

**Semaine de formation intensive NooJ**  
**Max Silberztein, Université de Franche-Comté**  
**INALCO, 9-13 janvier 2017**

NooJ est un environnement de développement utilisé pour formaliser huit niveaux de phénomènes linguistiques : orthographe et typographie, morphologie flexionnelle et dérivationnelle, syntaxe locale et structurelle, grammaire transformationnelle et analyse sémantique. NooJ contient des outils formels adaptés à chaque type de phénomène (grammaires rationnelles, hors contexte, contextuelles et non restreintes) et permet de développer des ressources à large couverture, pour de nombreuses langues.

NooJ est aussi utilisé pour analyser des corpus de textes par des linguistes mais aussi pour effectuer des analyses techniques, historiques, littéraires, sociologiques ou psychologiques de corpus. NooJ est gratuit et open source et fonctionne sous Windows, Mac OSX, LINUX et Unix, cf. : [www.nooj4nlp.net](http://www.nooj4nlp.net).

Le stage s'adresse particulièrement aux étudiants de Master, doctorants et universitaires intéressés par la linguistique et la linguistique de corpus ainsi que par l'analyse automatique de textes dans une perspective de sciences humaines. Les étudiants en Master qui peuvent et désirent valider le stage auprès de leur université doivent me prévenir dès leur inscription. Le stage dure une semaine : les matins sont dédiés au cours et aux travaux dirigés ; pendant les après-midis, des chercheurs présenteront diverses applications de NooJ.

ATTENTION : le stage est gratuit, mais les inscriptions sont obligatoires et les places sont limitées : il faut s'inscrire en envoyant son nom, son statut (étudiant / doctorant / universitaire) et le nom de son institution à [max.silberztein@univ-fcomte.fr](mailto:max.silberztein@univ-fcomte.fr).

Lundi 9 janvier

9H-12H30 : **Fouille de texte** : textes et corpus, requêtes et concordances, analyses statistiques.  
14H-14H40 : Cai Zhen (Univ. de Franche-Comté) : Dictionnaire électronique chinois et applications.  
14H40-15H20 : Alena Saint Joanis (INALCO) : Construction du dictionnaire de l'ukrainien.  
15H20-16H : Masako Watabe (Univ. Paris-Sorbonne) : Module NooJ polylectal pour la langue rromani.

Mardi 10 janvier

9H-12H30 : **Lexique & morphologie** : Unités Linguistiques Atomiques, dictionnaires, morphologie.  
14H-14H40 : Hamid Annouz (INALCO) : Dérivations nominales en kabyle.  
14H40-15H20 : Chantal Enguehard (Univ. de Nantes) : Morphologie des noms en tamajaq.  
15H20-16H : Maximiliano Duran (Univ. de Franche-Comté) : Dérivation des verbes quechua.

Mercredi 11 janvier

9H-12H30 : **Syntaxe locale** : requêtes syntaxiques, grammaires locales, annotation automatique de textes, extraction d'entités nommées.  
14H-14H40 : Héra Fehri (Univ. de Gabès, Tunisie) : Outil d'aide à la décision pour l'identification des plantes médicinales.  
14H40-15H20 : Med Aly Fall Seideh (Univ. de Gabès) : Les noms latins de plantes : reconnaissance des entités nommées et traduction français-arabe  
15H20-16H : Agnès Tutin (Univ. Stendhal) : Annotation semi-automatique d'expressions polylexicales.

Jeudi 12 janvier

9H-12H30 : **Syntaxe structurelle** : Grammaires hors contexte, contraintes d'accord et arbres.  
14H-14H40 : Mario Monteleone (Univ. de Salerne, Italie). *Résolution d'anaphores : méthodes linguistiques et statistiques*.  
14H40-15H20 : Denis Le Pesant (Modyco). Présentation de deux nouveaux dictionnaires NooJ.  
15H20-16H : Mourad Aouini (CNRS, UPS-CLT). Reconnaissance des entités nommées en moyen-français.

Vendredi 13 janvier

9H-12H30 : **Sémantique** : Analyse transformationnelle et génération de paraphrase automatique, levée d'ambiguïté sémantique, traduction  
14H-14H40 : Slim Mesfar (Univ. de Carthage, Tunisie) : Analyse sémantique automatique d'un corpus de tweets  
14H40-15H20 : Andrea Rodrigo (Univ. Nacional de Rosario, Argentine) : Analysing different Spanish adjectives using NooJ.