

International Symposium on Signed and Spoken Linguistics (1)

‘Description, Documentation and Conservation’

National Museum of Ethnology, Osaka, Japan. July 28–29, 2012

Abstract

Documentating British Sign Language: The BSL Corpus Project

Adam SCHEMBRI, Jordan FENLON, Ramas RENTELIS, Rosemary STAMP

and Kearsy CORMIER

La Trobe University

The British Sign Language Corpus Project (BSLCP) was a project funded by the UK Economic and Social Research Council (ESRC, 2008-2011) that aimed to create a machine-readable corpus of spontaneous and elicited British Sign Language (BSL) digital video data collected from deaf native, near-native and fluent signers across the United Kingdom. Researchers at University College London led the project, with co-investigators based at Bangor University (Wales), Heriot-Watt University (Scotland), Queens University Belfast (Northern Ireland) and the University of Bristol (England). In the field of sign language studies, the BSLCP represents a unique combination of methodology from variationist sociolinguistics, language documentation, and corpus linguistics. The project undertook studies on sociolinguistic variation, language change and lexical frequency in BSL while simultaneously documenting BSL via creation of a corpus. Thus, unlike previous large-scale sign language sociolinguistic projects, the dataset has been given metadata descriptions, and has been archived and made accessible online. Since June 2011, video data have been available online and searchable via key metadata at <http://www.bslcorpusproject.org/data/>.

As annotation is added to the online digital video dataset (including gloss-level and more detailed linguistic annotations as well as English translations via ELAN annotation files), we hope that it will become a standard reference and core data source for all researchers investigating BSL structure and use.