

Visit of corpus researcher from Tokyo

Posted on Friday 16th August 2013



Takehiko Maruyama from the Center for Corpus Development, at the National Institute for Japanese Language and Linguistics in Tokyo visited the Centre for Corpus Research on 5 August 2013.

He gave a short presentation about the work of the National Institute for Japanese Language and Linguistics, and in particular two major corpus development projects he has been involved in: the 'Corpus of Spontaneous Japanese' (which contains 650 hours of speech with 7.52 million tokens) and the 'Balanced Corpus of Contemporary Written Japanese', containing 100 million words randomly taken from various written texts.