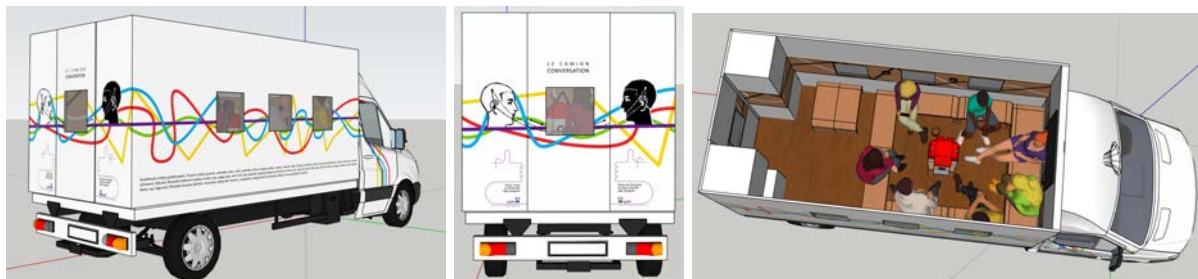


Projet Le Camion Conversation

(Laboratoire mobile de la Cité internationale de la langue française)
V1 15 septembre 2019



Le camion conversationnel est un projet qui s'inscrit dans un vaste programme qui associe recherches linguistiques, arts, politiques culturelles et valorisation.

Partie émergée du programme, ce dispositif innovant permet de recueillir des données uniques sur la langue française, dans la diversité de ses usages, et les langues en contact avec celle-ci. Ces données permettent d'assurer ainsi un fort développement de la recherche et de l'innovation en linguistique fondamentale, sociolinguistique, traitement automatique du langage, didactique, tout en s'inscrivant dans une démarche de conservation d'un patrimoine sonore et de sa valorisation.

L'histoire des Archives de la parole

C'est l'auteur de la grande Histoire de la langue française, Ferdinand Brunot, qui le premier s'est préoccupé d'enregistrer et de conserver les traces sonores de faits de langue, en créant les fameuses Archives de la parole en 1911 à partir d'un projet de collectes sur l'ensemble du territoire français. Il s'agissait, grâce au phonographe, d'enregistrer, d'étudier et de conserver des témoignages oraux de la langue parlée. Guillaume Apollinaire, Émile Durkheim, Alfred Dreyfus, etc., mais aussi nombre de locuteurs « anonymes » ou étrangers vont laisser le témoignage de leur voix aux Archives de la Parole. Trois cents enregistrements sont ainsi réalisés entre 1911 et 1914. Ces enregistrements sur rouleaux de cire, pieusement conservés, forment le fonds premier du département de l'audiovisuel à la BnF. Ils sont maintenant un élément phare de la bibliothèque numérique Gallica.

<https://gallica.bnf.fr/html/und/enregistrements-sonores/archives-de-la-parole-ferdinand-brunot-1911-1914?mode=desktop>



Le programme *Corpus de la parole* DGLFLF-CNRS-BnF

Le programme Corpus de la parole, initié par la DGLFLF en 2004 en collaboration avec le CNRS et la BnF, permet de conserver et valoriser les corpus oraux en français et langues de France. C'est un enjeu capital pour la politique linguistique contemporaine de notre pays. Alors que la plupart des langues européennes disposent de corpus oraux accessibles en ligne, et souvent gratuitement, un tel outil n'existait pas en France jusqu'à récemment, ce qui n'était pas sans conséquences néfastes

pour la visibilité et la vitalité du français et des langues parlées en France. Enjeu pour la recherche linguistique et pour le développement de l'ingénierie linguistique, mais aussi pour l'enseignement de ces langues, pour la sauvegarde et la diffusion du patrimoine oral, il prend actuellement la forme du site Trésors de la parole qui donne accès à plusieurs milliers de documents sonores dans un projet innovant au cœur du Web de Données : <http://tresorsdelap parole.culture.gouv.fr/>

Les données du site Trésors de la parole sont un élément central du projet C088-Instrument numérique de la DGLFLF, portail des données qualitatives et quantitatives issues des recherches sur les pratiques du français et des langues de France. L'ensemble des données sont sauvegardées et pérennisées dans la plateforme Cocoon (CNRS-TGIR Huma-Num) utilisée par les laboratoires de linguistique du CNRS et des universités.



