ELAN > Toolbox > ELAN

ELAN sert à définir les frontières des segments Audio à transcrire, et à les préparer pour l'annotation. Une ligne *ref* est créée pour référencer (numéroter) chaque segment, une ligne *tx* pour la transciption, une ligne *mot* dans laquelle chaque mot est isolé dans une cellule et une ligne *ft* pour la traduction libre. Ce fichier ELAN est exporté dans le format Toolbox.

Toolbox est utilisé pour annoter le fichier issu de ELAN. Les lignes de segmentation *mb* et d'annotation *ps* et *ge* seront créées interactivement grâce à un dictionnaire. Ce fichier Toolbox sera importé en retour dans ELAN afin de pouvoir être exploité en mode recherche.

Pour assurer une homogénéité des configurations au sein de l'équipe, les fichiers **refSenel.typ**, **mdf.typ** seront préalablement copiés dans le sous-dossier *Settings* du sous-dossier *Toolbox* du dossier *Senelangues* ainsi que le fichier **Senel1.etf** dans le sous-dossier ELAN.

SENELANGUES
- AUDIO
  - Fichiers.wav…
- ELAN
  - **Senel1.etf**
  - Fichiers ELAN…
- TOOLBOX
  - Dictionnaire
  - Fichiers annotés Toolbox…
  - SETTINGS
    - **Mdf.typ**
    - **refSenel.typ**

Création d'un fichier **ELAN**

Le fichier ELAN contiendra les informations de segmentation et d'annotation du fichier Audio ou Vidéo. Il faut donc commencer par charger ce fichier *media*.

Créer un nouveau Document ELAN

- FICHIER, NOUVEAU
- *Fichier du type* : **Media files**, Sélectionner le fichier audio (**.wav**)
- Cliquer sur le bouton >> entre les 2 fenêtres, OK

Enregistrer le fichier ELAN

- FICHIER, ENREGISTRER SOUS : nom_du_fichier, ENREGISTRER

  le nom du fichier sera du type **codelangue_InitialesChercheur_type_num**
  type = conv(ersation) ou narr(ation) ; num = numéro du fichier dans la série

Importer le modèle des types linguistiques

- TYPE, IMPORTER TYPES, RECHERCHER, Rechercher *Senel1.etf*
- IMPORTER, FERMER

Supprimer le rang (acteur) *Default*

- CLIC-DROIT sur l'étiquette *Default*
- SUPPRIMER *Default*
- OK

Importer les Rangs (Acteurs) *ref, tx, mot* et *ft*

- ACTEUR**,** IMPORTER ACTEURS
- CLIQUER sur le bouton RECHERCHER et sélectionner le fichier *Senel1.etf*
- CLIQUER sur SELECT, puis IMPORTER, puis FERMER

Segmentation du fichier *Audio*

Deux méthodes: 1) la segmentation au kilomètre qui permet en une fois de définir les segments voulus. La frontière de ces segments pourra être corrigée par la suite. 2) la segmentation progressive avec annotation au fur et à mesure.

1 - Segmentation au kilomètre

- ACTEUR, SEGMENTATION
- SELECTIONNER le rang de segmentation ***ref*** en face de *Selection de l'acteur*
- SELECTIONNER '*Une frappe par annotation*' (chaque fois qu'on tapera la touche <Entrée>, il y aura une frontière de segment)
- CLIQUER sur le bouton LIRE LA SELECTION
- TAPER la touche <ENTREE> à la fin de chaque unité à segmenter
- CLIQUER sur le bouton PAUSE pour arrêter la segmentation
- CLIQUER sur le bouton APPLIQUER pour appliquer la segmentation

Les segments sont délimités dans le rang ***ref***. La segmentation peut être interrompue et reprise ultérieurement par ce même processus.

2 - Segmentation au fur et à mesure

- COCHER la case MODE DE SELECTION
- SELECTIONNER en balayant avec la souris le segment à écouter (la zone apparaît plus sombre) et l'écouter avec le bouton LIRE LA SELECTION
  - Pour repositionner la frontière de droite, MAINTENIR APPUYEE la touche *Majuscule* et CLIQUER à l'endroit voulu
- Une fois la zone à transcrire correctement définie, faire un CLIC-DROIT dans cette zone, **au niveau du rang *ref***, et sélectionner NOUVELLE ANNOTATION ICI. Une zone de texte s'ouvre. La fermer avec CTRL/ENTER. (le contenu du rang ref sera généré automatiquement, voir plus bas)
- Pour saisir la transcription, DOUBLE-CLIQUER sur le rang ***tx*** en dessous du segment ***ref*** correspondant. Fermer la fenêtre de saisie par CTRL/ENTER.
- Pour saisir la traduction libre, DOUBLE-CLIQUER sur le rang ***ft*** en dessous du segment ***ref*** correspondant. Fermer la fenêtre de saisie par CTRL/ENTER.

