par le biais d'une licence non-commerciale, à les mentionner sur toute production scientifique basée sur leur partage. Ce dispositif permet de faire émerger des communautés de producteurs et d'utilisateurs favorisant l'agrégation de projets au-delà des frontières géographiques ou institutionnelles. Le site SLDR est quadrilingue : anglais, français, espagnol, chinois, et s'appuie sur l'approche Web 2.0 pour les communautés d'utilisateurs, avec des liens vers les réseaux sociaux.



Speech & Language Data Repository (SLDR) www.sldr.org

Speech & Language Data Repository (SLDR) www.sldr.org

Primary data (corpus)

Flash Player plugin is required to read previews. If the preview cell is empty or the player is not active, you may download Adobe Flash Player.

Display corpus in: Arabic, Standard (العربية القياسية) [AR] (1) [OK]

ID	Name	Description	Downloading	Preview
[AMK] enregistrements	Standard ARABIC statements (Amina CHENTRI)	language_description/narrative # Laboratoire parole et langage - UMR 7309 (LPL, Aix-en-Provence FR) Affirmative sentences in standard Arabic recorded in a sound-proof studio by 5 North-African speakers (2 men and 3 women). These sentences contain all the phonemes of standard Arabic. جمل مفيدة باللغة العربية القياسية مسجلة داخل غرفة معزولة من طرف خمسة أشخاص من شمال إفريقيا (2 رجال و 3 نساء) - تحتوي على جميع فونيمات العربية القياسية (نساء العربية القياسية)	Download	Preview

Contact : <http://sldr.org/contact>