

Les nombreuses théories sur le développement du langage de l'enfant et l'importance de la morphosyntaxe dans ce développement méritent d'être mises à l'épreuve à l'aide de corpus réels obtenus dans des situations d'interactions naturelles entre enfants et adultes ou entre enfants. Pour cela, cet atelier a un but double : montrer comment gérer et transcrire des corpus de langage oral obtenus dans des situations naturelles ; montrer quel type d'analyse on peut réaliser sur les corpus obtenus.

Pour mettre ce projet en application, dans le prolongement du cours sur la morphologie et la syntaxe dans les corpus d'enfants et les corpus oraux, on utilisera le programme CLAN de CHILDES [1-2]. Ce programme est gratuit (offert à la communauté scientifique par l'université Carnegie Mellon de Pittsburgh) et permet de réaliser et vérifier des transcriptions de langage spontané en mettant en correspondance texte et vidéo, ainsi que texte et son. On utilisera des exemples de corpus, sons et vidéo (fournis aux participants) pour mettre en pratique le travail de transcription.

La première partie de l'atelier portera sur les problèmes de transcription à fin d'analyse lexicale, morphologique ou syntaxique. Ensuite, on utilisera les corpus transcrits pour réaliser tout d'abord des analyses lexicales. Puis on montrera comment réaliser une analyse morphosyntaxique [3] et comment utiliser cette analyse pour améliorer l'analyse lexicale et analyser le contexte des mots. Pour cela, il faut tirer parti des conventions de transcription, ce qui permet alors de faire des statistiques et recherches sur les corpus analysés. Enfin, on montrera comment créer de nouveaux corpus d'apprentissage afin de faire évoluer et améliorer l'analyse morphosyntaxique.

[1] MacWhinney, B. (2000). *The CHILDES project : Tools for analyzing talk* (3rd) (2nd ed.). Hillsdale, N.J: Lawrence Erlbaum.

[2] <http://childes.psy.cmu.edu/>

[3] Parisse, C., & Le Normand, M. T. (2000). Automatic disambiguation of morphosyntax in spoken language corpora. *Behavior Research Methods, Instruments, & Computers*, 32(3), 468-481.