

# Expériences et recommandations pour la structuration des données sonores, physiologiques et cliniques dans le cas des dysfonctionnements de la parole

*Alain Ghio<sup>1</sup>, Gilles Pouchoulin<sup>1,2</sup>, Lise Crevier<sup>3</sup>, Cécile Fougeron<sup>3</sup>, Corinne Fredouille<sup>2</sup>, Antoine Giovanni<sup>1</sup>, Danièle Robert<sup>1</sup>, Antonia Simon<sup>3</sup>, Bernard Teston<sup>1</sup>, François Viallet<sup>1</sup>*

<sup>1</sup>Laboratoire Parole et Langage, CNRS, Université Aix-Marseille, France

<sup>2</sup>Université d'Avignon, CERI/LIA, Avignon, France

<sup>3</sup>Lab. de Phonétique et Phonologie, UMR 7018 CNRS-Paris3/Sorbonne Nouvelle, Paris, France  
alain.ghio@lpl-aix.fr, corinne.fredouille@univ-avignon.fr, cecile.fougeron@univ-paris3.fr

## ABSTRACT

Research in clinical phonetics needs a structured organisation of various data. In such a domain, speech signal is useless if it is unlinked to the clinical state of the speaker or to the speech production context. In this paper, we present a model of database which is designed for the storage and accessibility of various speech disorder data. Information contained in this model can be considered as recommendations based on our own experience in recording data in a clinical context.

**Keywords:** clinical phonetics, voice/speech disorders, database

## 1. INTRODUCTION

Depuis une quinzaine d'années, l'étude des dysfonctionnements de la voix et de la parole est sortie du simple cadre de la recherche clinique et intéresse les laboratoires de recherche issus des sciences du langage ou du traitement automatique de la parole. Par l'observation des dysfonctionnements, les chercheurs non cliniciens confrontent les résultats de leur recherche établis sur des corpus de parole "normale" à des situations de dysfonctionnement. Ces recherches permettent un enrichissement des connaissances générales sur la parole et bénéficient de la complémentarité des expertises des communautés de cliniciens et des autres scientifiques. A ce titre, on peut citer les travaux de Hardcastle et Gibbon [1] où des techniques d'Electropalatographie conçues au départ pour les études de phonétique, ont été utilisées pour des locuteurs souffrant de troubles de l'articulation, initiative ayant donné lieu à un grand nombre de publications ainsi qu'à un réseau de santé CleftNet.

## 2. PARTICULARITÉS DES DONNÉES DE « PAROLE PATHOLOGIQUE »

L'expression "parole pathologique" est un raccourci pour désigner la parole produite par des locuteurs atteints de dysfonctionnements de la voix et de la parole. L'étude multidisciplinaire de la "parole pathologique" nécessite trois exigences : (1) des signaux de très bonne qualité pour ne pas imputer des distorsions, bruits... aux dysfonctionnements; (2) un matériau linguistique (énoncé produit par les locuteurs) suffisamment riche ; (3) des renseignements cliniques sur les locuteurs suffisamment précis (pathologie, thérapies, contexte clinique d'enregistrement). Actuellement, les études sur le

dysfonctionnement de la voix et de la parole souffrent cruellement d'une dispersion et d'une hétérogénéité des données. Souvent, les analyses portent sur quelques locuteurs enregistrés pour les besoins ponctuels d'une étude. L'enregistrement des signaux et le stockage sont souvent effectués par du personnel non expérimenté à certains aspects techniques de la prise et du formatage de données. A cela s'ajoute la perte fréquente des métadonnées (ex: pathologie, durée de la maladie, âge, contexte d'enregistrement...): ces informations sont utilisées comme critères d'inclusion dans une étude particulière puis ne sont plus conservées de façon pérenne et sont dissociées des données sonores.

