ECONOMIC AND SOCIAL RESEARCH COUNCIL
END OF AWARD REPORT

For awards ending on or after 1 November 2009

This End of Award Report should be completed and submitted using the **grant reference** as the email subject, to **reportsofficer@esrc.ac.uk** on or before the due date.

The final instalment of the grant will not be paid until an End of Award Report is completed in full and accepted by ESRC.

Grant holders whose End of Award Report is overdue or incomplete will not be eligible for further ESRC funding until the Report is accepted. ESRC reserves the right to recover a sum of the expenditure incurred on the grant if the End of Award Report is overdue. (Please see Section 5 of the ESRC Research Funding Guide for details.)

Please refer to the Guidance notes when completing this End of Award Report.

Grant Reference	RES-062-23-0825		
Grant Title	The British Sign Language (BSL) Corpus Project: Sociolinguistic variation, language change, language contact and lexical frequency in BSL		
Grant Start Date	1 January 2008	**Total Amount Expended:**	£ 914,836.05
Grant End Date	30 June 2011		
Grant holding Institution	University College London		
Grant Holder	Dr. Kearsy Cormier, Jan-June 2011 (Co-investigator, 2008-2010)		
Grant Holder's Contact Details	**Address**	**Email**	
	Deafness Cognition and Language Research Centre, 49 Gordon Square, London WC1H 0PD	k.cormier@ucl.ac.uk	
		Telephone	
		0207 679 8674	

Co-Investigators (as per project application):	Institution
Prof. Margaret Deuchar	Bangor University
Ms. Frances Elton	University College London
Prof. Donall O'Baoill	Queens University Belfast
Dr. Adam Schembri (Grant holder, 2008-2010)	University College London (2008-2010) La Trobe University, Melbourne (2011)
Dr. Rachel Sutton-Spence	University of Bristol
Prof. Graham H. Turner	Heriot-Watt University
Prof. Bencie Woll	University College London

1. NON-TECHNICAL SUMMARY

Please provide below a project summary written in non-technical language. The summary may be used by ESRC to publicise your work and should explain the aims and findings of the project. *[Max 250 words]*

The overall aim of this project was to create a corpus (i.e., a computerised collection of language recordings) of British Sign Language (the language of the British Deaf community, commonly known as 'BSL'), and to conduct corpus-based investigations of aspects of the vocabulary, grammar and sociolinguistics of BSL. The first major aim of this project was to create a centralised source of data for ongoing research efforts that aim to understand BSL and sign languages of Deaf communities in general. The first national web-based and publicly accessible BSL corpus, consisting of video recordings of 249 Deaf signers of BSL from 8 sites across the United Kingdom, is now available online at http://bslcorpusproject.org/data. The second major aim was to use the data collected to investigate variation and change in the phonology, vocabulary and grammar of BSL and to investigate frequency of lexical items in the language (i.e., to find out which signs are the most common in conversation). Overall the studies on sociolinguistic variation and change and lexical frequency have clarified previously held assumptions about how BSL varies across different social groups and in different linguistic contexts, how it is changing, and which signs occur frequently in natural conversation. These findings will lead directly to improved sign language teaching resources that will more accurately describe how the language is used by a range of subgroups within the British Deaf community. This will in turn bring about improvements to the training of BSL teachers, sign language interpreters and educators of deaf children.

2. PROJECT OVERVIEW

a) Objectives

Please state the aims and objectives of your project as outlined in your proposal to the ESRC. *[Max 200 words]*

The project aimed to create a machine-readable digital corpus of naturalistic and elicited British Sign Language (BSL) data collected from Deaf native and near-native signers across the United Kingdom in order to: (1) establish a permanent and representative record of BSL via a digital video corpus consisting of elicited grammatical and lexical data, elicited and spontaneous narratives and conversational and sociolinguistic interview data. It was to be partly annotated and tagged using ELAN software, given standard metadata descriptions, and archived to be made accessible and searchable on-line by researchers worldwide, serving as a model system for openly accessible video archiving; (2) enable the formulation of new observations on the use and structure of BSL that are corpus-based. Specifically we aimed to conduct empirical studies of (a) sociolinguistic variation in BSL phonology, lexis and grammar, (b) language change in BSL (focusing on grammaticalisation), (c) lexical frequency and (d) bilingualism and language contact; (3) establish a reference corpus of BSL data which can be compared to corpora of other signed languages; and (4) establish a reference corpus of BSL data against which future examples of the language can be compared to establish the type and rate of change that may be occurring.

b) Project Changes

Please describe any changes made to the original aims and objectives, and confirm that these were agreed with the ESRC. Please also detail any changes to the grant holder's institutional affiliation, project staffing or funding. *[Max 200 words]*

The overall aims and objectives did not change.

