Language documentation in Roam Research: Recognizing relationships in field notes

dflynn@ucalgary.ca
kody.tufts@ucalgary.ca

We report on our multi-team experience using Roam Research (https://roamresearch.com/) to document Dɛnɛsu̧łiné (chp). Roam is a note-taking application designed to help researchers establish and discover relationships between their notes, with rather minimal effort and friction. Roam's main attractions are *bidirectional hypertext* and *transclusion* — notions coined decades ago by information technology pioneer Ted Nelson, but never fully implemented until now (Nelson 1999, Appleton 2020).

For our purposes, bidirectional hypertext means that if we enter an utterance like the following (along with its recording) into a daily note, Roam is able to link its lexemes, stems, affixes, glosses, and metadata points, as well as any topics or concepts tagged in the surrounding discussion, to all other instances of these in our database.

naghi̧łná walíú?

na–Ø–ghɛ–NR–ł–ná	wa–Ø–lí	ʔu
again–3SG.OBJ–IPFV–2SG.SBJ–CAUS–be.alive(IPFV)	OPT–3SG.SBJ–COP	Q

'Are you sure you can heal him/her?'

For instance, clicking on *ghɛ-* displays a list of all field notes that are linked to this prefix form ("linked references") as well as a sublist of field notes that are potentially related to this prefix form ("unlinked references"). This helps us to make observations like the following: The prefix form *ghɛ-* is frequent and is normally used to signal perfectivity, except in progressive forms, where it can also signal imperfectivity. More rarely, *ghɛ-* also appears to serve as a thematic prefix that marks derivational (lexical) aspect. By scanning examples of the latter, and other(s') linked notes, we may decide to re-gloss *ghɛ-* in the above utterance as TH (thematic) and to add Ø- IPFV.

This brings us to transclusion. A lexeme like *naghi̧łná* need only be glossed once in the database because all other instances can be "mirrored." This means that if we revise this lexeme's glosses (as discussed above), our revisions can apply simultaneously to all instances of this lexeme in our database. Roam allows virtually anything to be transcluded in this sense. For instance, several parts of this abstract (including the example above and multiple phrases and sentences in the preceding paragraphs) are actually transclusions from separate research notes in our database. (Roam also allows edits to transclusions which leave the original research notes unchanged.)

In sum, Roam Research supports our documentation of Dɛnɛsu̧łiné by allowing us to connect and build on each other's field notes, analyses, and discussions. Our database will be rather large by 2021 so we plan to demonstrate Roam's use of queries at ICLDC7 and we will also provide a critical assessment of Roam relative to current standards in language documentation (e.g., Rice & Thieberger 2018).

References:

Appleton, Maggie. 2020. The pattern language of Project Xanadu. https://maggieappleton.com/xanadu-patterns (16 June, 2020.)

Nelson, Theodor Holm. 1999. Xanalogical structure, needed now more than ever: parallel documents, deep links to content, deep versioning, and deep re-use. *ACM Computing Surveys* 31(4es). 1–32. https://doi.org/10.1145/345966.346033

Rice, Keren & Nicholas Thieberger. 2018. Tools and technology for language documentation and revitalization. In Kenneth L. Rehg & Lyle Campbell (eds.), *The Oxford handbook of endangered languages*. Oxford: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780190610029.013.13