Délimiter un nouveau segment à la suite du précédent

- VERIFIER que la case MODE DE SELECTION est COCHEE
- CLIQUER sur le bouton , le fil rouge se déplace en fin de segment
- CLIQUER sur le bouton pour désélectionner le segment précédent
- CLIQUER à la fin du nouveau segment voulu
- CLIQUER sur le bouton pour écouter le segment sélectionné

- Pour étendre cette selection si nécessaire, MAINTENIR APPUYEE la touche *Majuscule* et cliquer plus à droite de la sélection (ou plus au gauche pour la réduire)

- Une fois le segment à transcrire correctement défini, faire un CLIC-DROIT dans cette zone, **au niveau du rang *ref***, et sélectionner NOUVELLE ANNOTATION ICI.

Déplacer la frontière entre deux segments

- COCHER la case MODE DE SELECTION
- Pour déplacer une frontière vers la droite, Sélectionner le segment de gauche en CLIQUANT à l'intérieur du segment, au niveau du rang *ref*. (Pour déplacer une frontière vers la gauche, Sélectionner le segment de droite.)
- Maintenir appuyée la touche ALT et amener le curseur sur la ligne de partage entre les deux segments, au niveau du rang *ref*.
- Cliquer (le curseur se transforme en une double-flèche) et déplacer la frontière vers la droite (ou la gauche), puis relâcher.

Annotation de segments prédéfinis

Si vous n'avez pas saisi les transcriptions au fur et à mesure

- DOUBLE-CLIQUER dans le segment à annoter, au niveau du rang ***tx*** ou ***ft*** . Une zone de texte s'ouvre
- TAPER le texte voulu (en vernaculaire dans *tx*, en français dans *ft*) et terminer par CTRL/ENTER.

Labeliser le rang *ref* (référencer les segments)

Il s'agit de numéroter automatiquement les segments, avec une étiquette (label)

- ACTEUR, LABELISER ET NUMEROTER…, Sélectionner **ref**
  - *Inclure la partie label* : **codelangue_InitialesChercheur_type_num**
    type = *conv*(ersation) ou *narr*(ation*) ;* num = *numéro du fichier dans la série*
  - *Insérer autre délimiteur* : TAPER _ (caractère souligné)
  - OK, FERMER

Ranger les rangs dans l'ordre voulu

- Cliquer sur l'étiquette du rang à déplacer et maintenir appuyé le bouton de la souris
- Déplacer l'étiquette à l'emplacement voulu, puis relâcher

Ranger les rangs dans l'ordre hiérarchique

- Cliquer avec le bouton de droite dans la zone des étiquettes des tiers
- SORT TIERS, SORT BY HIERARCHY

Tokeniser le rang *tx* dans le rang *mot* (une cellule par mot)

Il s'agit d'isoler chaque mot de la ligne *tx* dans une cellule individuelle

- ACTEUR, TOKENISER L'ACTEUR
  - *Source* : **tx**
  - *Destination* : **mot**
  - COMMENCER, FERMER

Exportation vers Toobox

- FICHIER, EXPORTER VERS : **Fichier Shoebox**
  - *Utiliser le type de base de données Shoebox* : Cliquer sur [...], Rechercher **refSenel.typ** dans le sous-dossier *Settings* du dossier Toolbox
  - Cocher *Encoder les marqueurs en Unicode*, OK
- Rechercher le dossier destinataire, donner un nom au fichier Toolbox, ENREGISTRER

TOOLBOX

Ouvrir le fichier issu de l'exportation ELAN

- FICHIER, OUVRIR
- Rechercher le dossier voulu, Sélectionner le fichier issu de l'exportation de ELAN
- OUVRIR
- Lancer l'interalignement (voir document *TOOLBOXELAN.PDF*)

RETOUR DANS ELAN

Ouvrir le fichier annoté avec TOOLBOX dans ELAN

- Fermer le projet TOOLBOX, Ouvrir ELAN
- FICHIER, IMPORTER, FICHIER SHOEBOX
- Sélectionner le fichier Toolbox en cliquant sur les 3 petits points [...]
- Cocher TOUS LES MARQUEURS SONT UNICODE
- OK
- A la demande, Rechercher le fichier Wav correspondant au texte annoté
- FICHIER, ENREGISTRER SOUS : nom_du_fichier_ELAN, ENREGISTRER