

Compte rendu sur le travail de documentation des enregistrements des enquêtes de l'Atlas Linguistique des Cotes de l'Atlantique et de la Manche (ALCAM)

travail effectué du 06-10-2014 au 26-02-2015

Adrien Desseigne

Plan du rapport

1. Rappel
2. La découpe des enregistrements
3. Le nommage des fichiers découpés
4. L'écriture des métadonnées
 - 4.1. La formulation des titres
 - 4.2. Le choix de la localisation
 - 4.3. Descriptions des informateurs
 - 4.4. Détails sur le contenu des pistes découpées
 - 4.5. Linguistic-field
5. Problèmes rencontrés
6. Listes des points d'enquêtes documentés
7. Listes de ce qu'il reste à faire
 - 7.1. Numériser les bandes suivantes
 - 7.2. Découper et documenter les bandes et cassettes suivantes
 - 7.3. Compléter les métadonnées des bandes et cassettes suivantes et éventuellement les redécouper en fonction des différentes parties de l'ALCAM
 - 7.4. Bande du NALBB
8. Listes des fichiers transmis
 - 8.1. Les fichiers sons et les métadonnées
 - 8.2. Les fichiers annexes

1. Rappel

L'ALCAM est un projet d'atlas linguistique, mis en place par Patrice Brasseur dans les années 1980 qui comprend 82 points d'enquête depuis Bray-Dunes en Flandre française jusqu'à Ciboure au Pays Basque français. Le questionnaire totalise 400 questions concernant la description du rivage et des éléments marins, la navigation, la pêche, la faune et la flore marines. En Bretagne celtique, une vingtaine d'enquêtes ont été réalisées à partir de 1983 par Jean Ropars, Claude Guillou, Armelle Audic et Jean-Michel Menesguen et coordonnées par Jean Le Dû alors professeur de celtique à l'université de Brest.

2. La découpe des enregistrements

Le questionnaire de l'ALCAM comporte plusieurs parties thématiques. Pour l'archivage et la documentation de ce fond, j'ai découpé le contenu des bandes en fonction de ces parties. Voici les noms proposés par Daniel Le Bris pour nommer ces parties :

- 1- Géomorphologie côtière et maritime (questions 1 à 29)
- 2- L'embarcation (questions 31a à 56)
- 3- Le pêcheur (questions 60a à 77b)
- 4- Aspects généraux du poisson (questions 80a à 96)

- 5- Les poissons (questions 101 à 185)
- 6- Les coquillages (questions 191 à 212)
- 7- Zoophytes, vers, crustacés, céphalopodes (questions 221 à 263)
- 8- Les algues (questions 271 à 283)
- 9- Les oiseaux marins (questions 290 à 308)
- 10- Techniques de pêche (questions 311a à 344)

Lorsque c'était possible, j'ai aussi distingué dans les découpes les présentations des informateurs et les discussions en marge de l'enquête (sur des thèmes souvent relatifs au milieu marin tout de même).

Parfois le déroulement de l'enquête se prêtait mal à une découpe en fonction d'une de ces parties (compléments d'enquêtes, mélange de questions et de discussions). C'est pourquoi parmi les extraits découpés certains contiennent du « contenu mixte » (cf. **4.1. La formulation des titres**).

3. Le nommage des fichiers découpés

Tous les fichiers découpés suivent le modèle de nommage suivant :

CRBC_ALCAM_COTEPRODUCTEUR_FACE_NUMERODEPISTE.wav

ex : CRBC_ALCAM_C028_A_1.wav

4. L'écriture des métadonnées

Pour l'écriture des métadonnées au format Dublin Core étendu selon les standards OLAC, j'ai essayé de suivre les recommandations de Séverine Guillaume et Jean-Baptiste Pressac.

Voici néanmoins quelques précisions sur certains choix que j'ai fait et qui pourraient être revus avant l'archivage des données.

