

# Guide pratique pour aligner une transcription sous ELAN

Version 2.0 – Juin 2009

Isabel Colón de Carvajal

Laboratoire ICAR

(CNRS & Université de Lyon)

<b>1</b>	<b>Introduction générale .....</b>	<b>1</b>
<b>2</b>	<b>Installation du logiciel.....</b>	<b>1</b>
<b>3</b>	<b>Présentation de l'interface d'ELAN.....</b>	<b>2</b>
<b>4</b>	<b>Création d'un nouveau fichier ELAN vierge.....</b>	<b>3</b>
<b>5</b>	<b>Mise en place de la partition de travail.....</b>	<b>4</b>
<b>6</b>	<b>Création et gestion des annotations.....</b>	<b>5</b>
6.1	<i>Description des barres de navigation et de player vidéo .....</i>	<i>5</i>
6.2	<i>Insertion d'une annotation .....</i>	<i>6</i>
6.3	<i>Modification du contenu d'une annotation .....</i>	<i>7</i>
6.4	<i>Augmentation ou diminution temporelle d'une annotation.....</i>	<i>7</i>
6.5	<i>Suppression d'une annotation .....</i>	<i>8</i>
6.6	<i>Déplacement d'une annotation .....</i>	<i>8</i>
<b>7</b>	<b>Enregistrement.....</b>	<b>9</b>
<b>8</b>	<b>Fonction Import &amp; Export .....</b>	<b>9</b>
8.1	<i>Importation .....</i>	<i>9</i>
8.2	<i>Exportation.....</i>	<i>10</i>
<b>9</b>	<b>Fonction « Multiscopes » pour aligner plusieurs sources vidéo .....</b>	<b>13</b>
9.1	<i>Solution 1 : Si la durée des clips sont strictement identiques .....</i>	<i>13</i>
9.2	<i>Solution 2 : Si la durée des clips n'est pas équivalente .....</i>	<i>14</i>

## 1 Introduction générale

ELAN signifie EUDICO Linguistic ANnotator.

ELAN est un instrument professionnel pour la création d'annotations complexes sur les ressources vidéo et audio. Un utilisateur peut ajouter un nombre illimité d'annotations aux sources audio et/ou vidéo. Une annotation peut être une phrase, un mot ou un gloss, un commentaire, une traduction ou une description de n'importe quelle caractéristique observée depuis la source. Les annotations peuvent être créées sur les couches multiples, appelées des tiers. Les tiers peuvent être hiérarchiquement interconnectés.

Sous ELAN, une annotation peut être alignée temporellement à la source audio ou vidéo ; ou il peut référer à d'autres annotations existantes. Le contenu écrit des annotations est toujours en Unicode et la transcription est conservée dans un format XML. ELAN fournit différentes vues des annotations, et chaque vue est reliée et synchronisée à la source concernée.

Il est possible d'associer jusqu'à quatre fichiers vidéo à un seul même document d'annotation. Chaque vidéo peut être intégrée dans la fenêtre du document principale ou être affichée dans sa propre fenêtre qui sera alors redimensionnable.

ELAN utilise le lecteur du fichier audio et/ou vidéo à un soft existant, comme Windows Media Player, QuickTime ou JMF (Java Media Framework). Par conséquent, une large variété de formats audio et vidéo est supportée par ELAN. Ce logiciel est écrit dans le langage de programmation de Java et les sources sont disponibles pour une utilisation non commerciale.

ELAN accepte en format d'entrée les *fichiers audio* .wav, les *fichiers vidéo* .mpg1, .mpg2, .mov, .avi, .wmv mais il est nécessaire que les vidéo aient été compressées avec un codec mpeg de préférence.