Pourtant, toute généralisation à une population clinique particulière nécessite l'observation d'un grand nombre de patients du fait de la très forte variation interindividuelle rencontrée (différentes évolutions de la maladie, stratégies de compensation individuelles, sévérité et spécificité des atteintes variées, ...). Certaines pathologies étant rares, la disponibilité des patients dans un contexte clinique n'étant pas toujours possible, l'acquisition de données de parole pathologique n'est pas chose aisée. Pour ces raisons, il est important de mutualiser les enregistrements existants. Or pour être utilisables, ces enregistrements doivent répondre aux exigences mentionnées ci-dessus, ce qui est rarement le cas en pratique. Si les problèmes de prise de son ou autres signaux physiologiques sont en passe de devenir anecdotiques grâce à la diffusion de matériels de qualité et à la meilleure formation des personnels en charge des enregistrements, si le stockage des signaux de parole ne constitue plus actuellement un obstacle, si le recours à du matériau linguistique suffisant se généralise, le maillon faible reste la normalisation et la structuration des données sur les locuteurs et leurs productions langagières. Concrètement, si les données sonores sont souvent accessibles, parfois au prix d'un important effort de numérisation quand le support de stockage est sous forme de bandes, de cassettes..., elles ne présentent au final aucun intérêt si les liens entre les enregistrements et les caractéristiques cliniques du locuteur sont rompus. Nous touchons là aux différences fondamentales entre recueil de données vs base de données, différences que nous aborderons plus loin. L'objectif de ce document est de présenter différentes actions de terrain et de proposer des recommandations pour la structuration des données sonores, physiologiques et cliniques.

### 3. CORPUS DE PAROLE PATHOLOGIQUE

#### 3.1. *Patients dysphoniques du service ORL du CHU Timone, Marseille*

Depuis plus de quinze ans, sous l'impulsion d'Antoine Giovanni, le service ORL du CHU de la Timone à Marseille enregistre régulièrement des patients dysphoniques qui se présentent à la consultation. Ces personnes sont à la fois enregistrées avec l'appareillage EVA [2] mais aussi sur des cassettes D.A.T. Pour des raisons de service, l'information sur les patients est stockée sur un registre sous la forme de cahiers sur lesquels sont notés l'identité du locuteur, sa pathologie, la date de l'examen, le contexte pré/post opératoire... Un important travail de numérisation, d'indexation et de saisie de renseignements a permis de disposer à présent d'une base de données de 1530 patients dysphoniques produisant des /a/ tenus, de la lecture de texte, de l'improvisation chantée..., pour un total de 1961 sessions d'enregistrements (certains locuteurs sont enregistrés plusieurs fois). Cette base est constituée de 504 hommes et 1026 femmes. Les pathologies principales sont : 314 nodules, 228 paralysies laryngées, 202 polypes, 138 œdèmes de Reinke, 139 dysphonies dysfonctionnelles, 81 kystes, 37 sulcus, 25 vergetures. A notre connaissance, cette base de données dépasse le contenu de la référence internationale du Massachusetts Eye and Ear Infirmary (MEEI), qui inclut 700 patients [3]. Nous prévoyons de compléter cette base en ajoutant les données physiologiques récoltées avec EVA sur ces mêmes patients et de l'enrichir avec les locuteurs récemment enregistrés. L'ajout de résultats d'évaluation perceptive (GRBAS) ou instrumentale est à l'étude, ce qui permettrait d'extraire facilement, par exemple, toutes les locutrices avec nodules en situation pré-opératoire avec un grade de sévérité de dysphonie entre 1 et 2.

#### 3.2. *Patients dysarthriques du service de neurologie du CH du Pays d'Aix*

Depuis une dizaine d'années, sous l'impulsion de François Viallet, le service de neurologie du CH du Pays d'Aix enregistre régulièrement des patients dysarthriques qui se présentent à la consultation. Ces personnes sont enregistrées avec l'appareillage EVA et les données cliniques sont saisies avec un tableur. Actuellement, nous disposons des enregistrements sonores et aérodynamiques de 990 patients et 160 sujets contrôles appariés en âge (âge moyen patients = 67,7 ans ; contrôles = 62 ans). La population de patients est composée pour la majorité de malades de Parkinson (601) et syndromes parkinsoniens (98). L'originalité de ce travail réside dans (1) la présence de signaux complémentaires au signal sonore (intensité SPL, débit d'air oral, pression sous-glottique estimée...), ce qui représente 75000 fichiers pour toute la base ; (2) la multiplication des contextes d'enregistrement des 601 malades de Parkinson (avec/sans médication, avec/sans stimulateur sous-thalamique...), ce qui représente 1616 sessions d'enregistrements ; (3) le recueil de

renseignements très précis sur le locuteur (date et lieu de naissance, langue maternelle...), et les conditions cliniques (date d'apparition de la maladie, localisation des symptômes, dosage médicamenteux d'usage et au moment de l'enregistrement, caractéristiques des éventuels stimulateurs électro physiologiques, scores des examens cliniques de type UPDRS, ...) Une telle précision est importante pour des études cliniques (ex : effet des thérapies sur la production de parole) mais aussi au niveau linguistique (recherche de caractéristiques phonéto-acoustiques spécifiques à un groupe homogène de locuteurs dysarthriques).

