Nêhiyawi-pîkiskwêwina maskwacîsihk : *Spoken Dictionary of Maskwacîs Cree*

Antti Arppe, Atticus G. Harrigan, Katherine Schmirler, Daniel Dacanay, Rose Makinaw

Dictionaries: Journal of the Dictionary Society of North America, Volume 44, Issue 2, 2023, pp. 127-142 (Article)

Published by Dictionary Society of North America
DOI: *https://doi.org/10.1353/dic.2023.a915068*

nêhiyawi-pîkiskwêwina maskwacîsihk: Spoken Dictionary of Maskwacîs Cree[1]

Antti Arppe
University of Alberta
arppe@ualberta.ca

Atticus G. Harrigan
University of Alberta
galvin@ualberta.ca

Katherine Schmirler
University of Lethbridge
katie.schmirler@uleth.ca

Daniel Dacanay
University of Alberta
dacanay@ualberta.ca

Rose Makinaw
Maskwacîs Education Schools Commission
rosemakinaw@maskwacised.ca

ABSTRACT

This paper details the development of *nêhiyawi-pîkiskwêwina maskwacîsihk: Spoken Dictionary of Maskwacîs Cree* (in progress). Since 2014, this joint project between the Maskwacîs Education and Schools

[1]This article appears in *Indigenous Lexicography*, a special issue of Dictionaries: *The Journal of the Dictionary Society of North America* 44(2) (2023), edited by Christine Schreyer and Mark Turin.

In the print version, all illustrations are rendered in grayscale. Any color illustrations can be found in the open-access online version at Project Muse: http://muse.jhu.edu/resolve/213

Commission (MESC) and the Alberta Language Technology Lab (ALTLab) has sought to record carefully pronounced, isolated spoken audio for the approximately 9,000 entries in the *Maskwacîs Dictionary of Cree Words* (Maskwachees Cultural College 2009), as well as to fill lexical gaps through elicitation, to record example sentences for as many of these entries as possible, and to make these recordings publicly available online. Between 2014 and 2018, approximately 700 hours of audio and close to 120,000 recordings for 20,300 carefully spoken word and phrase types were gathered in elicitation sessions. After extracting and annotating the relevant Cree vocabulary, these audio clips were compiled in a novel, publicly accessible online Speech Database as well as through *itwêwina*, the intelligent bilingual online Cree–English dictionary. The entries in this database are currently in the process of orthographic standardization, gloss standardization, and linguistic analysis. Simultaneously, native speakers of Cree are re-reviewing the database's entries to ensure pronunciation quality and verify definitions where needed. In this paper, we discuss the origins of this project; the original elicitation sessions; the postprocessing, standardization, and validation of the recordings; and means by which these recordings can be publicly accessed online.

Keywords: Plains Cree, nêhiyawêwin, spoken language, Indigenous lexicography, online dictionary

INTRODUCTION

Plains Cree (*nêhiyawêwin*, ISO: crk) is an Algonquian language spoken primarily throughout the Canadian provinces of Alberta and Saskatchewan. Although it is one of the most widely spoken Indigenous languages in North America, Plains Cree still has only a relatively marginal online presence, with few major lexical resources available to the digital public. This paper discusses the creation of one such resource, *nêhiyawi-pîkiskwêwina maskwacîsihk: Spoken Dictionary of Maskwacîs Cree* (in progress), to address this shortfall.

In 2014, Miyo Wahkohtowin Education (MWE: Ermineskin First Nation, Maskwacîs, Alberta), and Alberta Language Technology Lab (ALTLab) at the University of Alberta began a joint project to document the variant of Plains Cree (*nêhiyawêwin*) spoken in Maskwacîs. Since

2016, MWE has been succeeded by the Maskwacîs Education Schools Commission (MESC), an amalgamation of the school boards of all four Cree First Nations in Maskwacîs. In addition, the project integrates this documentation into various digital language tools. These tools largely revolve around the *Maskwacîs Dictionary of Cree Words / Nehiyaw Pîkiskwewinisa* (henceforth *Maskwacîs Dictionary*; Maskwachees Cultural College 2009), an almost 9,000-word Cree-to-English glossary which had been compiled by fluent Cree speakers and Elders in Maskwacîs at the turn of the millennium.