4.1. La formulation des titres

Ils suivent toujours le modèle suivant :

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à *[nom du point d'enquête]*, locuteur *[initiale]*, *[nom donnée à la partie de l'enquête]*

ex :

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à Roscoff, locuteur L.C., Les algues

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à l'Île de Sein, locuteur H.G., Présentation de l'informateur

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à l'Île de Sein, locutrice Y.B., Conversations

Lorsque le contenu d'une des parties du questionnaire se retrouve sur plusieurs pistes (déroulement de l'enquête discontinue, changement de face...), j'ai indiqué un ordre en ajoutant des numéros à la

fin des titres :

ex :

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à l'Île de Sein, locuteur H.G., Les poissons (1 sur 3)

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à l'Île de Sein, locuteur H.G., Les poissons (2 sur 3)

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à l'Île de Sein, locuteur H.G., Les poissons (3 sur 3)

Parfois, lorsque deux parties se trouvent sur une même piste et qu'aucune découpe n'a été réalisée, j'ai indiqué le nom des deux parties à la suite dans le titre :

ex :

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à Roscoff, locuteur L.C., Présentation de l'informateur, Les coquillages

Dans quelques cas, lorsque les contenus d'une piste comporte un ensemble de questions appartenant à plusieurs parties du questionnaire, j'ai simplement indiqué « Compléments d'enquête » dans le titre :

ex :

Atlas Linguistique des Côtes de l'Atlantique et de la Manche (ALCAM) : Enquête dialectologique à Landéda, locuteur V.C., Compléments d'enquête

4.2. Le choix de la localisation

Le plus souvent, il n'a pas été difficile d'associer l'enquête à un lieu géographique puisque l'informateur est originaire du lieu où s'est déroulée l'enquête. J'ai pu indiquer ce lieu dans les métadonnées de trois manières :

- en partant de l'échelle de la France à l'échelle de la commune (et parfois même du village) :

ex : France, Bretagne, Finistère, Roscoff
France, Bretagne, Finistère, Crozon, Morgat

- avec des coordonnées GPS :

ex : east=-3.9853; north=48.7261;

- quelques fois à partir du référentiel *Getty Thesaurus of Geographic Names (TGN)* :

ex : 1008187 (pour l'Île de Molène)

Parfois, l'association de l'enquête à un lieu géographique m'a posé problème, en particulier lorsque :

- le locuteur n'est pas originaire du point d'enquête

- l'enquête concernant un point de l'atlas a lieu dans une autre localité
- plusieurs informateurs originaires de lieux différents sont interrogés en même temps pour un point d'enquête d'où ils ne sont pas originaires

Puisqu'il s'agit de données de type dialectologique, j'ai fait le choix de toujours privilégier dans les métadonnées le lieu d'origine des informateurs même lorsqu'il ne correspond pas au point d'enquête de départ. Voici les cas particuliers sur lesquels je suis tombés et les « lieux d'enquêtes » que j'ai finalement retenus :

- point d'enquête « Roscoff » avec informateur de l'Île de Batz => « enquête à l'Île de Batz »
- point d'enquête « Le Conquet » avec informateur de Plouguerneau => « enquête à Plouguerneau »
- point d'enquête « Le Conquet » avec informateurs de Plouguerneau et de Plougonvelin en même temps => « enquête à Plouguerneau et Plougonvelin »
- point d'enquête « Le Conquet » avec un informateur de Plouider => « enquête à Plouider »
- point d'enquête « Le Conquet » avec un informateur de Ploumoguier => « enquête à Ploumoguier »
- point d'enquête « Le Conquet » avec un informateur né à Plouzané, mais qui a vécu très tôt à Plougonvelin le reste de sa vie => « enquête à Plougonvelin »

Un autre cas particulier a été le point d'enquête Portsall. Ce port qui appartient à la commune de Ploudalmézeau est limitrophe avec la commune de Landunvez d'où est originaire l'un des deux informateurs interrogés. J'ai indiqué dans les coordonnées GPS uniquement Portsall, mais j'ai gardé une distinction dans la forme écrite du lieu d'enquête :

France, Bretagne, Finistère, Ploudalmézeau, Portsall
France, Bretagne, Finistère, Landunvez, Portsall

Enfin, pour les enquêtes à « Morgat », un informateur est bien originaire de ce village de la commune de Crozon alors que l'autre est originaire d'un village plus au sud dans cette même commune. Dans les titres, j'ai donc indiqué qu'il s'agit plutôt d'enquêtes à Crozon alors dans les coordonnées GPS, j'ai donné les coordonnées des villages en question et non de la commune en général.