#### 3.3. *Patients neurologiques CCM*

Entre 1965 et 1997, Claude Chevrier-Muller (CCM) et son équipe du Laboratoire d'étude de la voix et de la parole (INSERM U3), ont enregistré des patients atteints de troubles neurologiques, qui lui étaient adressés dans un but d'évaluation des dysfonctionnements de la parole et de la voix. Ces enregistrements représentent plus de 1000 patients avec des pathologies diverses couvrant dysphonie et dysarthrie mais aussi anarthrie, aphasie, bégaiement et des pathologies psychiatriques. Les enregistrements ont été réalisés suivant un protocole standardisé (qui a un peu évolué dans le temps) d'environ 15 minutes de parole, incluant la production de listes de mots, de phrases, de séries automatiques, de voyelles tenues et de syllabes, une description d'image, la lecture d'un texte et de la parole spontanée. Les signaux audio et électroglottographiques étaient enregistrés en chambre sourde sur des bandes Revox avec indexation sur un cahier. Un dossier médical complet sur papier est associé à chaque enregistrement, comprenant les informations sur l'état civil ainsi que sur la nature de la pathologie, les symptômes et les traitements associés. Un important travail de numérisation des bandes Revox et des dossiers médicaux est actuellement en cours. Cette base est enrichie par des enregistrements réguliers de patients à l'Hôpital Européen Georges Pompidou par Lise Crevier-Buchman et son équipe, suivant le même protocole mais sur un support numérique de type D.A.T.

### 4. ORGANISATION EN BASE DE DONNÉES

L'intérêt majeur de construire une base de données mutualisant des ressources est de pérenniser ces informations, de permettre un échange et un enrichissement graduel via une plateforme web pour un groupe de travail. Pour cela, un modèle de BD a été conçu à partir d'une analyse fonctionnelle réalisée en milieu clinique, pour être ensuite affiné sur la disponibilité de données empiriques issues de différents corpus de parole pathologique comme ceux présentés dans la section 3.

Si les concepts autour des bases de données (BD) sont familiers aux informaticiens, il n'en est pas de même pour les non spécialistes. Il est fréquent de lire qu'un recueil d'enregistrements sonores constitue une base de données : on assimile ainsi corpus/collection de données avec BD. Cette dernière se distingue de la première par une structuration et une organisation cohérente régie par un

modèle, partageable par un groupe de personnes et mémorisable sur un support informatique permettant de sélectionner facilement les données répondant à des critères précis. Cela nous amène à aborder la notion de Système de Gestion de Bases de Données (SGBD) qui doit supporter les concepts définis au niveau du modèle de données, assurer le respect des règles de cohérence définies sur les données, rendre transparent le partage des données entre différents utilisateurs tout en assurant la confidentialité sélective d'une partie des données, pouvoir répondre à des requêtes avec un niveau de performances adapté, fournir différents langages d'accès selon le profil de l'utilisateur. Nous avons opté pour un modèle relationnel, considéré comme le plus simple et le plus élégant des modèles de BD. Sa simplicité provient de l'organisation tabulaire des données, atomiste et minimaliste, rendant intuitive l'architecture des données, les éléments de chaque table étant ensuite liés par des relations. Comme le montre la figure 1, la base de données est composée d'une cinquantaine de tables décrivant les locuteurs avec des informations civiles (date et lieu de naissance, lieu de résidence...), des informations sociolinguistiques (langue maternelle, professions...), médicales (prescriptions thérapeutiques, recueil des symptômes, diagnostics établis) et sur les sessions d'enregistrements. Pour standardiser certaines de ces informations, un ensemble de listes (Figure 1, tables "lst\_") permet de saisir de façon commune aux différents centres fournisseurs de données, des renseignements normalisés tels que professions, langues, pays/régions, symptômes, thérapies, diagnostics, facteurs de risque, localisation des pathologies, contextes expérimentaux, méthodes d'évaluations... L'intérêt de ces listes fermées est d'éviter la multiplication d'appellations pour un même terme (ex : MDP, maladie de Parkinson, Parkinson, Parkinsonien => [52] maladie de Parkinson).

