### Version définitive des rubriques des métadonnées remplies de AR03

=====================================================================

Sur les fichiers eux-mêmes (fichier audio unique .mp3 et le fichier .cha qui contient la transcription alignée avec l’audio)

L’alignement entre les fichiers texte et audio a été effectué avec un séquençage plus long que les 20 secondes prévues. Il peut être plus compliqué de naviguer dans le texte de transcription car le récit contient des monologues relativement longs.

Informations générales sur le mode de transcription

Contexte de l’enregistrement

Remarques sur l’oral (en relation avec le guide de transcription)

La prononciation des la lettre *v* est *b* (comme en espagnol), notamment en début de mots ex. *Valence*, *Vignolet*.

Ajouter les remarques de Carmen sur l'espagnol.

Incertitudes dans la transcription mais une interprétation est choisie

Choix de transcription dans la version déposée :

* *la veuve elle avait*, alors qu'initialement *la veuve* était transcrit par un mot espagnol non compris.
* *des ouï-dire <pron=cuidir>*, initialement transcrit en *des Cuidir* (hypothèse non retenue que ce pourrait être un lieu des cuidir(e)s)

Interprétation des Npr de lieux et de personnes

Pour les noms de lieux :

Des choix de transcription :

* *une ville de la Gérone de Gerona* (ajouter balise révision ?), initialement *une ville de la <pron=lieu> de <pron=lieu>*. L’identifiant du lieu est celui dans GeoNames de la *Provincia de Girona* avec nom alternatif *Gerona*, différent de la ville de Gijón a priori non citée.
* Bourg-Madame, mentionné deux fois, initialement *Burlada*. L’identifiant est celui de GeoNames, même si normalement les lieux en France ont un identifiant BDNyme (trouver l’identifiant dans BDNyme).
* Vignonet (avec le *v* prononcé *b*), ce qui semble logique par rapport à la mention de *Saint-Émilion*. L’identifiant est celui de GeoNames, même si normalement les lieux en France ont un identifiant BDNyme (trouver l’identifiant dans BDNyme).

Mention de *San Miguel de los Reyes*, mentionné deux fois : c’est un monastère à Valence en Espagne qui a servi de prison. Il n’est pas présent dans GeoNames donc pas d’identifiant à indiquer. Prononciation dans les deux mentions de *los* sans le *s* final.

Pour les noms de personnes :

* *monsieur Despagne*, initialement *monsieur d'Espagne*. Ce c’est pas certain que ce soit la bonne orthographe, mais c'est a priori un nom de famille. Il faudrait un code d'anonymisation (NP167 ?)

### Commentaires Déborah puis Carmen, Catherine, Laurence

=====================================================================

Commentaires Déborah/réponses, et Carmen, Catherine, Laurence

AR03

Balisage des NprLieu avec les identifiants de gazetier : ok

Vérification des codes de noms propres de personnes : ok

==================================================================================================================

Corrections systématiques :

==================================================================================================================

Tous les [] remplacés par des <> (<euh>, <silence>, <pause>, <pron=nonstd>, <pron=pi>...)

<mdd> ah/eh/hé/oh/hein/euh/ben/bah/beh </mdd> --- <ah/eh/hé/oh/hein/euh/bah/beh>

<mdd> et bien </mdd> --- <mdd> eh bien </mdd>

<mdd> mais/oui </mdd> >> corriger si nécessaire

<hum> --- <hmm>

<rire/pleur> --- <rireentendu/pleurentendu>

<silence> --- <si>

<revis> </revis> --- <revision> </revision>

<espagnol></espagnol> --- <lg=espagnol></lg=espagnol>

XXX là --- XXX-là

y'a/y'en > y a / y en

et caetera

<audiomanquantsep/fus> --- effacés

Ajouts de majuscules, par exemple :

les républicains --- les Républicains

les espagnols --- les Espagnols

==================================================================================================================

Corrections notables :

==================================================================================================================

Retranscription :

90 une ville de la <pron=lieu> de <pron=lieu>

une ville de la Gérone de Gerona

<LJ> J’ai l’impression que c’est une révision également, et j’ajoute au dictionnaire Npr lieux l’identifiant de la Província de Girona (nom alternatif dans GéoNames : Gerona) : une ville <revision> de la Gérone </revision> de <lieu id=6355230> Gerona </lieu>

264 la bonne sœur La Cornette

la bonne sœur à cornette

<LJ> oui !