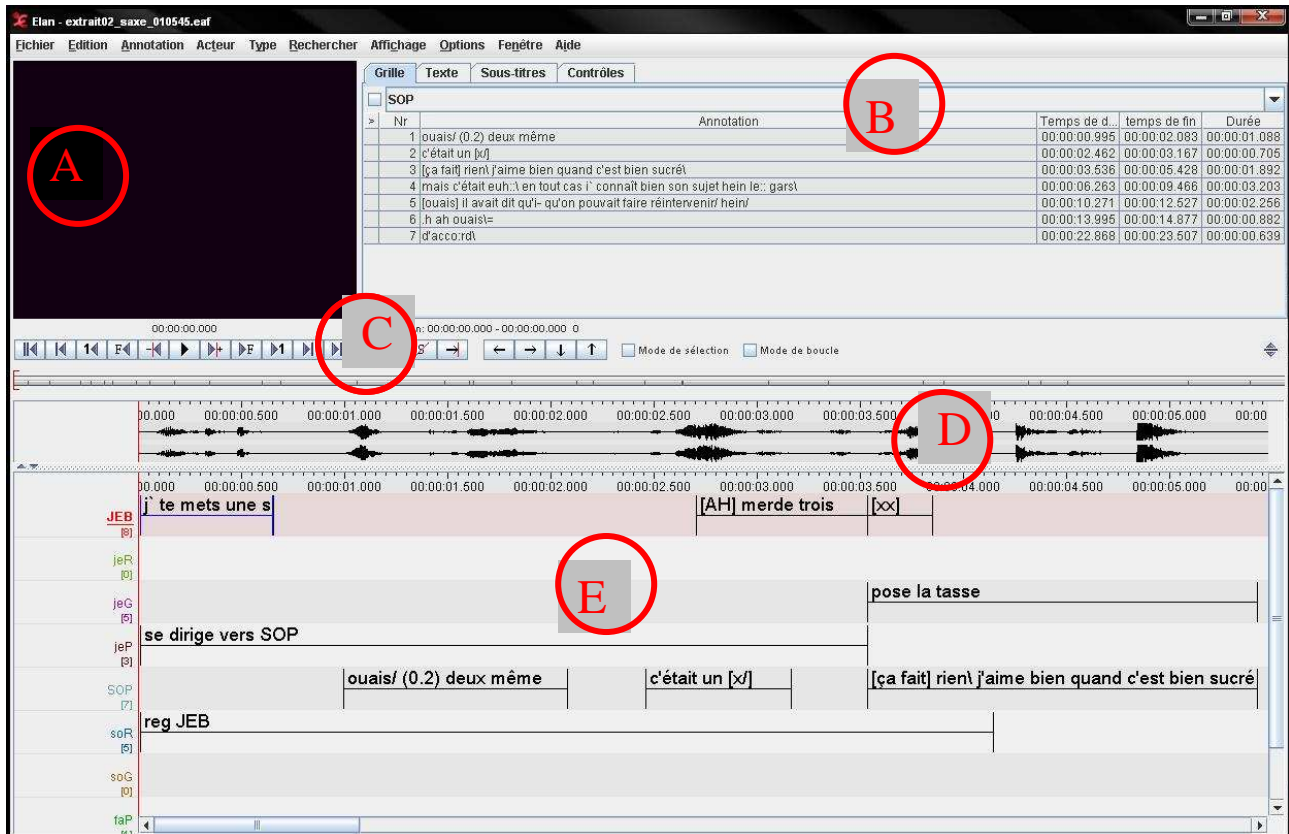
## 2 Installation du logiciel

Le logiciel est téléchargeable depuis ce site : <http://www.lat-mpi.eu/tools/elan/download>. On y trouve également le manuel complet en ligne ou consultable au format pdf.

Sous Windows, l'installation ne comporte aucun problème majeur, donc il suffit de lancer l'exécutable et de se laisser guider par les instructions.

Sous Mac, l'installation peut connaître quelques problèmes selon la version de java préalablement installée sur l'ordinateur. Il est conseillé d'installer la version d'ELAN compatible avec le même numéro de la version de Java.

### 3 Présentation de l'interface d'ELAN



Img. 1a : Fenêtre d'ELAN complète

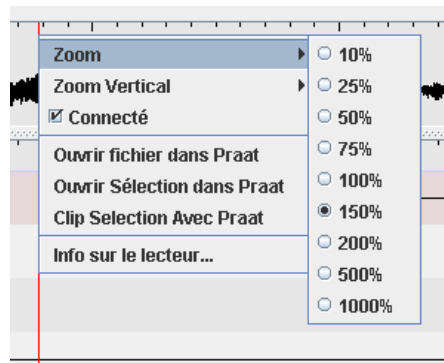
A = zone du lecteur vidéo.

B = zone où sont affichées les annotations sous différentes formes (avec le timing, sans le timing) puis l'onglet « **contrôles** » afin de jouer sur la vitesse du son et/ou son volume.

C = zone du contrôleur vidéo et/ou audio, ainsi que celui des sélections (cf. section 6.1).

D = zone de la wave form. Si l'on importe un fichier son avec son fichier vidéo, il est possible de voir l'oscillogramme de la wave form. En faisant apparaître le menu contextuel sur la wave form, il est possible de modifier le degré de zoom sur la partition de travail et bien sur, sur la wave form également (cf. Img. 1b ci après).

E = zone de la partition de travail pour les annotations futures.

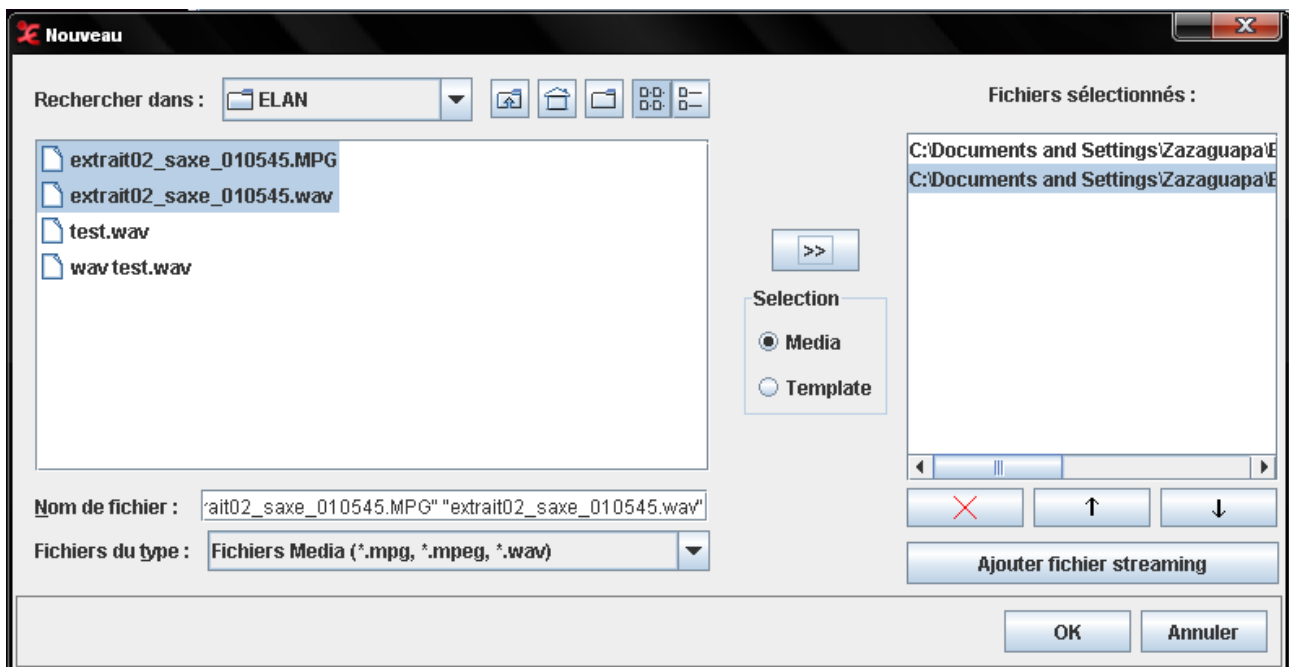


Img. 1b : Menu contextuel pour jouer sur le zoom de la wave form

## 4 Création d'un nouveau fichier ELAN vierge

Au préalable, il est recommandé de classer les fichiers vidéo/audio qui seront utilisés pour la même transcription dans un dossier unique.

