Corpus de Interlengua Española de Aprendices Sinohablantes

CINEAS

El CINEAS es un **corpus lingüístico** que se ha construido en el marco del **proyecto FFI2016-80280-R**, de MINECO, cuyo título es *Elaboración y catalogación de un corpus de textos escritos en ELE producidos por estudiantes sinohablantes* (véase sección Proyectos financiados).

Este corpus cuenta, en la versión 1.0, con **4378 textos** escritos por **estudiantes chinos de español** de los niveles **A1, A2, B1, B2 y C1** en contexto universitario de aprendizaje formal y en inmersión lingüística. Son textos recogidos en diversas universidades chinas y españolas. En estos momentos, el corpus ronda **unas 435.000 palabras**.

Los textos, escritos originalmene a mano, han sido **digitalizados** para facilitar su tratamiento informático. Asimismo, **se han catalogado** atendiendo a varios factores sociales y lingüísticos. Próximamente serán etiquetados los errores identificados en ellos para que puedan aplicarse búsquedas simples, cruzadas y avanzadas.

CINEAS cuenta con un acceso general a parte del corpus, que es de uso público, y con un acceso restringido a los investigadores e investigadoras del proyecto, que permite trabajar con la totalidad del corpus. En estos momentos, un usuario cualquiera puede hacer consultas simples, cruzadas y avanzadas, a partir de metadatos de carácter extralingüístico, en 2922 textos. En cambio, no puede todavía hacer consultas del etiquetado de errores que contiene el aplicativo *TextAnnot.*

Se puede acceder al corpus desde **aquí** [https://cineas.udl.cat/].