266-267 <mdd> vous savez </mdd> votre mère elle est entrée dans mon lycée quand elle sera morte

<mdd> vous savez vous savez </mdd> elle est en train d'agoniser quand elle sera morte

<LJ> en effet...

626 &478 il avait

la veuve elle avait

<LJ> ok pour ligne 626 (632 à présent), aussi ligne 478 ?

<DL> Il est écrit dans le fichier “MA01-FS15-VP12-DB04-CC07-EB06-AE02-FR14-JE09-AR03JV17-JO11-Remarques2411” que “&478” est l’un des codes employé pour indiquer un mot espagnol non compris. Quand je les rencontre, je remplace ces codes par ce qu’il faut. Dans ce cas-ci, ce n’était pas un mot espagnol, c’était “la veuve”.

==================================================================================================================

Questions/Réponses :

==================================================================================================================

Gazetier :

<DL> "Burlada" en double.

<DL> "Madrid" en double.

<DL> "Valence" en double.

<DL> Ajouter "Cuidir" ???

<DL> Remplacer "Gijon" par "Gérone/Gerona", quel id ?

<DL> Remplacer "Burlada" par "Bourg-Madame" (110, 134), quel id ?

<DL> Ajouter "Vignonet" (163, 238), quel id ?

<DL> Ajouter "San Miguel de Reyes/San Miguel de los Reyes" (545, 689), quel id ?

<LJ> Mise à jour sur le FTP :

* Suppression des doublons
* Pour “Cuidir” : rien, ce serait plutôt “ouï-dire” pour moi...
* Suppression de “Gijón:5-id=3121424” et ajout de “Gerona:5-id=6355230”
* Suppression de Burlada et ajout de “Bourg-Madame:5-id=6615247”. C’est un identifiant GeoNames mais de code ADM4 (pour une unité administrative) et non PPL (pour une villes normalement). Voir par la suite pour trouver l’identifiant dans BDNyme comme c’est une ville en France.
* Ajout de “Vignonet:5-id=2969095”, cela semble en effet logique par rapport à Saint-Émilion. Elle prononce souvent le “v” comme un “b” ex. “Valence” ligne 696.
* En fait pour San Miguel de los Reyes, c’est un monastère à Valence qui a servi de prison. Du coup, comme il n’est pas présent dans GeoNames, pas d’identifiant à indiquer. Ligne 545 : j’entends “San Miguel de los Reyes” avec “los” prononcé sans le “s” final. Même prononciation ligne 689.

<DL> 61 par les échos de ouï-dire <pron=cuidir>

par les échos <revision> de Cui- </revision> de Cuidir

"Cuidir" = lieu ? (Ou peut-être "des cuidir(e)s" ?)

<LJ> Je pencherai pour “ouï-dire” avec une prononciation particulière, ou si vraiment nous ne sommes pas sûres, indiquer <pron=pi>.

<DL> 175 "monsieur d'Espagne" que j'ai remplacé par "monsieur Despagne"

Pas sûr que ce soit la bonne orthographe, mais c'est un nom de famille, il faudrait un code d'anonymisation (NP167 ?)

<LJ> Indiquer dans les métadonnées ? Prononciation des “v” en “b”, par exemple en début de mots “Vignolet”, “Valence”.

<LJ> Voir pour diminuer la durée entre les bullets d’alignement, c’est assez compliqué de naviguer dans entre l’audio et le texte comme il y a d’assez longs monologues.

<DL> Pour ma part, j’alterne entre F4 et “Sonic Mode”. Quand je souhaite ne réécouter qu’un mot ou deux, quelques secondes seulement, j’utilise “Sonic Mode” le lecteur média de Clan. On peut lancer “Sonic Mode” :

- soit en allant dans “Mode” > “Sonic Mode”

- soit en utilisant le raccourci “Esc+0”