- Lancer le logiciel
- Créer un nouveau fichier d'annotation par le menu **Fichier > Nouveau**. ELAN va d'abord demander de charger les fichiers vidéo/audio correspondant à la future transcription.
- Dans la partie gauche de la boîte de dialogue, mettre en évidence le contenu du dossier qui contient les données vidéo et/ou audio :



Img. 2 : Boîte de dialogue de sélection des fichiers sources

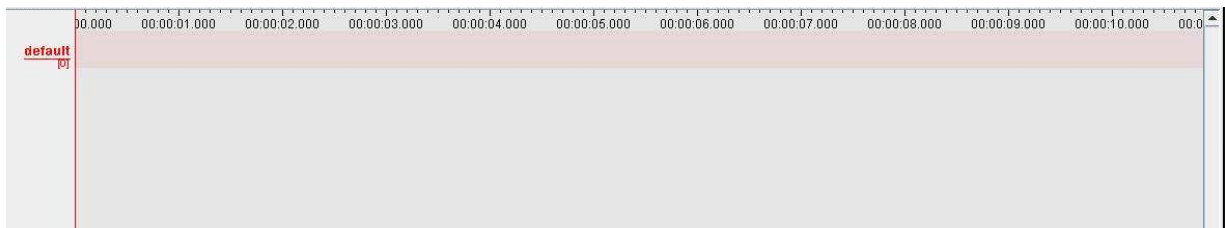
Faire basculer les fichiers souhaités dans le champ vide à droite « fichiers sélectionnés » grâce au bouton [ >> ].

NB : si l'on souhaite insérer des fichiers au format autre que .mpeg ou .wav, il est possible de modifier le type de fichiers à ajouter grâce **au menu déroulant** situé en bas à gauche de la boîte de dialogue.

d) Valider par **OK**

Votre fichier ELAN est créé, il faut passer à présent au travail proprement dit de transcriptions et d'annotations.

## 5 Mise en place de la partition de travail

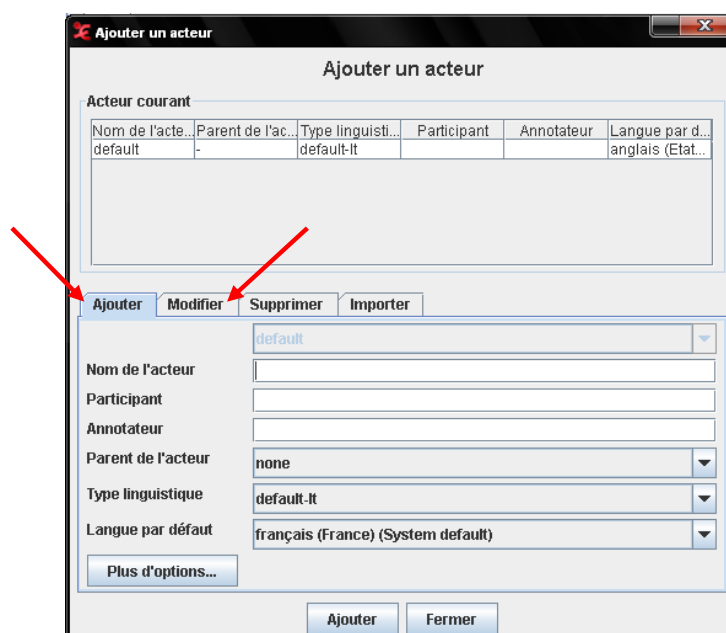


Img. 3 : Partition vierge pour la transcription

Afin d'insérer des annotations dans la transcription, il est conseillé de préparer la partition de travail en créant les lignes des participants. On appelle ces lignes des « **tiers** ». Par défaut lors de la création d'une transcription ELAN vierge, un tiers « default » est créé : soit on le supprime, soit on le modifie par un pseudo de participant pertinent pour l'extrait à aligner.

NB : dans ELAN, le participant est appelé **acteur**.

a) Par le menu **Acteur** > Ajouter un nouvel acteur (ou par le raccourci clavier **ctrl+T**).



Img. 4 : Boîte de dialogue nouvel acteur / modification acteur

Voici la description des différents champs à renseigner :

- Nom de l'acteur = Pseudo tel qu'il sera affiché au début de la ligne du tiers.
- Participant = Nom détaillé de l'acteur (informations supplémentaires non affichées dans la fenêtre active d'ELAN)
- Annotateur = champ facultatif à renseigner.

Ensuite, laisser par défaut les autres champs restants.

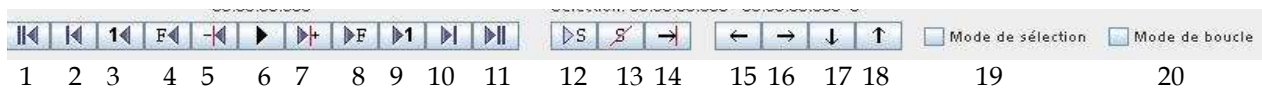
b) Valider en cliquant sur **Ajouter**.

*NB : la fenêtre d'ajout d'acteur ne se ferme pas automatiquement après l'ajout d'un acteur : il est possible d'ajouter d'autres acteurs, sinon, pour retourner à la fenêtre de travail, cliquer sur **Fermer**.*

Remarque : la fonction **Plus d'options** permet de modifier la couleur d'un tiers ou bien la police de ce même tiers.

## 6 Création et gestion des annotations

### 6.1 Description des barres de navigation et de player vidéo

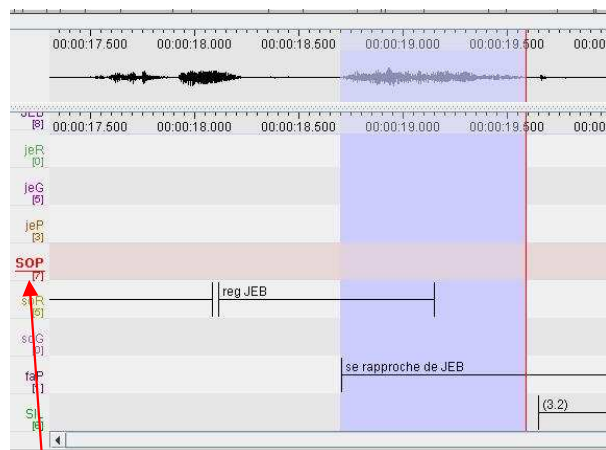


Description de chaque option en partant de la gauche :