Project extension and change in PI: In September 2010, a 6-month no-cost extension to June 2011 was requested. Because Dr. Schembri was taking up a new post in Australia from January 2011, we suggested that Dr. Kearsy Cormier (originally a project CI) take over as PI during the extension and that Dr. Schembri carry on as a project CI instead.

Staffing: Ms. Sally Reynolds, project research associate (1/2008-2/2009) was replaced by Dr. Jordan Fenlon (2/2009-6/2011). Ramas Rentelis, who began as a two-year part-time project research assistant 11/2008, changed to full-time (5/2010-6/2011). Deaf community fieldworkers were appointed for each of the 8 sites where filming occurred. Some fieldworkers initially appointed were unable to complete the task and were later replaced: Mark Nelson and Mischa Cooke were replaced by Melinda Napier for London and Jeff Bratton-Wilson was replaced by Sarah Lawrence for Cardiff.

Studentship(s): The original project proposal included two 3-year project-linked PhD studentships, including one on bilingualism and one on lexical variation. No suitable candidates for either studentship were found, so upon agreement with the ESRC, we offered a single 1+3 studentship focusing on the lexical variation study.

c) Methodology

Please describe the methodology that you employed in the project. Please also note any ethical issues that arose during the course of the work, the effects of this and any action taken. *[Max. 500 words]*

The British Sign Language Corpus Project aimed to create a machine-readable digital video corpus of spontaneous and elicited British Sign Language (BSL) collected from Deaf native and near-native signers. In total, 249 participants, of varying genders, ages, social classes, ethnicities, and language backgrounds, were filmed in 8 key regions across the UK (London, Bristol, Cardiff, Birmingham, Manchester, Newcastle, Glasgow and Belfast). Participants were filmed in pairs and were assembled in each of the sites by Deaf community fieldworkers who were all Deaf individuals living in the area with knowledge of the local Deaf community. The fieldworkers were responsible for identifying and recruiting native or near-native BSL users who have lived in their local community for at least ten years. Participants took part in four tasks: recounting a personal experience narrative, engaging in a 30 minute conversation, participating in an interview on language awareness and attitudes, and responding to a task designed to elicit 102 signs known to vary across the UK. Each filming session included three cameras: one focussing on each of the two participants, and a third on both participants.

Prior to data collection, ethical approval was obtained from the UCL Ethics Committee (ID 0864/001). One ethical issue that arose during the course of the project (discovered during annotation and translation of the data) was that in the interview task, participants named particular people and groups within the Deaf/sign language community. As a consequence, all of the interview data is restricted and viewable only by registered researchers upon signing a confidentiality agreement (for more detail see http://www.bslcorpusproject.org/cava/activities). The conversation data is also only available to registered researchers rather than the general public. This was planned before data collection to lessen the effects of the observer's paradox (i.e. the fact that linguists wish to observe the variety of language used when people are not being observed).

All of the video data have been catalogued with standard metadata, archived, and made accessible online for researchers (http://bslcorpusproject.org/cava). The corpus dataset was offered to the Economic and Social Data Service but the offer was declined due to the size of the dataset. The CAVA Archive is the result of collaboration between UCL, JISC and the UK Data Archive to archive large collections of audio-visual recordings of human communication (http://www.ucl.ac.uk/ls/cava/). Some video data (specifically the narrative and lexical elicitation tasks) have additionally been made available online for viewing by the general public (http://bslcorpusproject.org/data). The dataset has been partly annotated using ELAN software, with annotations being linked to specific studies (see Project Findings). Those video files which have been annotated using ELAN are machine-readable. (All annotations completed as of June 2011 are linked to specific studies on sociolinguistic variation and change and lexical frequency.) It is expected that ELAN annotation files will be made available via the CAVA Archive at a later date, as annotations are completed and checked in preparation for studies to be published.

d) Project Findings

Please summarise the findings of the project, referring where appropriate to outputs recorded on the ESRC website. Any future research plans should also be identified. *[Max 500 words]*

