### Version définitive des rubriques des métadonnées remplies de LC08

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

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.

L’alignement est un peu décalé.

Informations générales sur le mode de transcription

Contexte de l’enregistrement

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

Des prononciations hispanisantes non systématique mais fréquentes concernent les lettres et mots suivants : *le* [le], *v* [b], *j* [j]. Au vu de cette fréquence, aucune balise n’est ajoutée lorsque cette prononciation est effectivement constatée.

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 :

* Des noms modifiés, ajoutés ou supprimés par rapport aux premières transcriptions, même si une incertitude peut demeurer sur les choix de transcription :
  + *Condal*. Il pourrait en fait s’agir de Condat qui est à l’est de la commune de Fumel (comme Monsempron qui est mentionné comme une limite et est situé à l’ouest de Fumel). Condat est un village qui a été absorbé par la commune de Fumel et est donc à présent un quartier de Fumel. Il n’est donc pas associé à un identifiant car non présent dans les gazetiers utilisés. Remplacer par *Condat* ou *<pron=pi>* ? (2 mentions dans le récit)
  + *Siorac*. Transcrit Surac dans les premières versions sans identification. Il pourrait s’agir de Siorac juste au nord de Belvès et faisant partie du canton de Belvès. Le nom de cette commune fût Siorac, puis Siorac-et-Fongauffier, et à présent Siorac-en-Périgord.
  + *el Fuerte de Jaizkibel*. Jaizkibel étant la montagne à l’ouest de Fontarrabie, sur lequel est le Fort de Guadalupe.
  + *<pron=pi>* pour a priori un lieu, mais pas certain, mentionné près de la mention de la ville de Mosset. Dans les premières version de transcription : l*e désert* ou *les Deux-Sèvres*, remplacé donc par *<pron=pi>*.
* Une possible ambiguïté sur des formes de nom de lieux : *Villeneuve* correspond parfois à Villeneuve-sur-Yonne (plutôt dans la première moitié de la transcription), et parfois à Villeneuve-sur-Lot (plutôt dans la deuxième moitié de la transcription). Pour une annotation automatique à l’aide d’un gazetier, cela suppose de désambiguïser fgrâce au contexte.
* Pas d’annotation de *France* quand il apparaît dans *Banque de France*. À harmoniser ensuite avec les autres cas. Pour l’instant dans le guide de transcription, annotation des lieux comme *Radio <lieu id=xxx> Londres </lieu>* mais on pourra les supprimer ensuite.

Vérifier que Pampelune a été ajouté au dictionnaire des NPr du récit.

Pour les noms de personnes :

### Commentaires Déborah puis Carmen, Catherine, Laurence

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

Commentaires Déborah/réponses, et Carmen, Catherine, Laurence

LC08

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>

<pause/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 :

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

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

Questions/Réponses :

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

Gazetier :

<DL> "Cabezon de la Sal" --- "Cabezón de la Sal"

<DL> Ajouter "Méditerranée" ? (l. 599)

<DL> "Condal" : pas sûre de l'orthographe (l. 965, 1054). Remplacer par <pron=pi> ?

<DL> "Villeneuve" correspond parfois à "Villeneuve-sur-Yonne" (dans la première moitié de la transcription), et parfois à "Villeneuve-sur-Lot" (dans la deuxième moitié de la transcription).

<DL> Je n'ai pas annoté "France" quand il apparaît dans "Banque de France".

<LJ> Mise à jour du dictionnaire LC08 sur le FTP

* Ajout : Méditerranée:5-id=363196
* Modification : Cabezón de la Sal

Je pense que c’est “Condat” qui est à l’est de Fumel (// Monsempron à l’ouest de Fumel, mentionné comme limite). C’est un village absorbé par la commune de Fumel. Comme c’est un quartier de Fumel à présent, ne pas l’annoter pour l’instant. Sinon ce serait le même identifiant que Fumel.

Pour Villeneuve, ok on laisse comme cela dans le dictionnaire. Pour l’annotation automatique, cela supposera de préciser la désambiguïsation contextuelle.

Pour Banque de France : d’accord, je note qu’il faudra que l’on harmonise ensuite avec les autres cas. Pour l’instant dans le guide de transcription, j’avais ajouté l’annotation des lieux comme *Radio <lieu id=xxx> Londres </lieu>* mais on pourra les supprimer ensuite.

Lieux :

<DL> 321 --- "de Pampelune"

<LJ> oui, c’est ambigu ?

<DL> Non je l’ai ajouté : il n’y avait rien de transcris, alors que j’entendais “de Pampelune”.

<DL> 323 “<pron=lieu>” --- “el Fuerte de Rapitán”

<LJ> Je dirai “el Fuerte de Jaizkibel”, Jaizkibel est la montagne à l’ouest de Fontarrabie, sur lequel est le Fort de Guadalupe. Comme c’est une montagne, je l’ajoute au dictionnaire des Npr de lieux : Jaizkibel:5-id=3120203 pour annotation. Note pour plus tard : penser à l’ajouter aux cartes.

<DL> 1969 "le désert" --- "les Deux-Sèvres" : s'il parle bien des Deux-Sèvres, "Mosset" (ligne précédente) n'est peut-être pas la bonne ville.

<LJ> Oui, cela ne doit pas être les Deux-Sévres. Je mettrais <pron=pi> (pas sûr que ce soit un lieu non plus).

<DL> 2008, 2103 "Surac" ? Remplacer par <pron=pi> ?

<LJ> C’est peut-être Siorac, juste au nord de Belvès. La prononciation me semble coller.

Source Wikipedia : “souvent nommée « Siorac-de-Belvès »”. Commune de Siorac, puis Siorac-et-Fongauffier, et à présent Siorac-en-Périgord. Ajout dans le dictionnaire de : Siorac:5-id=2974480

Méta-données :

<DL> Alignement décalé

<DL> Prononciation hispanisante (pas systématiquement mais souvent) :

le [le]

v [b]

j [j]