- 1 - aller au départ du média
- 2 - aller au curseur précédent
- 3 - placer le temps 1 seconde en arrière
- 4 - aller à la frame précédente
- 5 - aller au pixel précédent
- 6 - lire / mettre en pause le média**
- 7 - aller au pixel suivant
- 8 - aller à la prochaine frame
- 9 - placer le temps 1 seconde en avant
- 10 - aller au prochain curseur
- 11 - aller à la fin du média
- 12 - lire la sélection**
- 13 - effacer la sélection
- 14 - placer le curseur à droite / gauche de la sélection

- 15 - aller à l'annotation précédente (horizontalement)
- 16 - aller à la prochaine annotation (horizontalement)
- 17 - annotation précédente (verticalement)
- 18 - annotation suivante (verticalement)
- 19 - mode sélection = lorsque la case est cochée, et que l'on lit le média, une sélection est automatiquement effectuée dans la partition de travail.
- 20 - mode boucle = lorsque la case est cochée, et qu'on lit une sélection, elle est jouée en boucle, en répétition.

## 6.2 Insertion d'une annotation



Img. 5 : Sélection d'un bloc temporel

- a) Vérifier que la ligne du tiers où l'annotation va être créée est bien sélectionnée. Le nom du participant doit être **surligné en rouge**.
- b) Sélectionner le bloc temporel correspondant au contenu textuel qui va être transcrit.
- c) Il est possible d'insérer l'annotation de deux manières différentes :
  - Par le menu **Annotation > Nouvelle annotation ici**
  - Par le menu contextuel → clic bouton droit (ou ctrl + clic sur mac) et choisir **Nouvelle annotation ici**
  - Par le raccourci clavier **Alt Gr + N**
- d) Taper le texte correspondant à l'annotation que l'on veut écrire ou décrire.
- e) Valider par **ctrl + Entrée** (ou **pomme + Entrée** sur mac)

*NB : si on valide par la touche Entrée uniquement, le contenu textuel ne sera pas sauvegardé et il apparaîtra uniquement une ligne dans le bloc temporel.*



Img. 6 : Résultat obtenu après validation de l'annotation


### 6.3 Modification du contenu d'une annotation

Lorsqu'une annotation a déjà été créée, et que l'on souhaite modifier son contenu textuel :

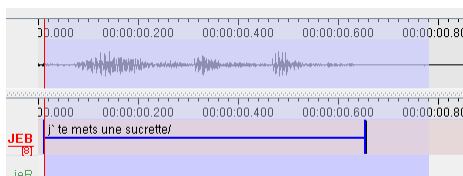
- soit on double-clique sur le bloc temporel correspondant afin de réactiver le mode « saisie de texte ». Il est important alors de bien valider les nouvelles corrections par la combinaison de touche **ctrl+Entrée** sur pc ou **pomme+Entrée** sur mac.
- soit on sélectionne l'annotation à modifier en cliquant UNE fois sur le bloc dont on souhaite modifier le contenu. La ligne horizontale du bloc doit alors être **surlignée en bleue** (cf. img. 6) ; et l'ensemble du bloc doit être bleuté ce qui marque sa sélection. Puis :
  - par le menu contextuel : **clic bouton droit** de la souris (sur l'annotation même) + **Modifier la valeur de l'annotation**.
  - ou par le menu **Annotation > Modifier la valeur de l'annotation**.
  - ou par le raccourci clavier **Alt+M** (ou **pomme+M** sur mac).

### 6.4 Augmentation ou diminution temporelle d'une annotation

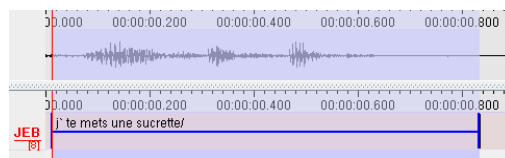
Lorsqu'une annotation a été créée, et que sa valeur temporelle a été déterminée, il est encore possible d'en modifier les frontières temporelles.

- a) Cliquer UNE fois sur le bloc dont on souhaite modifier la valeur temporelle. La ligne horizontale du bloc doit alors être **surlignée en bleue** (cf. img. 6) ; et l'ensemble du bloc doit être bleuté ce qui marque sa sélection.
- b) Vérifier que la ligne verticale rouge du curseur de la sélection soit placée à GAUCHE du bloc si l'on veut augmenter ou diminuer la valeur du temps de l'annotation. Si elle ne l'est pas, modifier sa position grâce au bouton située au dessus .
- c) Pour augmenter la valeur temporelle de l'annotation : maintenir la touche SHIFT du clavier (majuscule gauche) et cliquer à droite du bloc sélectionné.



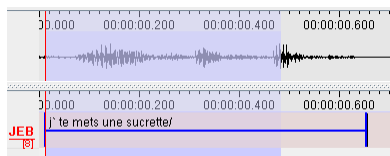


Img. 7a : sélection augmentée



Img. 7b : après validation

- d) Pour diminuer la valeur temporelle de l'annotation : maintenir la touche SHIFT du clavier (majuscule gauche) et cliquer à l'intérieur du bloc sélectionné.



Img. 7c : sélection diminuée



Img. 7d : après validation

NB : Cette manipulation afin d'augmenter ou diminuer un bloc de sélection est possible également sur une sélection vierge, c'est-à-dire vide de contenu d'une annotation. Il suffit d'insérer par la suite le contenu (cf. section 6.2).

### 6.5 Suppression d'une annotation

Pour supprimer une annotation qui a été créée, il faut d'abord la sélectionner, c'est-à-dire cliquer UNE fois sur le bloc dont on souhaite modifier la valeur temporelle. La ligne horizontale du bloc doit alors être **surlignée en bleue** (cf. Img. 6) ; et l'ensemble du bloc doit être bleuté ce qui marque sa sélection. Ensuite,

- soit on passe par le menu contextuel : **clic bouton droit** de la souris (sur l'annotation même) + **Effacer Annotation**.
- soit on passe par le menu **Annotation > Effacer Annotation**.
- Soit par le raccourci clavier **Alt+D** (ou **pomme+D** sur mac).

### 6.6 Déplacement d'une annotation

Pour déplacer une annotation dans un tiers différent, il faut la copier ou la dupliquer dans le tiers souhaité de la transcription.