Les projets Arts numériques et Sciences

Parallèlement au développement d'initiatives de diffusion et valorisation des données de la recherche, la DGLFLF a initié des projets au croisement des arts numériques et des sciences. Il s'agit notamment de favoriser un dialogue sur la collecte et l'utilisation de « pratiques linguistiques » à partir de démarches artistiques et de démarches scientifiques. Cette collaboration, très fructueuse avec l'engagement de l'artiste Guykayser, a permis de poser les bases d'un programme de dispositifs conversationnels.



Le programme « Dispositifs Conversationnels »

Le programme « Dispositifs Conversationnels » vise la constitution, la conservation et la valorisation d'un patrimoine sonore constitué de milliers d'enregistrements de conversations quotidiennes en français et en langues de France.

Il repose sur un partenariat avec les Infrastructures de recherche et les plateformes du MESRI et qui gèrent les données scientifiques dans le cadre des programmes nationaux et européens de la Science Ouverte. La plateforme Cocoon assure le stockage, la conservation et l'exposition des données linguistiques du programme en lien avec la BnF.

Les dispositifs conversationnels sont tous composés à partir de trois composants :

- Une base de données et une plateforme technologique gérant les données linguistiques à distance (la plateforme Cocoon).
- Une interface interactive pour écouter et naviguer dans ce riche patrimoine sonore.
- Un dispositif artistique de captation des pratiques linguistiques au sein d'un objet conversationnel.

Le Camion Conversation (laboratoire mobile de la Cité internationale de la langue française)

Le *Camion Conversation*, cent ans après le projet de Ferdinand Brunot doit permettre de sillonner la France pour collecter et traiter immédiatement des enregistrements de pratiques linguistiques en Français et langues de France. Il s'agira de constituer un trésor unique, témoignage permanent et dynamique de la langue française et des langues en contact avec celle-ci. Ce patrimoine, issu de projets artistiques et scientifiques avec la participation de centaines de citoyens, sera conservé par la BnF, et constituera une ressource unique pour la recherche en linguistique et didactique mais aussi pour les entreprises des technologies de la langue.

Pilotage du projet (DGLFLF et CNRS) :

Direction artistique : Guykayser

Direction scientifique : DGLFLF-Observatoire des pratiques linguistiques

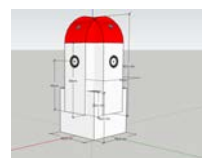
Infrastructure de recherche : TGIR Huma-Num et Plateforme CNRS Cocoon

Laboratoires de Recherche LLL UMR7270, MoDyCo UMR 7114

Le Camion Conversation, laboratoire mobile

Le Camion Conversation, laboratoire mobile unique en son genre est constitué de plusieurs modules.

Le premier, la borne interactive numérique, contient l'ensemble des microphones, haut-parleurs et matériel informatique permettant la captation et la restitution des enregistrements sonores. Cette borne, mobile, est réutilisable dans l'ensemble des dispositifs conversationnels (capsule, bateau, etc.)



Le second, la cellule aménagée, constitue le cadre situationnel favorisant le déroulement de conversations ou de tout autre situation de pratiques langagières



Le troisième, apporte une connexion entre le dispositif de captation-restitution et la plateforme technologique qui stockera et gèrera l'ensemble des données numériques dans le respect des pratiques actuelles des infrastructures de recherche françaises et européennes (principes européens FAIR pour une Science Ouverte). Interface de versement et



d'accès aux enregistrements connectée à la plateforme de traitements des données.

