### Version définitive des rubriques des métadonnées remplies de FS15

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

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

L’alignement a été effectué avec un séquençage d’environ 20 secondes.

Pour certaines séquences, en particulier celles du début qui sont particulièrement difficiles à transcrire, cette durée pourrait être raccourcie afin de faciliter la réécoute.

Il peut rester dans la marge des indications des débuts des fichiers audio séparés initiaux, par exemple [[01\_FS15]] [[02\_FS15]]. A SUPPRIMER A PRIORI

Informations générales sur le mode de transcription

Corrections automatiques de l’orthographe des mots “connait”, “parait” (qui correspond à l’orthographe réformée) en *connaît*, *paraît*.

Pour TLFI, l’orthographe “complétement” est parfois signalée comme fautive et parfois non ; toujours d’après le TLFI, cette forme est plus ancienne que les questions de réforme de l’orthographe. L'orthographe choisie dans les transcriptions est donc *complètement*. A VÉRIFIER QUE FAIT SINON A FAIRE

Contexte de l’enregistrement

Jeu de mots en clin d'œil au nom de l'interviewé : *<pron=pi> je mange du serrano* (serrano est par ailleurs dans le lexique de MATRICIEL).

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

Des emprunts fréquents à l’espagnol, non balisés :

* FS15 utilise la construction du passé composé espagnol, c’est-à-dire l’auxiliaire avoir au présent (he, has, ha , hemos, habeis, han) + participe passé, et le transpose en français, par exemple : *je suis <espagnol> he </espagnol> né* (où dans ce cas, la balise <espagnol> he </espagnol> n’est pas suivie d’une balise de traduction).
* *yo* (traduction “je”). Comme il n’y a pas de risque de confusion avec un mot français, une transformation avec une expression régulière a été effectuée : *yo* transformé en *<espagnol> yo </espagnol> <trad> je <trad>*. A FAIRE ? (avant vérifier que balises bien supprimées : ex. *<espagnol> yo </espagnol> <trad> j' </trad> ai grandi* → *yo ai grandi*)
* *no* (traduction “non”). Même remarque que pour *yo*. A FAIRE ? (si l’emprunt est bien systématique)
* *y* (traduction "et"). Même remarque que pour *yo*.
* *une <pron=[una]> baraque*. “una” est un mot espagnol et donc à ce titre il faudrait le transcrire comme un emprunt systématique. Même remarque que pour *yo* : *una pourrait être* transformé en *<espagnol> una</espagnol> <trad> une <trad>*. A FAIRE ??

Prononciations hispanisantes, non balisées :

* *un*, *de* [de] (traduction "de"), sont prononcés à l’espagnole.

Emprunts à l’espagnol et/ou prononciations hispanisantes :

* *le* : il est prononcé [lé] mais correspond ici soit à un emprunt soit à une prononciation hispanisante. Cela peut être difficile à déterminer, mais peut être possible en utilisant le contexte : par exemple, FS15 prononce "les Sénégalais", d’après le contexte la plupart du temps, il semble qu’il n'y a qu'un seul Sénégalais et donc il s’agirait d’une prononciation hispanisante.

Pour un emprunt, en espagnol “le” est le pronom personnel masculin, 3ème pers. du singulier, utilisé pour une personne ou un animal (ce sera “lo” pour un inanimé). Le déterminant défini masculin singulier est “el”, donc selon le cas il faudra le transcrire comme : “<espagnol> le </espagnol> <trad> lui <trad>” ou “ le <pron=[lé]>”.

Pour une prononciation hispanisante, *le* correspond au "le" français prononcé à l’espagnol.

A VERIFIER : si tous les “le” sont prononcés “lé” il faut l’indiquer dans les métadonnées.

La balise <pron=[API]> est utilisée à quelques reprises mais pas systématiquement.

INSÉRER LES COMMENTAIRES DE CARMEN sur la traduction espagnole et prononciations.

Incertitudes dans la transcription mais une interprétation est choisie

Cet entretien a été difficile à transcrire au vu de l'utilisation de mots espagnols et de la prononciation. Beaucoup de passages sont en effet en espagnol ou avec un fort accent espagnol. Les transcriptions, sauf deux, ont été effectuées par des personnes non hispanophones : il y a donc eu moins d’écoutes averties de ces passages. Des désaccords entre transcripteurs sont apparus mais des choix ont été faits en s’appuyant sur les écoutes d’hispanophones.