Members of ALTLab were introduced in 2013 to representatives of MWE by Ahmad Jawad, the managing director of Intellimedia, the developers of an existing digital language repository for Plains Cree (named, simply, the *Online Cree Dictionary*), and whose main line of business was in providing education administration software for First Nations school boards in Alberta. They hoped to shift the maintenance and development responsibility to computational linguists. ALTLab, which had been established at the University of Alberta in 2013 to develop language technology for Plains Cree and other Indigenous languages spoken in Western Canada, following the Giellatekno model (Trosterud 2006), happened to be in the right place at the right time. Initial discussions with MWE took place in 2013–2014 and concerned possible avenues of developing language technological tools for revitalization based on ALTLab competences and capabilities, for example, spell-checking, a morphologically enhanced (aka *intelligent*) online dictionary, and linguistically analyzed corpora for Cree (see Trosterud 2006, Arppe et al. 2016). But while MWE was known for embracing technological advances, the primary area where interests intersected was in the documentation of the specific variety of Plains Cree spoken in Maskwacîs. After more than twelve months of discussions, planning, and meetings, the relationship was formalized in a Memorandum of Understanding signed between MWE and ALTLab in 2014 and extended in a Memorandum of Partnership signed between MESC and ALTLab in 2018, as part of the research Partnership *21st Century Tools for Indigenous Languages* (21C: 21c.tools).

Objectives. The original objectives of this project were threefold. Firstly, we wanted to verify and audio-record all the lexical entries in the *Maskwacîs Dictionary of Cree Words* (Maskwachees Cultural College

2009), carefully pronounced by L1 Cree speakers in Maskwacîs. Secondly, we aimed at filling in possible lexical gaps in that dictionary by eliciting novel vocabulary. This process resulted in over ten thousand "new" Cree words, which are at present being subjected to a secondary validation process. Thirdly, we strived to record longer utterances exemplifying each Cree entry in the *Dictionary*—ultimately over two thousand such multiword utterances were recorded. Besides these immediate objectives, we also intended to record the informal discussions, whether in Cree or English, that occurred between the speakers and the elicitors during the process of elicitation. Crucially, all the recorded materials, in particular the carefully pronounced Cree words and phrases, would be transcribed according to Standard Roman Orthography (SRO) (Okimāsis and Wolvengrey 2008)[2] and translated into English to create a new text collection of Plains Cree.

The contents of this collection were intended to be made publicly available as part of a web-based electronic bilingual Cree–English dictionary, *itwêwina* (https://itwewina.altlab.app). ALTLab had been developing *itwêwina* as a parallel, complementary project to replace the pre-existing online dictionary, which MWE and MESC wanted to keep up to date. An online dictionary can be broadly and freely accessible for anyone with an Internet connection, making the lexicographical content available to all Cree language learners, instructors, and others, wherever they might be. Furthermore, it is much more straightforward to update and expand a web-based dictionary, in terms of content and new features, in comparison to a printed dictionary (cf., e.g., Garrett 2018). Among the online dictionary's advanced features was to integrate computational modeling of Cree word structure (Snoek et al. 2014, Harrigan et al. 2017, based on the aggregate of the lexical entries in the *Maskwacîs Dictionary* and the extensive lexical database underlying the largest Cree–English dictionary, *Cree : Words / nêhiyawêwin : itwêwina*, by Wolvengrey 2011, with currently over 25,000 entries) in order to recognize morphologically complex wordforms as well as generate extensive inflectional paradigms (Arppe et al. 2022).

[2]While we use the SRO as consistently as possible within this paper, we stay faithful to the orthography used within works we refer to, and so some variation is apparent.