## 5. RECOMMANDATIONS ET RÉALISATIONS TECHNIQUES

Contrairement aux bases de données orales de type patrimonial, dialogal, conversationnel, comme celles disponibles au CRDO [4], la parole pathologique nécessite la collecte et la mémorisation précise des informations sur les locuteurs et le contexte d'enregistrement. Par conséquent, il est recommandé de saisir le maximum de renseignements possible sur les aspects :

- (1) sociolinguistiques (ex : certaines dysarthries se manifestent par une élision des /r/, phénomène qui peut s'apparenter à un accent « créole » ; seules des informations sur le lieu de naissance ou de résidence du locuteur peuvent permettre de faire la part des choses) ;
- (2) médicaux (quelques commentaires sur l'état du patient peuvent permettre de faire des choix d'inclusion/exclusion pour certaines études) ;
- (3) symptomatiques (ex : date d'apparition et localisation des symptômes) ;
- (4) contextuels (ex : « le patient sort d'une bronchite, porte un appareil dentaire, a pris ses médicaments 4h auparavant ») ;

De même, toute forme d'évaluation (ex : UPDRS pour les malades de Parkinson, GRBAS pour les dysphoniques) constituent une source d'information à conserver précieusement.

Si l'utilisation d'un SGBD est recommandée pour la traçabilité et l'exploitation des métadonnées, la standardisation du protocole d'acquisition de données sonores ou physiologiques apparaît difficilement conciliable avec le contexte clinique. En effet, un protocole complet incluant la production de voyelles tenues, d'efforts vocaux, de phrases, de répétitions, de textes lus, de parole spontanée... est difficilement réalisable compte tenu de la batterie complète d'examen que subit le patient et de la fatigabilité engendrée par de trop longs efforts. Il est donc préférable d'adapter les tâches d'élocution à l'état dysfonctionnel du locuteur. Par exemple, l'étude de la nasalité est particulièrement intéressante dans le cas de dysarthries paralytiques du fait de l'immobilité du voile du palais, mais n'est pas centrale dans la maladie de Parkinson pour laquelle des exercices phonatoires peuvent être préférés du fait de l'hypophonie.

A propos de la réalisation technique, la base de données de parole pathologique est développée dans l'environnement PHP/MySQL sur un serveur Apache avec module SSL (Secure Sockets Layer) pour le cryptage des communications. Concernant la sécurité et l'accès aux données, il est recommandé de gérer des privilèges/rôles accordés aux utilisateurs et de crypter les données confidentielles. Ces recommandations nécessaires ne dispensent pas une déclaration à la Commission Nationale de l'Informatique et des Libertés (CNIL) pour le recueil, la saisie des informations cliniques et l'analyse statistique des données recueillies au cours de la recherche [5]. L'anonymisation des données ainsi que l'établissement de consentements éclairés auprès des locuteurs sont des aspects juridiques à ne pas négliger.

*Remerciements : Ce travail a été financé par un PHRC CH du Pays d'Aix et est financé actuellement par l'ANR-08-BLAN-0125 « DESPHO-APADY ».*

## BIBLIOGRAPHIE

- [1] W. Hardcastle, F. Gibbon, Electropalatography and its clinical applications. In *Instrumental Clinical Phonetics*. 149-195, 1997
- [2] A. Giovanni, N. Estublier, D. Robert, B. Teston, M. Zanaret, M. Cannoni, Evaluation vocale objective des dysphonies par la mesure simultanée des paramètres acoustiques et aérodynamiques à l'aide de l'appareillage EVA, *Ann. Otolaryngol. Chir. Cervicofac.*, 112, p. 85-90, 1995
- [3] Massachusetts Eye and Ear Infirmary. *Elementrics Disordered Voice Database*. Voice and Speech Lab, Boston, MA, Kay Elementrics Corp, 1994
- [4] B. Bel, P. Blache, Le Centre de Ressources pour la Description de l'Oral (CRDO). *Travaux interdisciplinaires du LPL d'Aix-en-Prov. (TIPA)*, vol. 25. 2006, p. 13-18.
- [5] CNIL, Méthodologie de référence pour les traitements de données personnelles opérés dans le cadre des recherches biomédicales, MR-001, 2006, www.cnil.fr

