

Le système CHILDES : brève présentation

CHILDES est un acronyme pour **C**hild **L**anguage **D**ata **E**xchange **S**ystem.

Le système CHILDES se compose essentiellement de deux outils informatiques (transcription et analyse) et d'une large base de données. Tous ces éléments sont disponibles gratuitement par téléchargement à partir du site du CHILDES <http://childes.psy.cmu.edu>. Tous ces outils existent en version Windows, Macintosh et UNIX.

Transcription (CHAT)

Le CHILDES propose un ensemble de normes de transcription spécialement construites pour l'étude des situations naturelles de dialogue. Ces normes sont connues sous le nom de CHAT (**C**odes for the **H**uman **A**nalysis of **T**ranscripts). Elles ont été adaptées à un grand nombre de langues naturelles, orales ou gestuelles.

Le CHILDES se compose ainsi d'un « Editeur » (Traitement de texte) automatisant un grand nombre de procédures de transcription. Cet Editeur permet la transcription selon les normes CHAT, mais aussi selon les normes **CA** (**C**onversation **A**nalysis). Le système fournit les polices servant à noter les caractères phonétiques (IPA) ou les langues utilisant des caractères différents de l'alphabet roman (UNICODE). L'Editeur comprend aussi un ensemble de possibilités additionnelles : il peut fonctionner en Mode « Codage », les fichiers transcrits peuvent être reliés à des fichiers audio ou vidéo numérisés, les énoncés transcrits peuvent être envoyés dans un analyseur de son (Praat), etc.

Analyses (CLAN)

Le CHILDES se compose aussi de programmes conçus pour analyser les fichiers transcrits en CHAT ; certains d'entre eux fonctionnent aussi sur les fichiers CA (Conversation Analysis). Ces programmes sont connus sous le nom **CLAN** (**C**omputerized **L**anguage **A**nalysis). Ils fonctionnent à partir d'une fenêtre « Commandes » de l'Editeur et sont ainsi disponibles à tout moment. Ils proposent des recherches à l'intérieur de larges corpus, des décomptes fréquentiels et des analyses de co-occurrence, des analyses phonétiques, lexicales, morphosyntaxiques, les calculs automatiques de divers indices utiles dans l'examen des productions langagières : MLU (Longueur Moyenne des Enoncés), DDS (Developmental Sentence Score), diverses mesures de diversité lexicale, des analyses d'interactions, etc.

Quelque soit l'origine des transcriptions (Windows, Mac, UNIX ou DOS), les programmes effectuent des transformations automatiques de l'une à l'autre.

Banque de données

Enfin la base de données comporte de nombreux corpus transcrits concernant diverses langues, avec parfois les fichiers audio et vidéo qui leur correspondent.

Pour plus de détail, consulter la présentation disponible sur le site de l'école, et télécharger les manuels Chat.pdf et Clan.pdf sur le site du CHILDES <http://childes.psy.cmu.edu>.