ALTLab also aimed to use the recordings to train a speech-synthesis model to create spoken forms for Cree words and sentences that would be impractical to record with native speakers, such as rarer individual wordforms in full inflectional paradigms (with potentially thousands of forms), or Cree language regular news bulletins or public health circulars, or to create lessons with spoken content in an intelligent computer-aided language learning (ICALL) application (Bontogon 2016). Such a model was indeed created (Harrigan et al. 2019); however, its use has been thus far limited to the creation of automatically generated spoken paradigms for noun and verb entries in *itwêwina*.

PREPARATION, ELICITATION, AND DOCUMENTATION PHASE

In this section we describe the process of preparing the dictionary content for elicitation, as well as the process of elicitation and documentation. We then offer a few observations about the processes.

Semantic classification of the content of the Maskwacîs Dictionary. In preparation for the elicitation sessions with the speakers, the entries of the *Maskwacîs Dictionary* were grouped into sets of semantic domains, based on those used in the *Rapid Word Collection Methodology* (or *RapidWords*) (Moe 2003), a vocabulary elicitation tool developed by the Summer Institute of Linguistics for rapid, intensive dictionary construction (Reule 2018). This semantic classification and organization work was undertaken by three undergraduate volunteers, Megan Bontogon, Elizabeth Pankratz, and the late Sarah Lamarche, during the summer of 2014.

Based on these semantic classifications, we programmatically assigned relevant entries in the *Maskwacîs Dictionary* to each of the elicitation questions listed in *RapidWords*, in effect providing us with elicitation sheets for each recording session which were partially pre-populated with relevant Cree answers to each semantic subdomain. The intention in grouping semantically similar words together and presenting these at the top of the elicitation sheet was to highlight possible lexical gaps and facilitate the recording of new/missing words inspired by exposure to related words in the same semantic domain. This approach served its purpose, with many thousands of new words, as well as several thousand sentences, being recorded during the project (Reule 2018).

Organization of the elicitation sessions. The elicitation and recording sessions involved two to four fluent native Cree speakers from Maskwacîs, who were recruited by MWE and MESC, and initially one discussion facilitator/elicitor and one technician/recordkeeper from ALTLab. Relatively early on the roles of the facilitator and technician became merged, as a single person familiar with the elicitation process could shift their attention between writing down new words and example sentences and monitoring recording quality. The elicitation session was organized under a succession of semantic subdomains in the following manner. First, the elicitor encouraged informal discussion concerning the meaning group, addressing whether all the relevant ideas and words were already represented in the subdomain, or whether some concepts and words were missing. This was followed by a gradual shift towards coming up with sentences that present natural uses of one or more of the words under discussion. After this, we proceeded to record individual words as well as example sentences, spoken carefully in isolation by the Cree native speakers, twice by each speaker. Finally, we advanced to the next semantic subdomain. Reule (2018) created detailed annotations for the various types of speech acts that occurred during these stages, as well as others that occurred during the sessions and often served to facilitate the process. These included: complaints (often humorous, though not always) that could cue the elicitor to move on from a topic before it became arduous; softening comments, such as the elicitor sharing that they themselves hardly know how to explain the English prompt words and thus they are difficult to translate; and jokes, which could be used to reunite speakers and elicitors towards their goals.

Altogether, between 2014 and 2018, we conducted 317 recording sessions, each lasting approximately two hours. At times, we considered whether this recording and elicitation process could have been undertaken in a faster and more efficient manner, but the structure and progress of the sessions were developed together with the speakers and thus took a form that they were agreeable to and glad to continue participating in. On the MESC side, as many as thirty-six speakers participated in the sessions, some in as many as over a hundred sessions and some in only a few. On the ALTLab side, this elicitation phase involved some twenty academics, from undergraduate students to faculty members.

Taking into account the already existing words in the *Maskwacîs Dictionary* and the newly elicited words and example sentences (amounting to altogether 20,300 word and sentence types), this amounted to just shy of 120,000 recorded tokens (each word or sentence type spoken usually twice by on average three speakers, i.e., six times in all).