Pour déplacer une annotation sur un même tiers, il faut sélectionner l'annotation, et en maintenant ensemble la touche du clavier **Alt**, faire glisser l'annotation vers la droite ou la gauche.

NB : la ligne horizontale devient alors **verte** lorsque le déplacement est possible.



Img. 8 : Déplacement d'une annotation sur un même tiers

## 7 Enregistrement

Le format natif d'ELAN est un fichier **.eaf**

- Lors du premier enregistrement sous ELAN, le logiciel va demander de donner un nom au fichier créé. Pour enregistrer le document pour la première fois, aller dans le menu **Fichier > Enregistrer sous**. Il faudra également spécifier l'endroit où le fichier va être stocké (de préférence dans le même dossier où sont stockées les sources vidéo et/ou audio).

- Lors des enregistrements suivants, soit on repasse par le menu **Fichier > Enregistrer**, ou bien on utilise le raccourci clavier **ctrl+S** (sur pc) ou **pomme+S** (sur mac).

*NB : il existe également une option d'enregistrement automatique où l'on peut spécifier la l'intervalle de répétition de l'enregistrement automatique. Afin de paramétrer cette option, il faut aller dans le menu **Fichier > Enregistrement automatique**, puis dans le menu déroulant qu'il propose, spécifier 1, 5, 10, 20... minutes.*

Une fois enregistré, on peut fermer le document et le ré-ouvrir ultérieurement, et cette fois ci, **en ouvrant uniquement le fichier .eaf**. ELAN ouvrira automatiquement les sources vidéo et/ou audio correspondantes. Si entre temps, elles ont changé d'emplacement dans l'ordinateur, ELAN demandera alors de spécifier de nouveau leurs emplacements.

## 8 Fonction Import & Export

### 8.1 Importation

ELAN donne la possibilité d'importer différents type de fichiers créés sous d'autres logiciels. Il est possible d'importer les formats suivant :

- Shoebox
- Toolbox
- **Fichier Textgrid (Praat)** : ELAN supporte des fichiers .Textgrid Praat encodes en UTF-8 et UTF-16.

- **Chat (Clan) (.cha uniquement)** : ELAN supporte les anciens fichiers CHAT et CHAT-UTF8, non les fichiers XML CHAT.
- Fichier Transcriber
- CSV (fichier texte délimité par des tabulations) (.txt ou .csv)

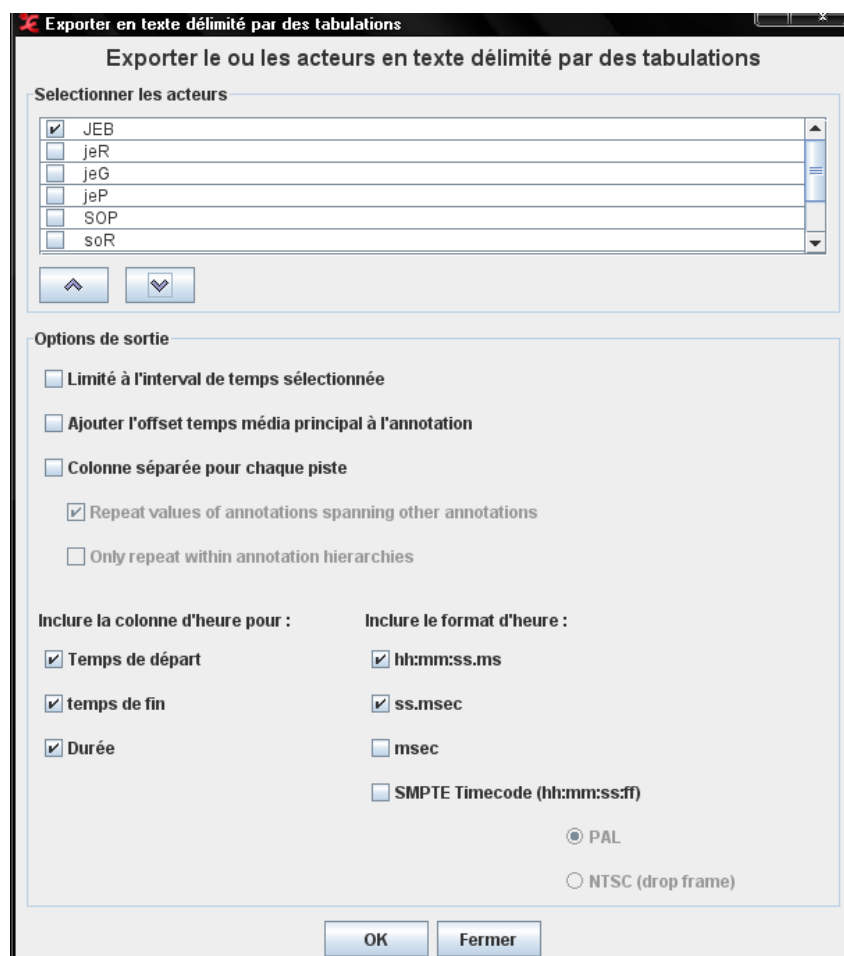
*NB : différents paramètres sont à renseigner en fonction du type d'import choisi.*

*Le résultat obtenu de l'importation n'est pas nécessairement fidèle aux informations présentes dans le fichier d'origine >> Il est donc souvent demandé de revoir l'alignement de la transcription une fois sous ELAN.*

