The Innu Language Documentation Project

Lynn Drapeau and Renée Lambert-Brétière
Université du Québec à Montréal

Background

Over the past twenty years linguists have been alerting the international community of the imminent threat to the world's minority languages and especially to oral aboriginal languages (e.g. Hale et al. 1992).

In reaction, institutions such as UNESCO have put forward a series of recommendations for the conservation and preservation of cultural knowledge through the archiving of oral materials.

The *Innu Language Documentation Project* is a step forward in the long term preservation of the Innu language. It aims at sustaining Innu knowledge about the culture, traditions, and language.

Context

Innu is an endangered Algonquian language spoken by approximately 13000 people in 11 communities spread out over Northeastern Quebec and Labrador in Canada.

It has been the subject of intensive grammatical and sociolinguistic analysis by Lynn Drapeau over the past thirty years (e.g. Drapeau, In press). Despite being one of the most studied languages of Quebec, there is no suitable linguistic documentation of the language to insure its lasting conservation and preservation (see Lambert-Brétière, 2011).

Objectives

The objectives of the *Innu Language Documentation Project* are:

❏ to document and describe the Innu language;

❏ to insure its long term conservation by providing access to a large corpus of transcribed and annotated oral recordings in formats exploitable by both the speakers and specialists;

❏ to support the speakers of Innu in their efforts and desire to maintain and promote their language.
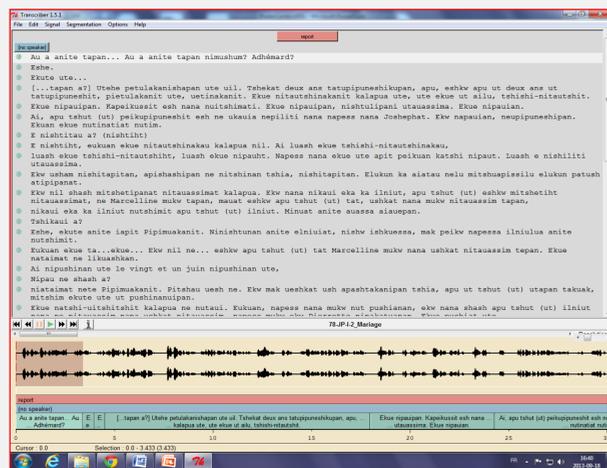
Methodology

PHASE 1 of the project is centered on developing tools for digitalizing, accessing, analyzing and archiving the corpus of oral recordings.

Stage 1: Digitization

The digitization of data is the transfer of all the linguistic material onto digital medium. The initial corpus comprises narratives by monolingual speakers recorded by Lynn Drapeau in the community of Pessamit in the 1970s and the 1980s. The original recordings were on audio cassettes and open-reel tapes, and are now converted to digital format. The Innu audio collection includes over 100 CDs, and the collection is expected to grow as more Innu audio recordings will be gathered from other locations, such as the Tshakapesh Institute.

Stage 2: Transcription

Audio files are processed with *Transcriber* which allows for a time-synchronized transcription. The transcription format used is the Innu standard orthography. This permit accessibility to speakers of all varieties of Innu.
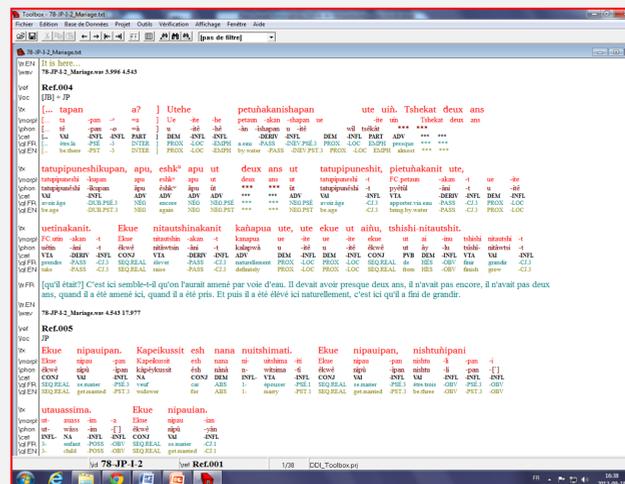
Transcription with
Transcriber

Methodology

Stage 3: Interlinearization

The parsing and interlinearization of texts with detailed grammatical information is done using *Toolbox*, a software designed by SIL International for the management and analysis of language data.

On the whole, the Innu digital corpus includes audio materials provided with time-aligned transcriptions in standard orthography, multi-tier annotations, morpheme-to-morpheme glossing, and free translations in French and English.

**Interlinearization with**
Toolbox

Stage 4: Archiving

All documentation materials, including sound files and recordings, transcriptions, linguistic analyses, as well as translations, are organized using *Arbil*, which associates metadata to the data into a format appropriate for archiving. This makes the documentation all the more valuable to linguists, including those interested in the study of language variation and language death, and to the communities interested in promoting language retention and revitalisation.

Partnership

The *Project* was initially set up in partnership with the Tshakapesh Institute, but participants from other linguistic communities of Quebec (Cree, Naskapi, Atikamekw) were later included.

The project has a flexible structure that insures fluid collaboration between linguists and the community. The 'scientific' team is mainly responsible for the linguistic analyses of the data, and for the production of training manuals for the benefit of the community, as well as for designing community programs to enhance Native capacity at language documentation. The second group acts as a 'steering' team, and its role is to evaluate the needs of the community, recommend specific formats for the outputs, and assess different scenarios for archiving the material.
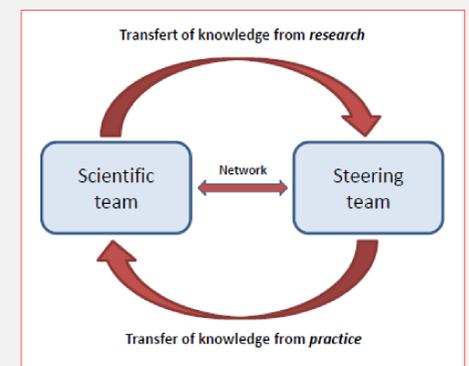
Conclusion

By providing a corpus of high quality original materials, the *Innu language documentation project* aims to make an important contribution to the ongoing developments in language documentation research and a major step in building valuable tools for language maintenance and revitalization. The most important outcome will be a linguistic archive sustaining Innu knowledge about the culture, traditions, and language.

References

Drapeau, L. (In press). *La grammaire de la langue innue*. Québec: Presses de l'Université du Québec.

Hale, K., et al. (1992). Endangered languages. *Language*, 68(1), 1-42.

Lambert-Brétière, R. (2011). La documentation linguistique au service de la préservation et de la conservation des langues autochtones du Québec. In L. Drapeau (Ed.), *Les langues autochtones du Québec, un patrimoine en danger* (pp. 195-218). Québec: Presses de l'Université du Québec.