4.3. Descriptions des informateurs

Dans la balise *description*, je n'ai donné que quelques informations succinctes sur les informateurs (nom et prénom, année de naissance, commune d'origine). Je n'ai pas indiqué toutes les informations disponibles sur les fiches d'enquêtes (notoriété dans la commune, religion, politique affirmée, marié ou non...). Certaines informations disponibles pourraient être ajoutées dans les métadonnées comme la profession, le niveau scolaire, la pratique du dialecte locale.

4.4. Détails sur le contenu des pistes découpées

Dans une balise *tableOfContents*, j'ai indiqué le contenu détaillé de chaque piste, à savoir :

- quelle partie du questionnaire ?
- quel(s) sujet(s) de conversations ?
- quel(s) numéro(s) de question ainsi que son intitulé
- les questions supplémentaires

J'ai utilisé le symbole / pour indiquer un changement de question et le symbole + pour indiquer la présence d'une question supplémentaire ou d'une précision sur un thème non prévu à la base dans le questionnaire de l'ALCAM :

ex : Les algues (271-Les algues marines: terme général / 272-Cladophora rupestris / +Cystoseira / 273-Laitue de mer (Ulva lactuca) / 275-Enteromorphes / 274-Codium / 275-Enteromorphes / 276-Chorda filum / 277-Laminaria digitata / +Laminaria hyperborea / +Saccorhiza polyschides / 279-Le fucus / +Ascophyllum nodosum / +Fucus serratus / +Fucus ceranoides / +Pelvetia canaliculata / +Fucus vesiculosus / 280-Le carragheen (Chondrus crispus) / 281-Zostera / +L'Herbier / +Le varech / 271-Les algues marines: terme général / 272-Cladophora rupestris / 273-Laitue de mer (Ulva lactuca) / 274-Codium / 275-Enteromorphes / 276-Chorda filum / 277-Laminaria digitata / +Laminaria hyperborea / +Saccorhiza polyschides / 278-Laminaria saccharina / 279-Le fucus / +Fucus serratus / +Ascophyllum nodosum / +Fucus vesiculosus / +Fucus ceranoides / 280-Le carragheen (Chondrus crispus) / 281-Zostera)

Toutes les questions sont mentionnées y compris celles qui n'ont pas amené de réponses de la part de l'informateur (souvent le cas !). L'intitulé des questions indiqué dans les métadonnées correspond le plus souvent à celui du questionnaire. J'ai noté parfois les différences de formulation de la part de l'enquêteur, mais pas systématiquement (ex : « 281-Les herbiers » au lieu de « 281-Zostera »). De même, je n'ai pas indiqué lorsque l'enquêteur demande à la fois le singulier et le pluriel d'un mot là où le questionnaire ne prévoit que le pluriel ou que le singulier.

4.5. Linguistic-field

Pour les extraits qui correspondent aux réponses du questionnaire j'ai mentionné un *linguistic-filed* OLAC avec pour valeur « *lexicography* » alors que les extraits de conversations porte la mention « *text_and_corpus_linguistics* ».

5. Problèmes rencontrés

Sur certaines pistes audio, on entend des bruits de téléphones portables captés au moment de la numérisation et qu'il faudrait sans doute renumériser avant archivage. Voici la liste de plusieurs de ces pistes :

CRBC_ALCAM_C020_A_2.wav
CRBC_ALCAM_C020_A_4.wav
CRBC_ALCAM_C020_A_6.wav
CRBC_ALCAM_C022_B_4.wav
CRBC_ALCAM_C024_A_1.wav
CRBC_ALCAM_C025_A_3.wav
CRBC_ALCAM_C044_A_1.wav

6. Listes des points d'enquêtes documentés

Je n'ai pu documenter qu'une partie des points d'enquêtes de l'ALCAM dont voici la liste :