## 8.2 Exportation

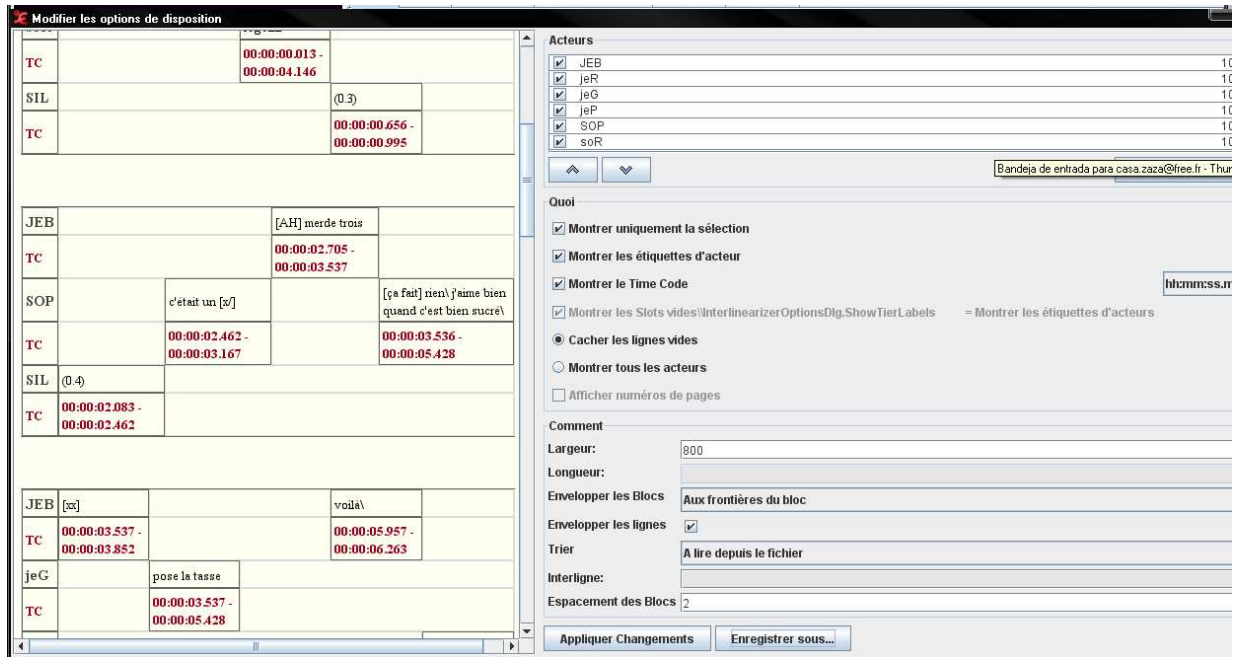
ELAN permet d'exporter sa transcription alignée sous différents formats de fichiers afin de les exploiter dans d'autres logiciels. Il est possible d'exporter dans les formats suivant :

- Fichier Shoebox
- Fichier Toolbox
- Fichier CHAT
- Texte délimité par des tabulations

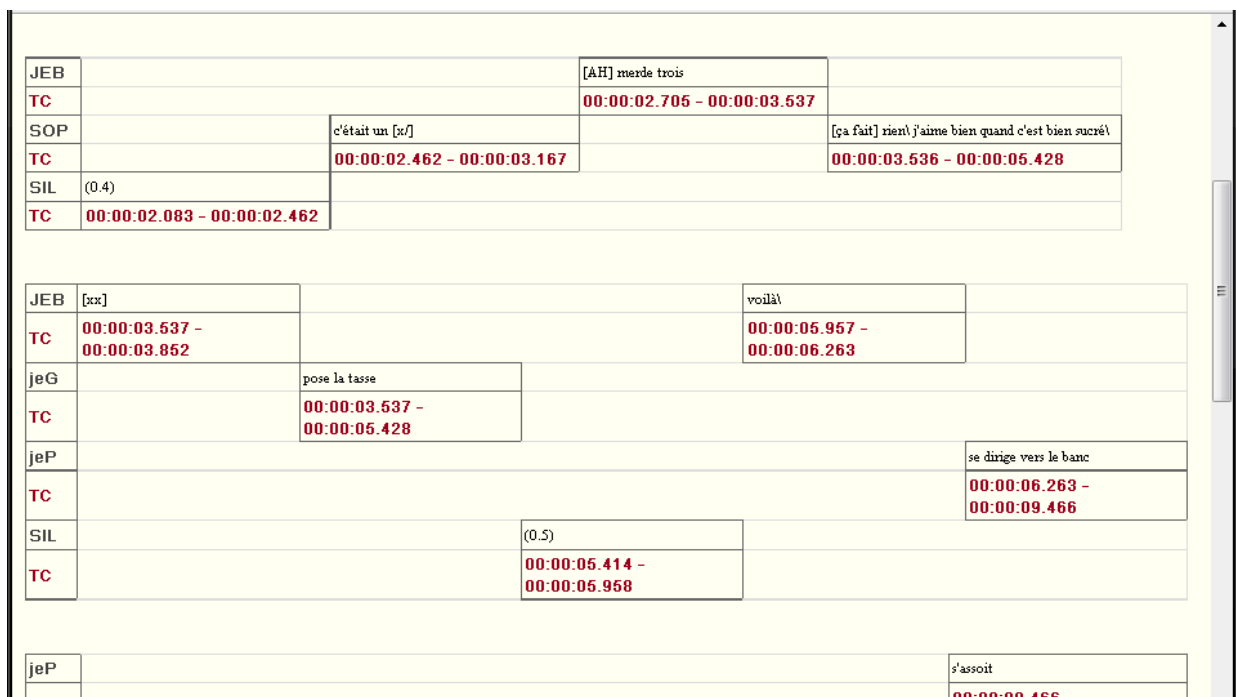


*Img. 9 : Options au moment de l'export en CSV*

- **Tiger-XML** : Il est possible d'exporter vers Synpathy, qui est un outil pour annoter, analyser et éditer graphiquement la structure syntaxique des énoncés. Cet outil a été développé par le Max Planck Institute de Psycholinguistique. Il est basé sur le Syntax Viewer provenant du projet de recherche TIGER.
- Texte interligne
- **HTML**

