Documenting endangered languages using LEXUS and ViCoS

Yélî Dnye is a Papuan language spoken on Rossel Island in the Louisiade Archipelago, Papua New Guinea. Language documentation has been a continuous process since 1995 within the project Pioneers of Island Melanesia (Stephen C. Levinson). Different linguistic and cultural topics are recorded on video, audio and images. For Yélî Dnye a toolbox dictionary has been created containing over 6000 lexical entries. A large part of these entries are nouns representing objects in the natural world. The original dictionaries are imported in LEXUS and enriched with multimedia fragments. These fragments represent the meaning of the words in their indigenous contexts.

Iwaidja is a language spoken by approximately 200 speakers on the Cobourg Peninsula, Northern Territory, Australia. The language has been documented within the framework of the DooSsé endangered languages project (Bruce Birch, Nick Evans and others).

As a follow up on the documentation project the Iwaidja team is aiming at the creation of an accessible and attractive online dictionary, with multimedia capability and cross-referenced conceptual spaces will contribute to conservation, revitalization and empowerment of the language and community. As a pilot we have imported the current Iwaidja lexicon into LEXUS. The lexicon consists of 3500 entries and was originally created in Toolbox. The lexicon has been formatted and initial enrichment with multimedia has started.

LEXUS
LEXUS is a tool for the creation and exploration of multi-media online lexica in the context of language description and documentation.

ViCoS
ViCoS (Visualising Conceptual Spaces) aims at providing a simple interface between the lexical space (giving access to lexical resources) and the ontological space (where words relate concepts with each other).

Major Goals
Language documentation aims at the creation of representative and long lasting, multi-purpose record of natural languages. It contributes to maintain, consolidate or revitalize endangered languages and thus also contributes to the description of cultural elements of a language community. Our aim is to increase this cultural aspect by allowing users to complement linguistic information with ontological information.

Two groups of users are supported:
- Scientists such as linguists and anthropologists may use LEXUS and ViCoS to contribute to and use resources to study the language and culture of the community, or to compare them to the ones of other communities. ViCoS complements LEXUS by making available easy-to-use technology to impose a conceptual view over the lexical information in LEXUS. Moreover ViCoS permits to define structures that are orthogonal to the ones defined in the standardized LEXUS-LMF lexicon schema.
- Speech community members may use LEXUS to browse the lexicon by word list. In addition ViCoS can be used to describe their languages and culture and to learn from such resources. For community members, words are keys to access and describe relevant parts of their life and cultural traditions, such as food preparation, medicine, ceremonies, legends etc. For them, words are of foremost practical rather than theoretical nature; and their understanding of words is best described by the various associations they evoke rather than in terms of any formal and abstract theory of meaning.

ViCoS can be accessed from the LEXUS workspace

ViCoS is designed to add value to LEXUS by allowing users to create arbitrary relations between lexical entries or their parts, such as images or example sentences. The resulting network of semantic relations between entries form a kind of conceptual space, where words denote increasingly complex concepts. While each concept is anchored in the lexical space, its meaning becomes culturally relevant in the context of the concepts it is connected to.

The main requirement for the construction of knowledge spaces is thus to enable community members to anchor the words of a linguistic resource according to their classifications, hence creating a kind of “ethno-ontology”. It seems clear that existing ontologies have only limited relevance here. Their proper and effective use requires considerable expertise or training, and they also induce a significant and usually Westernized bias of how the world should be modeled.

LEXUS and ViCoS are part of the LAT suite.

Max Planck Institute for Psycholinguistics
Technical Group
Jacquetijn Ringersma
Marc Kemps-Snijders
Claus Zinn
Hub Verweij