05- Roscoff
06- Brignogan
07- Landéda
08- Portsall
09- Île de Molène
10- Île de Ouessant
11- Le Conquet
12- Plougastel

13- Crozon
15- Île de Sein

7. Listes de ce qu'il reste à faire

Voici le détail des tâches qui resterait à effectuer pour la suite du projet de documentation du fond ALCAM au CRBC :

7.1. Numériser les bandes suivantes :

ALCAM_021 (Groix)
ALCAM_024 (Quiberon)
ALCAM_025 (Séné)

7.2. Découper et documenter les bandes et cassettes suivantes :

ALCAM_021 (Groix)
ALCAM_022 (Belle-Ile)
ALCAM_023 (Belle-Ile)
ALCAM_024 (Quiberon)
ALCAM_025 (Séné)

ALCAM_C046 (Le Guilvinec)
ALCAM_C047 (Le Guilvinec)
ALCAM_C048 (Le Guilvinec)
ALCAM_C049 (Le Guilvinec)
ALCAM_C050 (Le Guilvinec)
ALCAM_C051 (Le Guilvinec)
ALCAM_C052 (Le Guilvinec)
ALCAM_C053 (Le Guilvinec)
ALCAM_C054 (Le Guilvinec)
ALCAM_C055 (Le Guilvinec)
ALCAM_C056 (Le Guilvinec)

7.3. Compléter les métadonnées des bandes et cassettes suivantes et éventuellement les redécouper en fonction des différentes parties de l'ALCAM :

ALCAM_013 (Plougrescant)

ALCAM_C030 (Plougasnou)
ALCAM_C031 (Plougasnou)
ALCAM_C032 (Plougasnou)
ALCAM_C033 (Plougasnou)
ALCAM_C034 (Locquémeau)
ALCAM_C035 (Locquémeau)
ALCAM_C036 (Locquémeau)
ALCAM_C037 (Plougrescant)
ALCAM_C038 (Plougrescant)
ALCAM_C039 (Plougrescant)
ALCAM_C040 (Plougrescant)
ALCAM_C041 (Plougrescant)
ALCAM_C042 (Loguivy-Mer)

Il s'agit des bandes que Jérémie Moysan avait commencé à documenter et qui ne sont pas sur le même modèle de métadonnées. Il faut voir si en plus de relire les métadonnées, on redécoupe aussi les extraits en fonction des sous-parties du questionnaire de l'ALCAM. Attention ! Les extraits de la bande ALCAM_C021 ont déjà été redécoupés et intégrés au reste des données documentées sur Portsall.

7.4. Bande du NALBB

Voir ce qu'il faut faire de la bande ALCAM_005 contenant des enquêtes du NALBB à Molène

8. Listes des fichiers transmis

8.1. Les fichiers sons et les métadonnées

J'ai commencé à mettre en ligne sur sharedocs l'ensemble des extraits que j'ai découpés ainsi que leurs métadonnées. Pour qu'on puisse s'y retrouver plus facilement, j'ai créé un dossier CDD_ADESSEIGNE à part dans le dossier ALCAM pour y déposer mon travail. J'ai regroupé les extraits dans des dossiers par point d'enquête. Un fichier XML de métadonnées accompagne chaque dossier qu'il faudra fusionner avant l'archivage.

8.2. Les fichiers annexes

- CCD_ALCAM_ADESSEIGNE_COMPTE-RENDU.odt : Le présent compte-rendu sur le travail effectué.
- CCD_ALCAM_ADESSEIGNE_HORAIRES.txt : Le détail des heures de travail effectuées. Il pourra être utile pour estimer le temps nécessaire pour finir la documentation du fond.
- CCD_ALCAM_ADESSEIGNE_SOMMAIRE_DES_BANDES.odt : Un listing simplifié des découpes effectuées sur chaque bande.
- CCD_ALCAM_QUESTIONNAIRE_BRASSEUR.odt : Le questionnaire de l'ALCAM mis au propre avec le nom des parties proposées par Daniel Le Bris