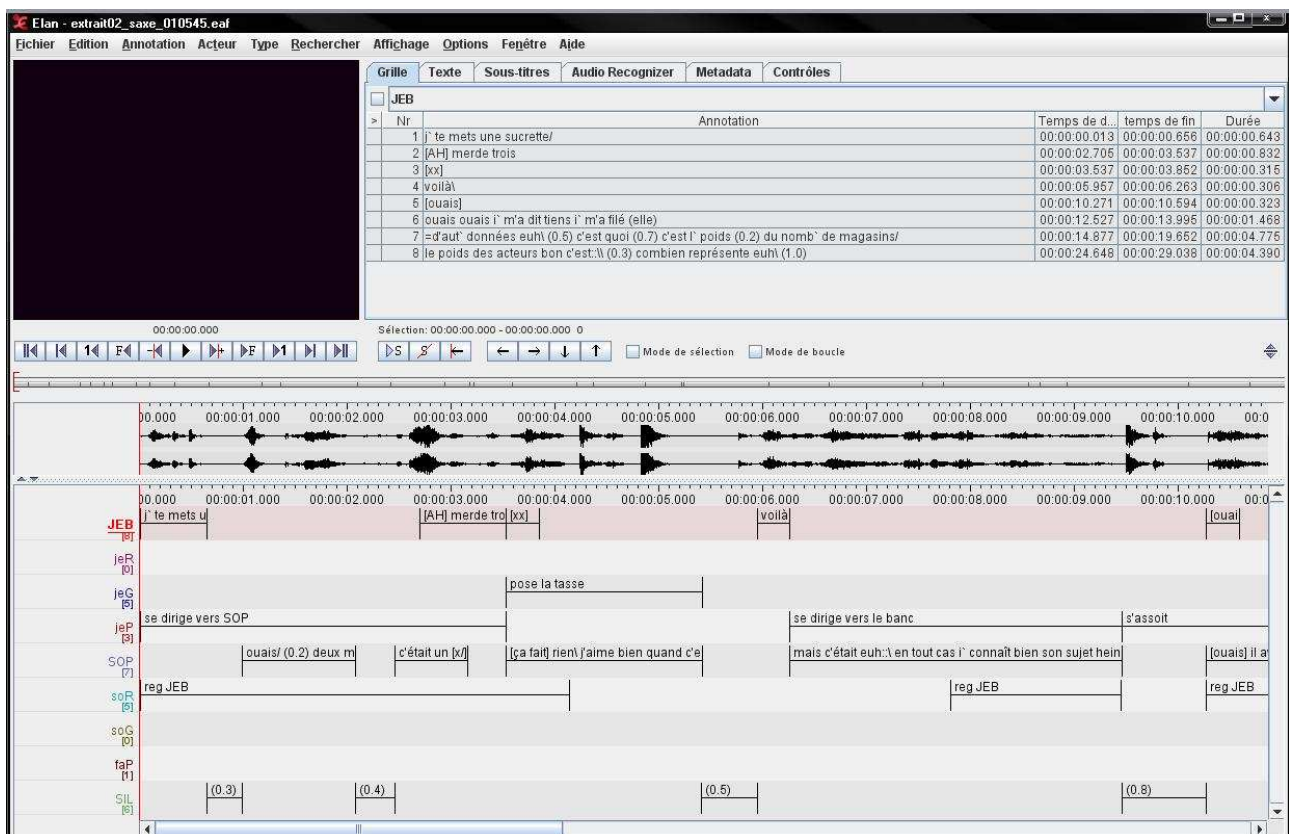


Img. 10 : Options au moment de l'export en HTML



Img. 11 : Résultat de l'export dans le navigateur

- Transcription de texte traditionnel
- **Textgrid Praat**
- **Liste de mots**
- Fichier SMIL : Le **Synchronized Multimedia Integration Language** (SMIL, prononcé "smile") permet de faire des présentations audiovisuelles interactives. SMIL est typiquement utilisé pour des "medias enrichis", des présentations multimédia qui intégrant des flux audio et vidéo avec des images, du texte ou tout autre type de média. SMIL est un outil facile à apprendre, comme le langage HTML, et de nombreuses présentations SMIL sont écrits en utilisant un simple éditeur de texte.
- Texte Quicktime
- Texte des sous-titres
- **Image depuis fenêtre ELAN**



Img. 12 : Résultat de l'export en image

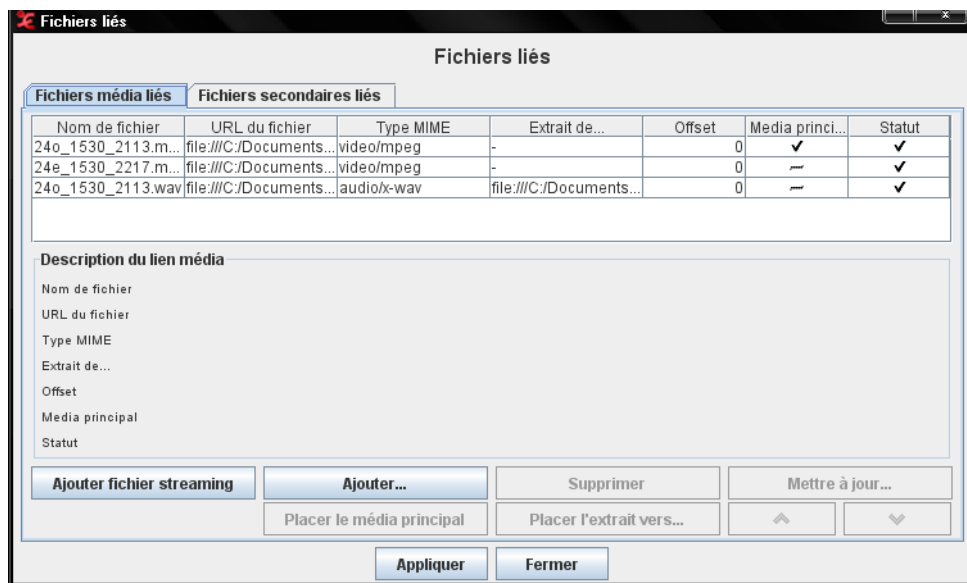
## 9 Fonction « Multiscopie » pour aligner plusieurs sources vidéo

ELAN permet d'aligner une source vidéo et audio simultanément. Il est également possible d'associer par exemple 2 vues vidéo d'une même scène (lié à une source audio aussi).