> All BSL Corpus Project video data have been made available online (see Methodology). In addition, the following studies have been completed.
>
> *Phonological variation in the 1 handshape study:* This study examined variation in BSL signs produced with the 1 handshape (i.e., a typical pointing handshape with extended index finger). Analysis of 2110 signs indicates that variation in the 1 handshape is conditioned by linguistic factors including (a) grammatical category of the sign, with functor signs showing significantly more variation than content signs, and (b) phonological environment, with the preceding and following segments conditioning the variation. Conference presentations on this study were given at Linguistics Association of Great Britain (2009), New Ways of Analysing Variation (2009), and International Society for Gestures Studies (2010).
>
> *Lexical variation and change in numbers, colours, and placename signs:* In this study on lexical variation conducted by project-linked student Rose Stamp, we have found that a) younger signers, b) signers who attended a school outside of their region, and c) signers with non-signing parents were all less likely to use traditional signs (which may be specific to the region they lived in), compared to a) older signers, b) signers who attended school in their region and c) signers with signing parents. We also found that sign category was an important factor (signs for countries are changing faster than signs for colours). Conference presentations on this study were given in 2009 (UK Language Variation and Change), 2010 (Theoretical Issues in Sign Language Research Conference) and 2011 (International Conference on Language Variation in Europe).
>
> *Lexical frequency:* The study on lexical frequency includes 24,920 sign tokens from the corpus conversation data: approximately 500 signs each from 50 participants from two regions, balanced as much as possible for age, gender, and language background. Preliminary results (presented at the Symposium in Applied Sign Linguistics in 2011) indicate that 60% of the data consists of signs from the core lexicon, with high frequency items including more content words than is typical of a spoken language like English. This data provides the first lexical frequency information for a sign language based on a large conversational dataset, and complements existing work on sign familiarity.
>
> *Outputs:* In addition to the peer-reviewed conferences and workshops noted above (in the UK and internationally), we have given presentations invited by local universities and training centres (e.g. Middlesex University and City Lit Centre for Adult Learning in London), and at Deaf community events hosted by DCAL and/or BSLCP (DCAL Deaf Open Day in London, 2010, and 6 UK-wide DCAL Roadshow events in 2011).
>
> *Future plans:* Publications are in preparation for the phonological/lexical variation studies and the lexical frequency studies. Other studies are underway with project CIs (e.g. sociolinguistic variation and change in backchanneling with Rachel Sutton-Spence, and translation issues with Graham Turner). These will be submitted to high-impact journals in theoretical and applied/corpus linguistics, sociolinguistics, and gesture studies. Also a proposal is in preparation for BSL Corpus Project 2 (2012-2017), which will focus on grammatical variation in the BSL Corpus.

e) Contributions to wider ESRC initiatives (eg Research Programmes or Networks)

If your project was part of a wider ESRC initiative, please describe your contributions to the initiative's objectives and activities and note any effect on your project resulting from participation. *[Max. 200 words]*

> The BSL Corpus Project included three Co-Investigators (Woll, Cormier & Elton) from the Deafness, Cognition and Language Research Centre (DCAL), an ESRC-funded research centre, and all core project staff were physically based at DCAL. The studies on variation and change in BSL help extend DCAL's work on production, comprehension, processing, acquisition and neural bases of BSL. For example, the study on lexical frequency in BSL directly feeds into work done by researchers at DCAL by providing frequency information needed for designing language processing experiments. Significantly, the BSL lexical database created under BSLCP will be developed into an online dictionary under DCAL (2011-2015).
>
> The project has also benefited from being part of DCAL. For example, DCAL has employed several junior trainee research staff with the aim of exposing them to different research projects,

several of whom (Breish Rowe and Indie Johal) have contributed to work on the project. Also, the 2009 corpus workshop "Sign language corpora: Linguistic issues" was hosted at UCL and run by project staff/students working in conjunction with DCAL staff and students.

Additionally, at Heriot-Watt University, one of the project partner institutions, BSLCP data informs the research programme of RCUK Academic Fellow, Elaine Farrow.

3. EARLY AND ANTICIPATED IMPACTS

a) Summary of Impacts to date

Please summarise any impacts of the project to date, referring where appropriate to associated outputs recorded on the ESRC website. This should include both scientific impacts (relevant to the academic community) and economic and societal impacts (relevant to broader society). The impact can be relevant to any organisation, community or individual. *[Max. 400 words]*

Most immediately, the BSL Corpus Project has increased capacity building relating to junior researchers (both Deaf and hearing) within sign language linguistics; this is evident from the placements/progress of project staff/students. Jordan Fenlon has won a prestigious Fulbright Scholar-in-Residence award to teach 4 courses in the Linguistics Department at Gallaudet University in Washington DC for 2011-2012. Ramas Rentelis is completing an MSc in Deafhood Studies at Bristol University in 2011; for his MSc dissertation, he is examining the use of mouth patterns in the BSLCP data. Rose Stamp successfully upgraded from MPhil to PhD in 2011 and is due to complete her PhD on lexical variation and change in BSL in 2012. Heriot-Watt's lead BSLCP translator Kyra Pollitt commenced an AHRC PhD studentship on BSL at Bristol University in 2011.

In terms of impact on BSL teaching and research, work is underway by DCAL (2011-2015) to produce an online dictionary and reference grammar for BSL, both of which will be based on the lexical database created for the BSLCP study on lexical frequency (results of this study have been presented to teachers, interpreters and researchers at DCAL and at Bristol University in 2011). This dictionary and reference grammar will prove invaluable references for research and teaching/learning of BSL and for linguistics more generally.

