COLLOQUIA

Colloquium on Dictionaries and Endangered Languages: Technology, Revitalization, and Collaboration

Organizer: Sarah Ogilvie University of Cambridge (svo21@cam.ac.uk)

Abstract:
This colloquium is dedicated to the topic of endangered languages and dictionaries, and focuses particularly on issues relating to dictionaries and technology, revitalization, and collaboration. Six international experts who specialize in the documentation and description of endangered languages will present papers that advocate the use and development of new technologies in lexicography; methodologies of collaboration and capacity building; and lexicographic techniques for language learning and revitalization.

Dictionaries and Endangered Languages: Technology, Revitalization, and Collaboration
Sarah Ogilvie, University of Cambridge (svo21@cam.ac.uk)

This paper opens the Colloquium, and presents the results of a survey of over two hundred dictionary projects conducted by the Endangered Languages and Dictionaries Project at the University of Cambridge. It provides a general overview of how dictionaries can move beyond being mere static records of language towards being more dynamic tools for language maintenance and learning.

Dealing with variation in a trilingual dictionary
Peter Austin, SOAS (pa2@soas.ac.uk)

Drawing on the experience of compiling a trilingual dictionary of the Sasak language, spoken on Lombok in eastern Indonesia, this paper proposes innovative lexicographic techniques for dealing with linguistic variation on a host of levels whether it be regional, social, literary, or contact variation.

Documenting indigenous knowledge in dictionaries: the Marquesas Islands
Gabriele Cablitz, University of Kiel (gcablitz@googlemail.com)

This paper looks at the creation of multimedia encyclopedic lexicons which are able to represent folk taxonomies and ethno-ontologies. It is based on the results of a DoBeS project that aims to document, maintain, and preserve the indigenous knowledge of the Marquesan speech community in the southern Pacific.
Syntax and text in an online lexicon database: Linguist-learner collaborations
Andrew Garrett, University of California, Berkeley (garrett@berkeley.edu)

Based on Garrett’s experience of creating online dictionaries of Karuk and Yurok, two indigenous languages of northern California, this paper discusses the lexicographic benefits of lexical tagging and syntactic parsing of text corpora using a dependency-grammar treebank. These methods facilitate optimal functionality of online learners’ dictionaries such as the ability to generate automatically collocational information for all lexical items.

Offline collaborative dictionary making with WeSay
John Hatton, SIL International (john_hatton@sil.org)

More and more dictionary projects today have multiple authors. This is especially the case for projects which seek to involve community members in the compilation process. But how can teams collaborate if they don’t all live in the same location and they don’t all have unlimited access to the Internet? This paper presents a new solution: Chorus developed by the creators of Wesay.

Building a lexical database with multiple outputs: examples from legacy data and from multimodal fieldwork
Nick Thieberger, University of Melbourne / University of Hawai'i (thieberg@hawaii.edu)

Based on ongoing work on dictionaries of South Efate (Oceanic, Central Vanuatu) and Ngarrindjeri (Pama-Nyungan, Australia), this paper presents two ways of creating structured lexical data suitable for subsequent reuse depending on different contexts. One model focuses on the process of building a contemporary multimedia corpus and dictionary, and the other focuses on an historical paper-based dictionary.