### 9.1 Solution 1 : Si la durée des clips sont strictement identiques

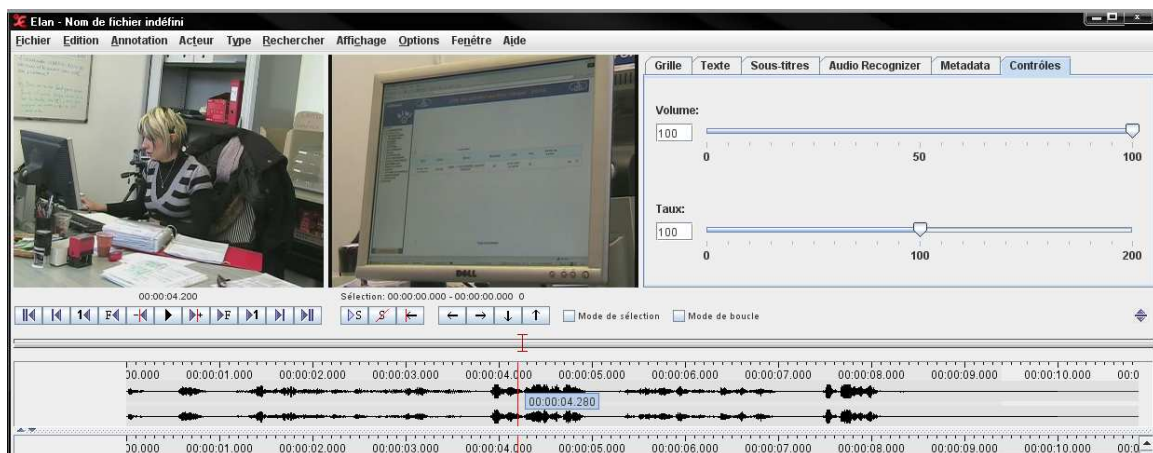
NB : Le timing de départ des différents clips doit être le même sinon cela provoquera un décalage dans la lecture des vues simultanées.

La gestion des fichiers multimédias liés se fait par le menu **Edition > Fichiers liés** :



Img. 13 : Boîte de dialogue Fichiers liés

Grâce à l'option **Ajouter**, on peut rajouter une autre source vidéo si on ne l'a pas déjà associée lors de la création du fichier ELAN vierge de départ.



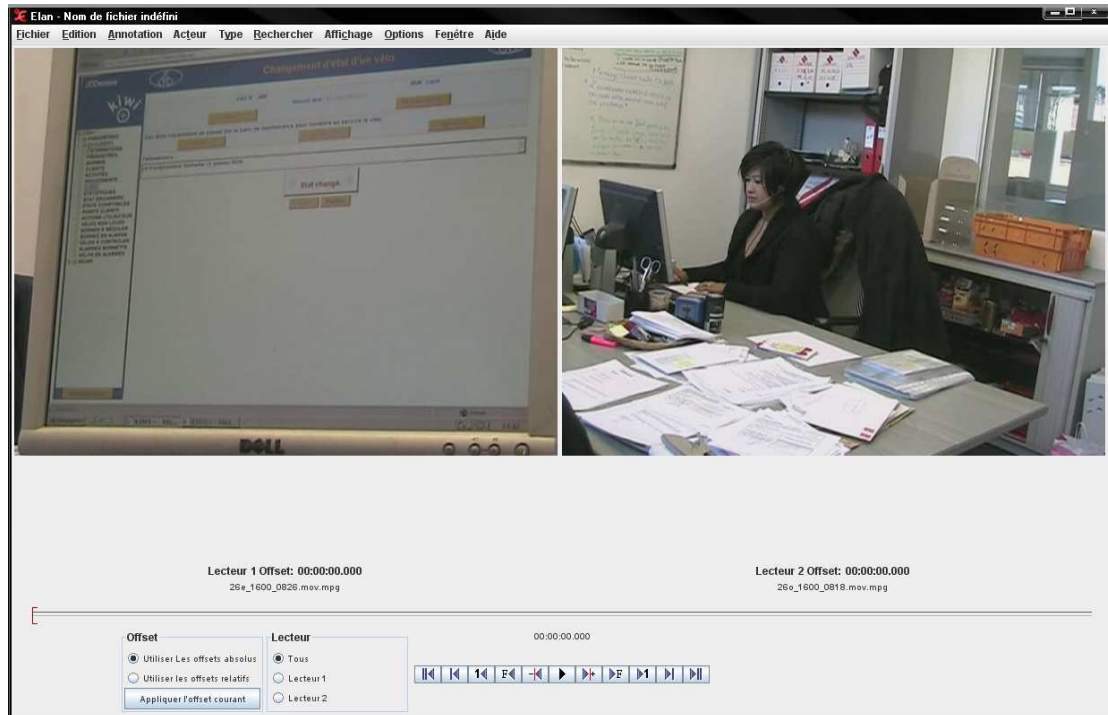
Img. 14 : Résultat d'association de 2 vues vidéo et 1 audio



## 9.2 Solution 2 : Si la durée des clips n'est pas équivalente

NB : Ici, le timing de départ des clips peut être différent, l'option permet de définir un timing de départ sur une ou plusieurs vues vidéo secondaires.

L'option nécessaire s'appelle « Mode synchronisation de média ». Après avoir ouvert un nouveau fichier ELAN composé des différentes vues vidéo à synchroniser (minimum 2), il est possible d'activer le mode de synchronisation par le menu **Options > Mode synchronisation de média** :



Img. 15 : Boîte de dialogue Mode synchronisation de média

- a) Faire le choix dont le time code sera définie :
  - - Offsets absolus = pour chaque vidéo, son propre timing sera affiché
  - - Offsets relatifs = la vidéo du player 1 est défini comme « le maître » (le principal), c'est-à-dire que la position du timing des autres vidéos seront exprimées au point de départ de cette vidéo principale, qui commencera au timing 00:00:00.000.
- b) Sélectionner le player 1 uniquement, puis définir un moment distinct de la vidéo qui est facile à repérer (point d'ancrage commun aux différentes vues)
- c) Sélectionner ensuite le player 2 et faire ce même repérage du même moment distinct de la vidéo 1.
- d) Enfin, valider par le bouton Appliquer l'offset courant.
- e) En sélectionnant le bouton de l'ensemble des vidéos, il est possible de les jouer simultanément et de vérifier si la synchronisation qui vient d'être définie est correcte.

Si elle ne l'est pas, il faut alors répéter les étapes b) à d) jusqu'à ce que le résultat soit correct.

- f) Pour quitter le mode synchronisation afin de procéder aux annotations de ces vues à présent synchronisées, il faut aller dans le menu **Options > Mode annotation**. Il est possible de créer les annotations sur la partition.