Internationally, the BSL Corpus is one of very few large sign language corpus projects (along with projects in Australia, The Netherlands and Germany) and only the second to have video data available online (after The Netherlands). The BSL Corpus team was involved in the Sign Language Corpora Network (2009-2010) and maintains strong links with researchers working on these corpus projects. The BSLCP-hosted 2009 workshop "Sign Language Corpora: Linguistic Issues" had international attendance (>100 academics), and many were researchers starting or planning sign language corpora in their own countries. Since then, the BSL Corpus team have worked and consulted with such researchers, including those from USA, Italy, India and Japan.

The BSL video data and web interfaces were gradually curated/developed and uploaded from early 2011. From July 2011, the form for requesting access to the conversation and interview data has been available. Since then (as of Sept 2011) we have begun to receive requests for access from researchers, including a student at Bristol University and two sign language researchers in Japan planning to create a Japanese Sign Language corpus. We expect to receive many more in the future.

b) Anticipated/Potential Future Impacts

Please outline any anticipated or potential impacts (scientific or economic and societal) that you believe your project might have in future. *[Max. 200 words]*

An accessible online BSL corpus will have a significant and tangible impact on sign language research and also linguistics more generally. Eventually, the BSL corpus will be accessible and searchable by researchers world-wide, thus facilitating access to primary data for meaningful peer-review of claims (by both sign linguists and linguists not familiar with sign languages) and allowing subsequent researchers to make their own new corpus-based analyses. Although the ELAN annotation files created under this project may not be available online immediately, we have established a system for collaborating with researchers wishing to annotate the BSL Corpus data themselves (http://www.bslcorpusproject.org/cava/cava-repository-end-user-terms-and-conditions/). Thus other researchers will be able to create their own annotation files, and/or take annotation files created in this project and add further levels of annotation, enabling the extraction of patterns in the data that are extremely difficult to arrive at in any other way. From a practical perspective, the corpus will also be of enormous benefit to students and teachers of BSL and interpreter trainers and their students, serving as an invaluable resource in sign language/interpreting classrooms across the country. Better sign language/interpreting resources will result in improved services for deaf people, better ensuring their full participation in society.

You will be asked to complete an ESRC Impact Report 12 months after the end date of your award. The Impact Report will ask for details of any impacts that have arisen since the completion of the End of Award Report.

4. DECLARATIONS

Please ensure that sections A, B and C below are completed and signed by the appropriate individuals. The End of Award Report will not be accepted unless all sections are signed.

Please note hard copies are NOT required; electronic signatures are accepted and should be used.

A: To be completed by Grant Holder

Please read the following statements. Tick ONE statement under ii) and iii), then sign with an electronic signature at the end of the section (this should be a image of your actual signature).

i) The Project

This Report is an accurate overview of the project, its findings and impacts. All co-investigators named in the proposal to ESRC or appointed subsequently have seen and approved the Report.	☒

ii) Submissions to the ESRC website (research catalogue)

Output and impact information has been submitted to the ESRC website. Details of any future outputs and impacts will be submitted as soon as they become available.	☒
OR	
This grant has not yet produced any outputs or impacts. Details of any future outputs and impacts will be submitted to the ESRC website as soon as they become available.	☐
OR	
This grant is not listed on the ESRC website.	☐

iii) Submission of Datasets

Datasets arising from this grant have been offered for deposit with the Economic and Social Data Service. *<See Methodology section>*	☒
OR	
Datasets that were anticipated in the grant proposal have not been produced and the Economic and Social Data Service has been notified.	☐
OR	
No datasets were proposed or produced from this grant.	☐

SIGNATURE:

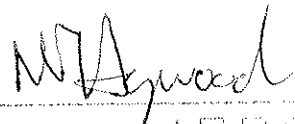NAME: Kearsy Cormier   DATE: 26 September 2011

B: To be completed by Head of Department, School or Faculty

Please read the statement below then sign with an electronic signature to confirm your agreement.

This Report is an accurate overview of the project, its findings and impacts.

SIGNATURE:

NAME: Prof David Shanks

POSITION: Head, Division of Psychology and Language Sciences   DATE:26/09/11

C: To be completed by Finance Officer of Grant-Holding Research Organisation

Please read the statement below then sign with an electronic signature to confirm your agreement.

ESRC funds have been used in accordance with the ESRC Research Funding Guide. All co-investigators named in the proposal to ESRC or appointed subsequently have seen and approved the Report.

SIGNATURE:

NAME:

POSITION: DATE: 29/9/11

Mark Hopwood
Research Finance Supervisor
Research Administration