Le portrait sonore de la langue française et des langues en France

Le Camion Conversation effectuera des collectes sur l'ensemble du territoire afin de constituer la plus grande base de données sur le français et les langues parlées en France. Il s'agit de réaliser le portrait sonore des pratiques langagières en France en collectant des enregistrements de conversations au gré du parcours du camion à travers la France. La collecte principale nécessitera un premier circuit, dans la perspective du projet Villers-Cotterêts, d'une cinquantaine d'étapes d'ici 2022. Chaque étape impliquera un lieu (musée, médiathèque, école, marché, maison de quartier, fête foraine...) et pourra donner lieu à un projet local avec un artiste afin de construire un réseau de recherches culturelles participatives.

Un animateur scientifique aura en charge la préparation et la réalisation du circuit de collecte et la médiation scientifique et culturelle de l'ensemble du programme.



Les exploitations

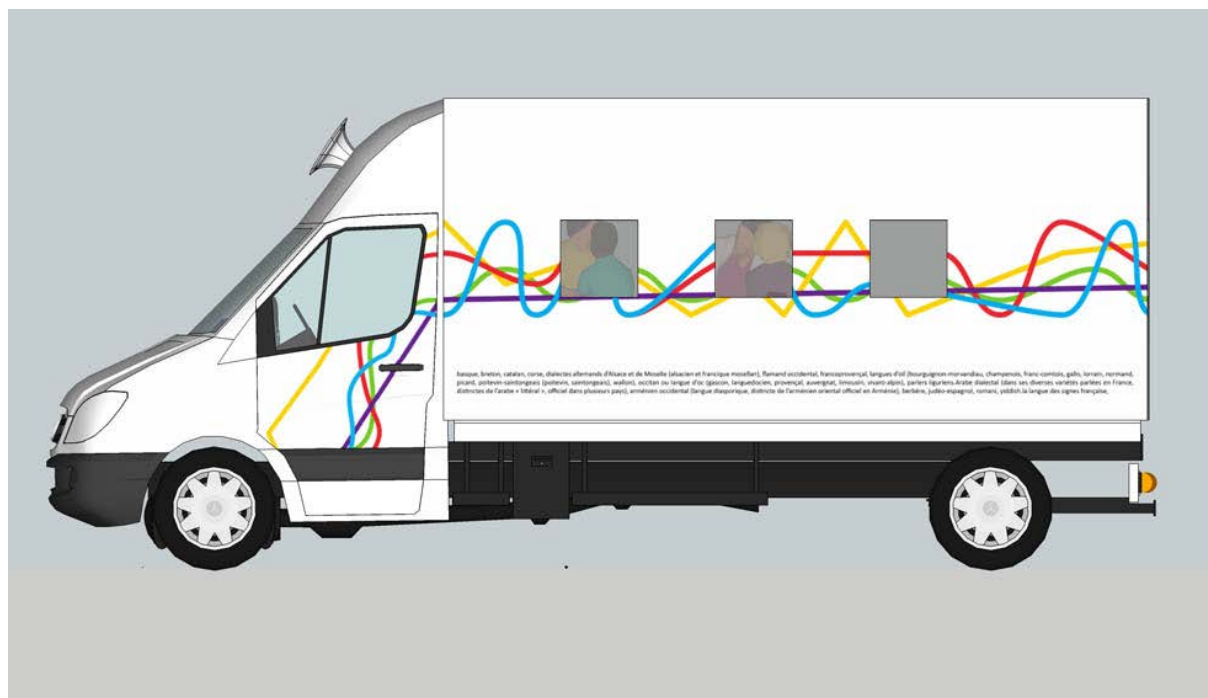
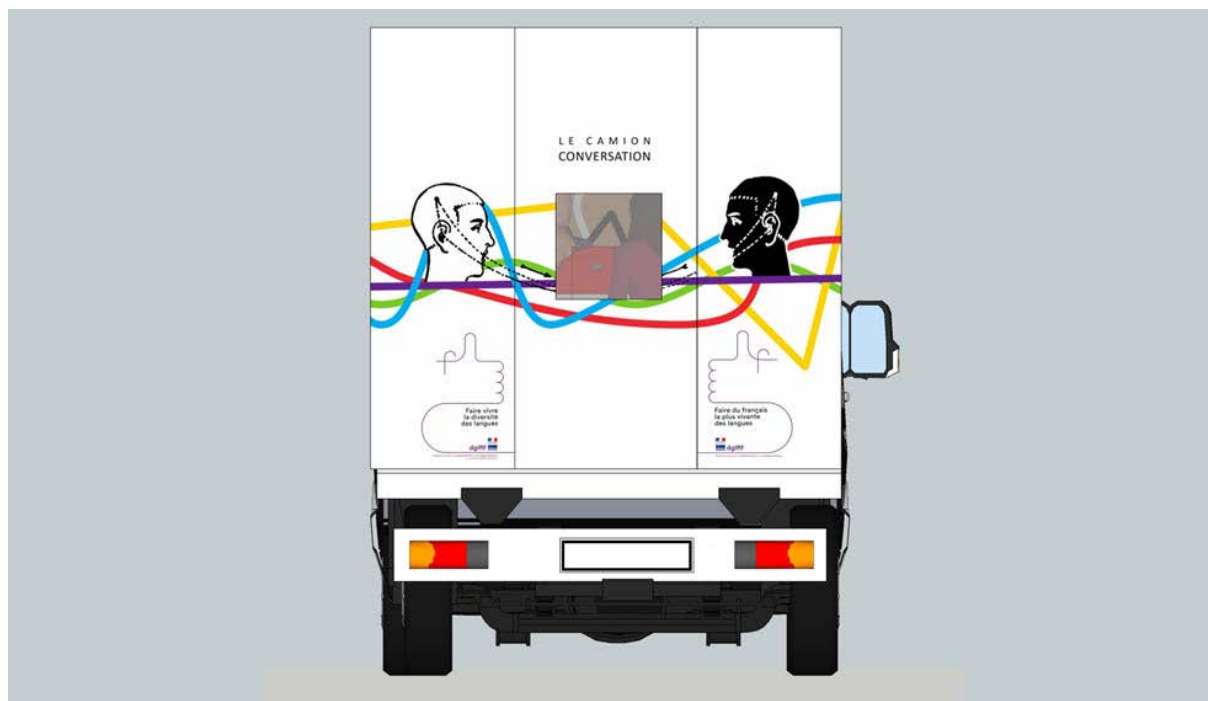
Premières exploitations innovantes prévues

