



**NOTATION, ANNOTATION ET ANALYSE
DE CORPUS MULTIMODAUX AVEC ELAN**
ATELIER DE FORMATION **19 et 20 NOVEMBRE 2014**
MSHS Bât A5, 5 rue Théodore Lefebvre 86 000 POITIERS

OBJECTIF

Le consortium « Corpus oraux et multimodaux » (IRCOM) coordonné par la fédération Typologie et Universaux Linguistiques sous l'égide de la très grande infrastructure de recherche HUMA-NUM a vocation à fédérer les équipes, laboratoires, chercheurs et enseignants-chercheurs engagés dans la constitution et le traitement de corpus oraux et multimodaux, afin de faire converger les pratiques. A cette fin, l'IRCOM propose un certain nombre de formations dans différents domaines liés à l'utilisation de corpus en linguistique. L'un de ces domaines concerne la notation, l'annotation et l'analyse de corpus multimodaux. Une session de formation à l'annotation avec le logiciel ELAN se déroulera les **mercredi 19 et jeudi 20 novembre 2014 à Poitiers.**

COMITE D'ORGANISATION ET FORMATEURS

L'atelier « Notation, annotation et analyse de corpus multimodaux avec le logiciel ELAN » est organisé par le groupe de travail *multimodalité et modalité gestuo-visuelle* en charge de ces questions au sein du consortium IRCOM. Trois formateurs interviendront dans cet atelier :

- Christophe Parisse (Modyco)
- Isabelle Rousset (Lidilem)
- Dominique Boutet (SFL)



DEROULEMENT DE L'ATELIER

La formation se déroulera sur 2 journées. Les participants, limités à 28 au total, seront divisés en deux groupes de travail selon le niveau d'expertise des participants : niveau débutant et niveau avancé. Les participants devront apporter leur ordinateur portable (configurations requises, voir « Avant d'assister à la formation »). Les corpus et schémas d'annotation utilisés seront fournis par l'IRCOM. La formation théorique alternera avec des sessions de mise en pratique et ce tout au long des deux journées.

INSCRIPTION ET PRISE EN CHARGE DES FRAIS

Le nombre de participants étant limité à 28 participants, nous demandons aux personnes intéressées de bien vouloir faire parvenir à **Claudia S. Bianchini** (claudia.savina.bianchini@univ-poitiers.fr) un bref CV, une lettre indiquant les intérêts et motivations du candidat pour participer à l'atelier, ainsi que la fiche de pré-inscription dûment complétée **avant le 15 avril 2014**. Les candidatures seront examinées par un comité *ad hoc* du Conseil Scientifique.

L'IRCOM prendra en charge les déjeuners et les pauses café qui seront pris sur place.

Programme préliminaire de l'atelier :

GROUPE DEBUTANTS

Jour 1 (19 novembre 2014) Prise en main du logiciel ELAN

- Chargement du logiciel. Présentation des principales fonctionnalités d'ELAN (découverte du menu), chargement de données primaires multimodales (vidéo et audio), création de pistes (acteurs), segmentations (manuelle, à la volée), annotations. Synchronisation des données primaires. (C. Parisse)
- Sensibilisation aux formats et codecs. Comment importer une annotation effectuée sous excel, ou d'autres logiciels d'annotation (Praat, Transcriber, Toolbox, Clan). Bâtir un template sous ELAN : stéréotypes et types associés aux pistes (acteurs). (C. Parisse)

Jour 2 (20 novembre 2014) Template et Fonctions avancées

- Bâtir un template sous ELAN : vocabulaire contrôlé. Associations VC, types linguistiques et pistes d'annotation. Fonctions avancées d'ELAN (merge et tokeniser). Répartir les annotations entre annotateurs et fusionner les transcriptions. (D. Boutet)
- Corriger une annotation : Fonction rechercher/remplacer (fichier simple ou multiple). Faire de requêtes sous ELAN, sauvegarde des requêtes et des résultats. Exporter vers Excel. (I. Rousset et C. Parisse)

GROUPE AVANCES

Jour 1 (19 novembre 2014) Prise en main du logiciel ELAN (Isabelle Rousset)

- Questions présidant à l'établissement d'un schéma d'annotation, élaboration d'un template : structuration de l'annotation (type des pistes et des attributs). Constitution de vocabulaires contrôlés, insertion de différents médias, modalités de segmentation, annotations. (I. Rousset)
- Import d'annotations, Fonctions avancées (merge, filtres, copy, tokeniser), Modifications/ réutilisation d'un template. Sensibilisation à la constitution d'un manuel de codage. (I. Rousset)

Jour 2 (20 novembre 2014) Requêtes et analyse

- Requêtes, utilisation d'expressions régulières, enregistrement des requêtes, recherches sur plusieurs fichiers, exports des résultats. Utilisation d'Excel : import, export des données, utilisation des fonctions d'Excel pour modifier des données, pour générer des tiers ou des données à réimporter le cas échéant. (C. Parisse)
- Insertion de données primaires de type biomécanique (fichiers liés), affichage dans ELAN, transformation des données en annotations (import des données depuis excel). (D. Boutet)



Avant d'assister à la formation,
il est nécessaire d'avoir téléchargé et installé :

Pour Windows et Mac OS :

- OS Windows requis : XP ou Vista ou Windows 7 ou Windows 8.
- 4 Go d'espace disque disponible pour l'installation des applications 5 Go supplémentaires pour le téléchargement du contenu
- Carte graphique avec les dernières versions des pilotes mises à jour ; Résolution d'écran 1 024 x 768
- Logiciel QuickTime 7
- 2 Go de mémoire RAM
- Ecran couleur avec carte vidéo couleurs 16 bits

Mac OS (Privilégiez tout de même un ordinateur PC)

- Processeur Intel® multicoeur 64 bits
- Mac OS X v10.6 à v10.
- 8.2 Go de mémoire RAM

Téléchargez les logiciels ELAN, Excel ou Libre Office, Notepad++ ou JEdit (pour les Mac)

**NOTATION, ANNOTATION ET ANALYSE
DE CORPUS MULTIMODAUX AVEC ELAN
ATELIER DE FORMATION 19 et 20 NOVEMBRE 2014
MSHS Bât A5, 5 rue Théodore Lefebvre 86 000 POITIERS**

FICHE DE PRE-INSCRIPTION

à renvoyer **avant le 15 avril 2014** à
claudia.savina.bianchini@univ-poitiers.fr

L'inscription sera validée après l'examen du dossier de candidature composé : 1) d'un bref CV ; 2) d'une lettre où le candidat expose ses intérêts et motivations à participer à l'atelier ; 3) du présent formulaire.

Nom, prénom :

Adresse personnelle :

E-Mail :

Téléphone :

Statut : Doctorant Chercheur / ITA Post-doctorant
 Enseignant-Chercheur Autre (précisez) :

Organisme :