### Version définitive des rubriques des métadonnées remplies de MA01

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

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 de 20 secondes environ.

Informations générales sur le mode de transcription

Deux courts passages n’étaient pas transcrits dans les premières versions de transcription et le sont à présent.

A ajouter dans "lexiqueOnomatopeesEtInterjectionsMATRICIEL" : "justement", "ou".

Contexte de l’enregistrement

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

Ajouter les remarques de Carmen sur l'espagnol.

Incertitudes dans la transcription mais une interprétation est choisie

Interprétation des Npr de lieux et de personnes

Pour les noms de lieux :

* *Neuvic sur l'Isle* : le nom officiel est Neuvic
* Une mention de la *Dordogne* le département (et non pas la rivière normalement)
* Les identifiants des communes françaises citées Monpazier et Les Lèches sont issues de Geonames (et non de la BD Nyme)

Pour les noms de personnes :

* *Angel* --- *AV16* (Je pense qu’il parle de AV16) <LJ> j’ai l’impression que c’est bien une personne à anonymiser.

### Commentaires Déborah puis Carmen, Catherine, Laurence

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

Commentaires Déborah/réponses, et Carmen, Catherine, Laurence

MA01

Ajout du header (@Begin ... @End)

Ajout d'une \* à chaque tier

Alignement avec séquençage de 20 secondes environ

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> euh/hein </mdd> --- <euh/hein>

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

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

<mdd> ah/eh/hé/oh/bah </mdd> --- <ah/eh/hé/oh/bah>

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

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

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

XXX là --- XXX-là

<audiomanquantsep/fus> --- effacés

Ajouts de majuscules, par exemple :

les républicains --- les Républicains

les espagnols --- les Espagnols

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

Corrections notables :

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

Transcription : [deux passages non transcrits dans la transcription originale]

7 \*MA01FIE : alors comment on commence là ?

14-19 elle va témoigner de d'MA01 non non c'est pas ça MA01 ah c'est au- toi aussi

c'est MA01 <bc> ah ah l'autre jour je disais c'était c'était un jambon d'accord

\*MA01FIE : ouais mais c'est pas un jambon <rireentendu>

\*CH\_OD1 : c'est pas que j'ai comparé <rireentendu> d'accord d'accord

Correction :

127 C<mdd> ah </mdd>ors

Cahors

<DL> Juste une pensée émue pour une maladroite expression régulière !

<LJ> oui !

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

Questions/Réponses :

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

<DL> 61 : "TODT" est écrit en lettres capitales. C'est voulu ou je remplace par "Todt" ?

<LJ> C’est vrai qu’a priori ce n’est pas un acronyme donc “Todt”.

<DL> A ajouter dans "lexiqueOnomatopeesEtInterjectionsMATRICIEL" : "justement", "ou".

<DL> 433 : "Angel" --- “AV16” (Je pense qu’il parle de AV16).

<LJ> j’ai l’impression que c’est bien une personne à anonymiser.

Gazetier :

<DL> "Echourgnac" ou plutôt "Échourgnac" avec accent aigu ? (occurrence l. 95)

<DL> "Evreux" ou plutôt "Évreux" avec accent aigu ?

<DL> Ajouter "Monpazier:id=2993240" ? (occurrence l. 102)

<DL> Ajouter "Les Lèches:id=3000501" ? (occurrence l. 127)

<DL> Deux orthographes "Neuvic-sur-l'Isle" et "Neuvic sur l'Isle" : garder seulement "Neuvic sur l'Isle" ?

<DL> Il y a deux id pour la Dordogne : "id=DEPARTEM0000000000000027" et "id=PAIHABIT0000000049526714". A quoi correspond le 2e id ?

<LJ> La nouvelle version du dictionnaire est sur le FTP avec : "Échourgnac", "Évreux", “Monpazier”, “Les Lèches”, “Neuvic sur l'Isle” (note : le nom officiel est Neuvic).

Pour la Dordogne, le premier correspond au département et le second à la rivière normalement. J’enlève le second comme il ne doit pas y avoir d’occurrence.

Note : demander l’accès au serveur de données pour lire les identifiants des communes en France (“Monpazier”, “Les Lèches”) et vérifier que le second identifiant de la Dordogne est bien celui de la rivière.