Interprétation des Npr de lieux et de personnes

Pour les noms de lieux :

Pour les noms de personnes :

* Le code Npr NP135 n’apparaît pas dans la transcription. A SUPPRIMER DE LA LISTE POUR CE RÉCIT?

### Commentaires Déborah puis Carmen, Catherine, Laurence

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

Commentaires Déborah/réponses, et Carmen, Catherine, Laurence

Notes Déborah

===========

FS15

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

Ajout d'une \* à chaque tier

Alignement avec séquençage de 20 secondes environ

<CD> Pour certaines séquences, en particulier celles du début qui sont particulièrement difficiles à transcrire, il me semble que leur durée est trop longue pour faciliter la réécoute.

Balisage des NprLieu avec les identifiants de gazetier : ok

Vérification des codes de noms propres de personnes : ok

<CD> Autre emprunt systématique -- il me semble et donc à ce titre à noter dans les métadonnées de l’entretien : FS15 utilise la construction du passé composé espagnol, i.e. auxiliaire avoir au présent (he, has, ha , hemos, habeis, han) + participe passé, et le transpose en français, ce qui donnerait ligne 45 par exemple : “je suis <espagnol> he </espagnol> né”. Dans ce cas, la balise <espagnol> he </espagnol> ne serait pas suivie d’une balise de traduction.

<CD> Pour être plus précis, et donc faciliter la transcription, il faudrait faire la différence entre les emprunts comme “yo” et les prononciations hispanisantes comme “un” et “le” prononcés à l’espagnole.

<CD> Pour les emprunts systématiques comme “yo”, après avoir signalé le phénomène dans les métadonnées de l’entretien, on peut le transcrire tel quel et, comme il n’y a pas de risque de confusion avec un mot français, on fera une transformation avec une expression régulière “yo” -> “<espagnol> yo </espagnol> <trad> je <trad>” ;

<CD> même remarque pour “no” si l’emprunt est systématique.

<CD> Pour “le” prononcé “lé”, ça peut être une question de prononciation ou d’emprunt. “le” est le pronom personnel masculin, 3ème pers. du singulier, utilisé pour une personne ou un animal (ce sera “lo” pour un inanimé). Le déterminant défini masculin singulier est “el”, donc selon le cas il faudra le transcrire comme :

“<espagnol> le </espagnol> <trad> lui <trad>” ou “ le <pron=[lé]>”

D’autre part, si tous les “le” sont prononcés “lé” il faut l’indiquer dans les métadonnées.

Prononciations récurrentes non balisées :

yo (trad "je")

de [de] (trad "de")

y (trad "et")

no (trad "non")

le (prononcé [lé] mais correspond parfois à "le")

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

Corrections effectuées :

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

Corrections systématiques :

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

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

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

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

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

<mdd> mais </mdd> --- mais

<mdd> oui </mdd> --- oui

rire / [rireentendu] --- <rireentendu>

<silence> --- <silenceentendu>

<LJ> ok, en fonction et si c'est ambigu, indiquer une balise <si> --- </si> (signal interrompu)

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

XXX là --- XXX-là

<audiomanquantsep> </audiomanquantsep> <audiomanquantfus> </audiomanquantfus> --- effacés

Ajouts de majuscules, par exemple :

les républicains --- les Républicains

les espagnols --- les Espagnols

Autres corrections :

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

Questions :

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

[[01\_FS15]] [[02\_FS15]] ... (dans la "marge") A quoi cela correspond-il ? Est-ce que les efface ?

<LJ> ce sont les indications des débuts des fichiers audio séparés. À conserver et on ne supprimera qu'à la fin ? <CD>oui

70-72 <repetition> <espagnol> yo </espagnol> <trad> j' </trad> ai grandi </repetition> <espagnol> yo </espagnol> <trad> j' </trad> ai grandi <espagnol> yo </espagnol> <trad> j' </trad> me mis

<repetition> yo ai grandi </repetition> yo ai grandi yo me mis

>> je supprime donc tous les balisages de "yo" dans la transcription ?

<LJ> je pense que l'on conserve les balises <espagnol> et <trad> dans ce cas-là. Ou alors c'est proche de "je" français en prononciation ? Dans d'autres cas, indiquer uniquement <espagnol> et ne pas faire la traduction qui sera faite ensuite.

J'ai corrigé les "connait/parait/complètement" en "connaît/paraît/complétement". Mais je précise au cas où, que si on applique l'orthographe réformée, dans ce cas, il faut laisser connait/parait/complètement".