# INFOS SOCIO-LINGUISTIQUES

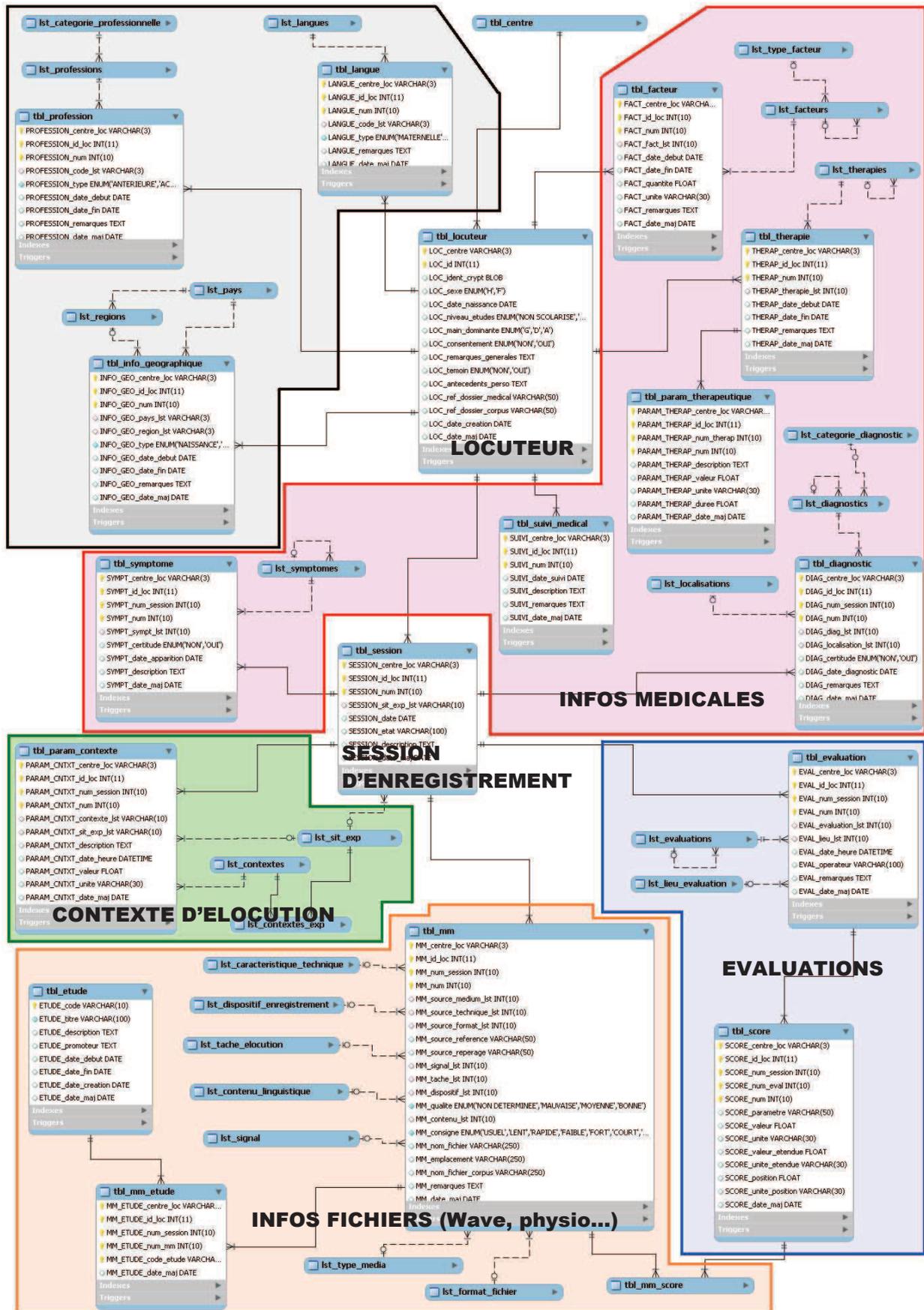


Figure 1 : Modèle conceptuel généralisable d'une base de données de « parole pathologique »