

Macros & balisage

Annexe au cahier des charges pour la transcription d'entretiens audios à visée terminologique

Contact :

Stéphane Riou, ingénieur de recherche - termino-icima.institut@marionnette.com - 03 24 33 72 50

Version actualisée le 26 février 2019

1. Balises communes

1.1. Rire

Phénomène : rire

Macro TRJS : Rire

Encodage TEI : `<shift feature="voice" new="laugh"/> <shift feature="voice" new="normal"/>`

Description : Si des mots sont prononcés en riant, mettre les termes à l'intérieur de la balise, sinon laisser vide.

1.2. Soupir

Phénomène : soupir

Macro TRJS : Soupir

Encodage TEI : `<shift feature="voice" new="sigh"/> <shift feature="voice" new="normal"/>`

Description : Si des mots sont prononcés en soupirant, mettre les termes à l'intérieur de la balise, sinon laisser vide.

1.3. Incertain

Phénomène : Incertain (doute)

Macro TRJS : Incertain

Encodage TEI : `<unclear>(transcription incertaine)</unclear>`

Description : mettre un terme par balise `<unclear>`. Si plusieurs possibilités, les séparer par un pipe | dans la balise.

2. Balises entités nommées

2.1. Lieu

Phénomène : Lieu

Macro TRJS : EN_Lieu

Encodage TEI : `<placeName>(nom de lieu)</placeName>`

Description : nom d'un lieu : ville, salle de spectacle, pays...

2.2. Date

Attention : deux informations à compléter : l'attribut *@when* et la date.

2.2.1. Date entière

Phénomène : Date entière

Macro TRJS : Date

Encodage TEI : `<date when="(JJ-MM-AAAA)">(la date)</date>`

Description : `<date when="24-12-2017">24 décembre de cette année</date>`

2.2.2. Date mois

Phénomène : Date avec mois et année

Macro TRJS : Date_Mois

Encodage TEI : `<date when="(MM-AAAA)">(la date)</date>`

Description : `<date when="01-2019">janvier 2019</date>`

2.2.3. Date année

Phénomène : Seulement l'année

Macro TRJS : Date_Année

Encodage TEI : <date when="(AAAA)">(la date)</date>

Description : <date when="1968">68</date>

2.3. Identité

Phénomène : Identités

Macro TRJS : EN_Identite

Encodage TEI : <name>(nom de la personne physique, fictive ou morale)</name>

Description : nom, prénom, surnom de personnes, de personnages, d'organisation...

2.4. Spectacle

Phénomène : Titre et nom de spectacles

Macro TRJS : EN_Spectacle

Encodage TEI : <name type="work">(nom du spectacle)</name>

Description : titre d'un spectacle, d'une oeuvre...

3. Balise terme métier

Phénomène : Terme métier (*ne pas confondre avec le tier transcription fine*)

Macro TRJS : Terme_Metier

Encodage TEI : <term>(le terme ou le groupe de mots)</term>

Description : <term>marionnette à tige fixe</term>

4. Balises pauses

Les pauses courtes et la pause longue s'appliquent à tout type de transcription.

4.1. Pause longue

Phénomène : pause chronométrée (transcription fine)

Macro TRJS : pause_longue

Encodage TRJS: ###

Encodage TEI : ((mécanisme intégré à TRJS))

Description : ### ((TRJS convertit directement en TEI)). Pour une pause de plus de 2 secondes

4.2. Pause courte

Phénomène : pause courte (tout type de transcription)

Macro TRJS : pause_courte

Encodage TRJS : #

Encodage TEI : ((mécanisme intégré à TRJS))

Description : # ((TRJS convertit directement en TEI))

5. Balises conversationnelles

5.1. Transcription phonétique

Phénomène : Transcription phonétique

Macro TRJS : API

Encodage TEI : <seg type="API">(le texte en API)</seg>

Description <seg type="API">R0b0R</seg>

5.2. Auto-correction

Phénomène : correction, reprise de son propre terme

Macro TRJS : Correction

Encodage TEI : <choice><sic>(correction)-</sic><corr>(terme visé)</corr></choice>

Description des <choice><sic>pou</sic><corr>poupées</corr></choice> des marionnettes

S'il est impossible de trouver le terme visé, recopier à l'identique le contenu de la balise <sic> dans <corr>

5.3. Geste

Phénomène : Geste

Macro TRJS : Geste

Encodage TEI : <kinesic> <desc></desc> </kinesic>

Description : <kinesic> <desc>prise entre le pouce et l'index</desc> </kinesic>

5.4. Indications contextuelles

Phénomène : Indications contextuelles

Macro TRJS : Contexte

Encodage TEI : <incident type="comment"><desc></desc></incident>

Description : <incident type="comment"><desc>il y a un changement de place des intervenants</desc></incident>

6. Passage transcription fine

Les passages en transcription fine (voir le cahier des charges en 3.1.1) sont marqués en sélectionnant par un clic-droit sur “Locuteur” et en choisissant “fine” dans la catégorie Template. Ces passages, étant sur un tier différent, se superposent aux transcriptions des locuteurs.