<LJ> "complétement" n'est pas une orthographe plus ancienne que "connaît/paraît" ? De mon côté, je laisserais "connaît/paraît" ; pour "complétement" ou "complètement" : "complètement" sauf si cela est incohérent avec le fait d'appliquer l'orthographe non réformée (selon la date des réformes de l'orthographe ?)

<CD> en regardant le TLFI l’orthographe “complétement” est parfois signalée comme fautive et parfois non ; et toujours d’après le TLFI, cette forme est plus ancienne que les questions de réforme de l’orthographe, je mettrais donc systématiquement “complètement”.

558 <pron=pi> je mange du serrano

Ce tour n'était pas du tout transcrit car <pron=pi>, mais j'ai pu transcrire la fin.

C'est une blague en clin d'oeil au nom de l'interviewé. Est-ce que je laisse tel quel ?

<LJ> oui je laisserais, c'est vrai qu'on l'entend bien

<CD> signaler le jeu de mots dans les métadonnées ; “serrano” rajouté dans le lexique MATRICIEL

340 une <pron=[una]> baraque

384 non <pron=[no]>

>> la balise <pron=[API]> est utilisée à quelques reprises mais vraiment pas systématiquement.

<LJ> Ok pour API, mais il ne faut pas que ce soit trop long à prendre en compte. Si cela revient souvent pour la même personne, remonter cette prononciation non standard aux commentaires généraux par entretien.

<CD> “una” est un mot espagnol et donc à ce titre il faudrait le transcrire comme un emprunt systématique, i.e. l’indiquer dans les métadonnées de l’entretien, et comme il n’y a pas de risque de confusion avec un mot français le transcrire tel quel “una” et le transformer ensuite avec une expression régulière en “<espagnol> una</espagnol> <trad> une <trad>”

761-800 De par l'accent espagnol de FS15, on entend "les Sénégalais", mais je crois comprendre que, la plupart du temps, il n'y a qu'un seul Sénégalais. J'ai donc mis beaucoup de pronoms et de verbes au singulier. Qu'en pensez-vous ?

<LJ> oui, effectivement, au singulier. Indiquer en commentaires généraux de cet entretien cette

prononciation non standard.

<CD> voir la remarque en début de fichier

Gazetier : selon ce que Laurence m'expliquait, il me faudra donc la version non temporaire du gazetier ?

Egalement, j'ai relevé "Fos" ligne 698, que je n'ai pas retrouvé dans le gazetier temporaire.

<LJ> j'ai mis sur le FTP les nouvelles versions. Il se peut que des noms de lieux ne soient pas présents dans le gazetier alors qu'ils le sont dans la transcription mais cela peut être dû à une erreur dans la transcription.

<DL> Je n’avais pas encore appliqué le dictionnaire sur cette transcription. C’est chose faite, donc :

- A supprimer ?

Landes

los Pirineos

Italie

Espana

Espaga

El Mediterraneo

- A remplacer :

Cordoba --- Córdoba

- A ajouter :

Fos

<LJ> Dictionnaire mis à jour sur le FTP le 18/11

Ajouts :

Fos:5-id=3017679

Modifications :

Cordoba --- Córdoba

Suppressions :

Landes:5-id=DEPARTEM0000000000000012

los Pirineos:5-id=3112616

Italie:5-id=3175395

Espana:5-id=2510769

Espaga:5-id=2510769

El Mediterraneo:5-id=363196

<DL> Code Npr : NP135 n’apparaît pas dans la transcription

Concernant la transcription dans sa globalité, effectivement il y a beaucoup de passages en espagnol et de passages avec un fort accent espagnol. Je ne pense pas que je puisse vraiment faire mieux. J'aurais parfois transcrits autrement, mais je n'ai pas une très bonne oreille pour l'accent espagnol. Je suppose qu'il vaudrait mieux laisser un traducteur hispanophone se charger, non seulement des traductions, mais aussi des passages avec un fort accent espagnol. Qu'en pensez-vous ?

<LJ> Oui, juste indiquer entre les balises <espagnol> </espagnol> sans la traduction.

<CD> oui, cet entretien est particulièrement difficile et je ne suis pas toujours d’accord avec votre transcription, mais vu les difficultés de cet entretien, j’ai plusieurs propositions :

* constituer une catégorie d’entretiens “difficiles” en justifiant chaque fois dans les métadonnées de l’entretien ;
* pour ces entretiens, segmenter en séquences plus courtes pour faciliter la réécoute.(cf. remarque en début de fichier) ;
* programmer si on a le temps une séance de travail commune pour réécouter ensemble et transcrire de manière plus efficace les passages difficiles.