Transcribing nłeʔkepmxcín stories and conversations

Overview: This talk reports on an in-progress project to transcribe stories and conversations in nłeʔkepmxcín (Interior Salish; British Columbia). We discuss challenges identified by both academic and community partners, and outline paths to solutions.

Context: Prior to our project there was only one published volume of nłeʔkepmxcín stories (Egesdal et al. 2011), plus a few other isolated texts. Conversational data is even rarer; Egesdal et al. (2011) contains a partial transcription of one informal conversation, and Koch (2008) presents some short prompted conversations.

Team and workflow: Our project is a collaboration between nłeʔkepmxcín fluent speakers, linguistics faculty and students, participants in a nłeʔkepmxcín postsecondary program, and a community cultural organization. The entire team meets regularly on Zoom and several times a year in person.

Data are recorded over Zoom or in person. Recorded conversations are time-stamped and transcribed using the orthography of Thompson & Thompson (1996), which is also used in postsecondary nłeʔkepmxcín programs. Conversations are then played back to consultants, who provide translations of their own speech into English.

Results and plans: Our team has recorded more than 21 hours of spontaneous conversational data since August 2023, of which 13 hours have already been transcribed. We have also recorded and transcribed more than an hour's worth of stories. We will make stories and conversations accessible to community members in print form and electronically.

Challenges: There are several technological challenges in this work, both audio-related and font-related. Here we focus on orthographical challenges, which we deem the most important both linguistically and for our community partners. Orthographical challenges arise in particular because there is a lack of standardization in many aspects of the nłeʔkepmxcín writing system.The near-complete absence of prior written connected speech in the language means that many of the questions our work raises have never been explicitly tackled before.

The first major challenge is that there are several different orthographies in use in nłeʔkepmxcín territory. The orthography we use is the one used in linguistic documentation and has the support of key community language experts, but inevitably some community members cannot so far make use of our results. Other issues include: (i) how to represent rhetorical lengthening; (ii) how to indicate intonation or pitch, which speakers express is important; (iii) word segmentation (e.g., whether to attach clitics to their hosts); (iv) punctuation; the community prefers not to use punctuation, so how do we represent pauses or self-corrections, or distinguish direct from indirect quotation? (v) rendition of vowel quality (phonetic vs. 'dictionary-accurate' transcription); (vi) whether to reflect dialectal variation in pronunciation; (vii) how to deal with segments that are deleted in fast speech.

Paths to solutions: Our approach involves constant consultation with fluent speakers and other community members involved in language work. Academic and community partners brainstorm decisions at bi-weekly meetings. As a result of these sessions, we have for example ceased using punctuation and capitalization. We have decided not to represent fine phonetic variation but to stick closely to dictionary spellings unless the pronunciation is markedly different. Decisions we make will likely impact future documentation of the language, but ultimately the community as a whole will decide how their language will be recorded in written form.

References

Egesdal, Steven, M. Terry Thompson and Naz'inek Mandy Jimmie. 2011. *Nłeʔkepmxcín Thompson River Salish Speech.*

Koch, Karsten. 2008. Intonation and focus in Nłeʔkepmxcín (Thompson River Salish). PhD dissertation, University of British Columbia.

Thompson, Laurence C. and M. Terry Thompson. 1996. *Thompson River Salish Dictionary: nłeʔkepmxcín.* Missoula, MT: UMOPL.