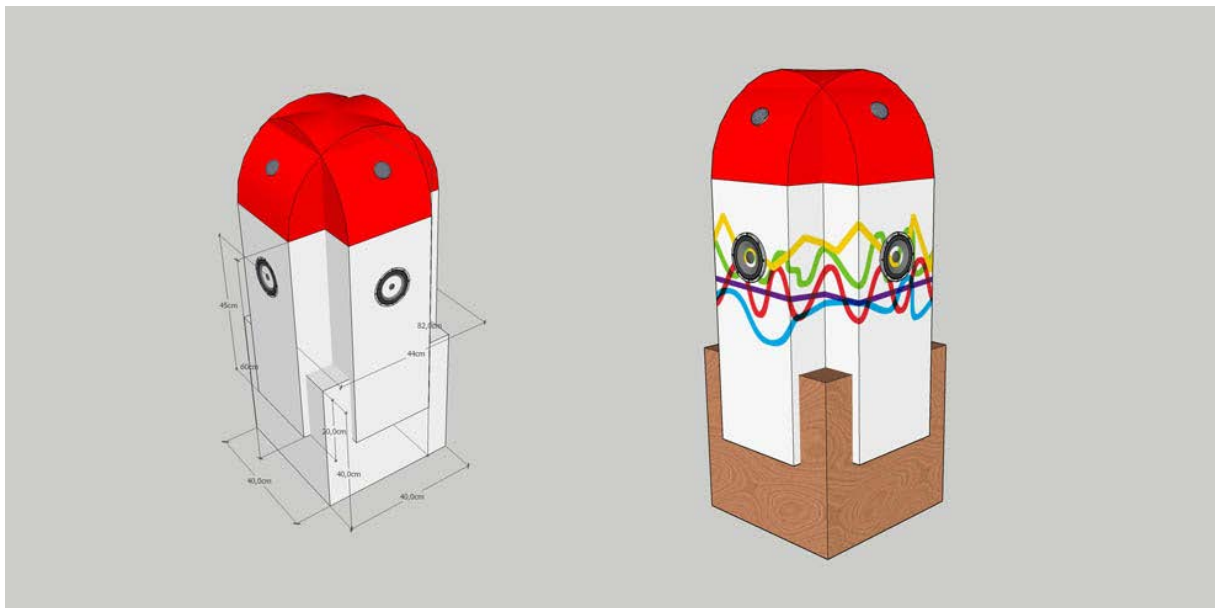
- Analyses linguistiques et sociolinguistiques
- Applications didactiques pour l'enseignement du français
- Ressources pour l'Intelligence Artificielle
- Applications en technologie de la langue (assistants vocaux, reconnaissance de la parole, synthèse de la parole,...)
- Mise à disposition des données (données ouvertes) à partir du portail « C088-Instrument numérique » de la DGLFLF.

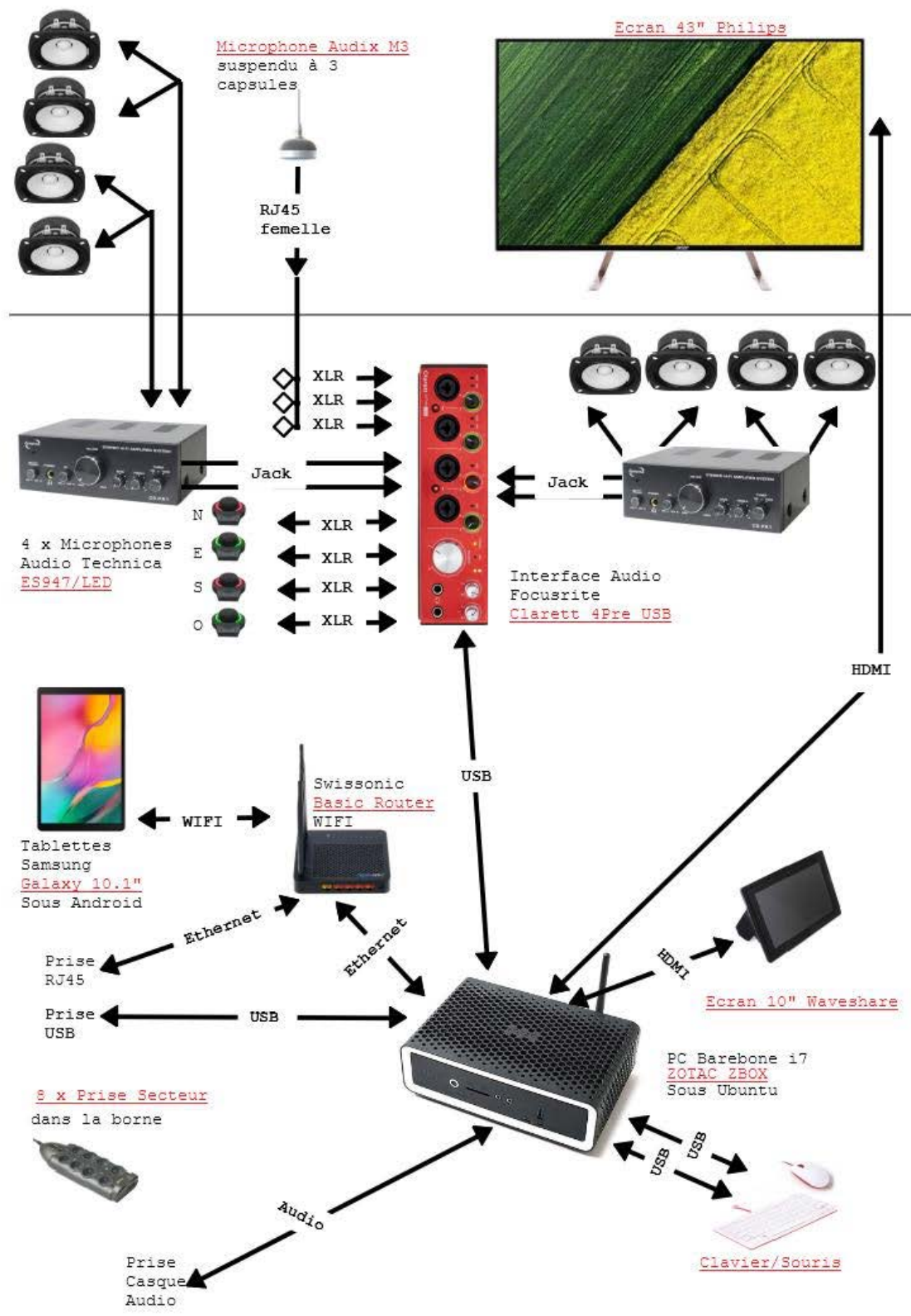
Calendrier du projet

Calendrier projet (trimestres)	4eT 2019	1eT 2020	2eT 2020	3eT 2020	4eT 2020	1eT 2021	2eT 2021	3eT 2021	4eT 2021	1eT 2022	2eT 2022	3eT 2022	4eT 2022
Directions artistique et scientifique													
Conception de la borne interactive													
Fabrication de la borne interactive													
Conception du camion-laboratoire													
Fabrication du camion-laboratoire													
Aménagement du camion-laboratoire													
Préfiguration, tests collectes et animations													
Préfiguration tests scientifiques et didactiques													
Collectes													
Exposition													
Exploitation scientifique													
Valorisation des technologies													

Annexes







Annexe : Liens vers les vidéos et documents

Le camion en 3D : <http://bit.ly/2ZQghdT>

La capsule sur FR3 : <http://bit.ly/2LMf0iT>

Liens du dossier

Les Archives de la parole

- Présentation dans Gallica : <https://gallica.bnf.fr/html/und/enregistrements-sonores/archives-de-la-parole-ferdinand-brunot-1911-1914?mode=desktop>
- Description du fonds à la BnF : <https://archivesetmanuscripts.bnf.fr/ark:/12148/cc78617w>

Arts numériques et sciences

- Le site de Guykayser <https://guykayser.autoportrait.com/autoportraits-collectifs>
- Le site de Gérard Parésys <http://gerard.paresys.free.fr/>
- Le site de Laurent Malys <https://www.laurent-malys.fr/>

Dispositifs Conversationnels

- Trous de mémoire
 - <http://www.autoportrait.com/trousdememoire/>
 - <http://gerard.paresys.free.fr/TrousDeMemoire/index.html>
- Les voix du bac à sable
 - <https://guykayser.autoportrait.com/autoportrait-collec/les-voix-du-bac-a-sable-les-paysages-du-larynx>
- La capsule conversationnelle
 - <https://guykayser.autoportrait.com/autoportrait-collec/la-capsule-conversationnelle-creation-2018>
 - <http://gerard.paresys.free.fr/Journal/Capsule.html>