Observations. The process of recording the L1 speakers, although simple in principle, was in practice considerably more complex than initially anticipated. This process ostensibly required only the review of previously documented Cree words, the elicitation of as-yet-unrecorded vocabulary in each semantic domain, and the creation of example sentences for each entry. One complication was that recordings were performed on location in Maskwacîs, rather than in a laboratory setting, as necessitated by the need for multiple speakers in any given session (in order to record pronunciations by both female and male speakers as well as allow for examples of possible variation), as well as the advanced age of many of the consulting L1 speakers. As such, ambient noise, conversational interruptions, and microphone distortions were relatively commonplace. We addressed this by eliciting considerable repetition during and between sessions, and we expected to record core lexical items in more than one recording session.

In addition to these practical challenges in recording, the process of elicitation itself also appeared to pose at times some degree of confusion or even frustration to our L1 consultants. Many of the elicitation questions in *RapidWords* concerned semantic domains deemed by speakers to be irrelevant to (traditional) Cree culture on the North American plains (e.g., the domain *6.6.2.9 Working with chemicals*), and many questions appeared to be intended to elicit large sets of lexicalized synonyms, such as are found in English, with its synonym sets of corresponding Germanic and Greco-Latinate vocabulary (e.g., *begin*, *start*, *commence*, and so forth). Such degrees of synonymy are often absent in polysynthetic languages such as Cree, which are more fluid in their means of word formation. Eventually we started pruning such overly repetitive or irrelevant categories from the elicitation prompt sheets and would skip such prompts when the speakers showed signs of irritation, thus following a practice which has since been recommended by Boerger and Stutzman (2018).

Similarly, many of the *RapidWords* prompt questions had no corresponding (lexicalized) Cree word, leading to either no Cree translation being provided by the consultants, or descriptive sentences being provided instead, as in (1). We have not yet determined how best to include these examples in the online dictionary.

(1) prompt: atheist
descriptive sentence: *namôya ê-omanitômit*
literal translation: 's/he doesn't have a god, s/he doesn't believe in God'

Another challenge in the elicitation process was that verbal inflections that the speakers preferred to provide as citation forms for new entries did not consistently match with those already in the original Maskwacîs dictionary. In Cree, aside from imperatives, verbs may be inflected for person and number using one of two distinct sets of agreement affixes, termed Independent and Conjunct. Broadly speaking, affixes from the Independent set are used when the verb is the head of a main clause, and affixes from the Conjunct set are used when the verb is the head of a subordinate clause. However, verbs may also take Conjunct affixation in a main clause when the action or state described by the verb is either not bound to present moment in time or not firsthand knowledge of the speaker—this additional function of the Conjunct is likely a vestigial trace of a now-defunct evidentiality paradigm. The entries in the *Maskwacîs Dictionary* were provided in their Independent forms (the conventional dictionary lemma forms for Cree verbs), but when responding to elicitation questions, the L1 speakers typically had a preference for providing these verb forms with Conjunct morphology, for example, *ê-nipât* for 's/he sleeps', likely because the states and actions described were not "real" in the context of the elicitation session, instead being elicited hypotheticals. At times, the speakers also expressed that the corresponding Conjunct forms would be better matches for the English translations included in the original Maskwacîs dictionary, in comparison to the Independent forms that had been used. Nevertheless, when queried whether one could also use the Independent form, for example, *nipâw* for the same 's/he sleeps', the speakers usually agreed, though often noting that the action in question would have just ended. As such, the context-impoverished nature of the elicitation session as a theater of language use must be taken into account when examining the vocabulary items elicited through these means.

POSTPROCESSING, STANDARDIZATION, AND VALIDATION PHASE

In this section we describe the postprocessing of the recorded data, standardization to the SRO, and validation with fluent speakers.

Postprocessing the recordings. A substantial amount of work yet remained to turn these recordings into the intended publicly accessible resource. To start, the interspersing of informal discussion with careful pronunciations necessitated postprocessing, so that the targeted words and sentences would be identified, delineated, and matched with the Cree transcriptions and English translations documented within elicitation sheets. This was undertaken with the ELAN audio annotation tool (ELAN 2023) by over ten undergraduate linguistics students during 2019–2021.

After this process, these extracted and annotated recordings needed to be made easily accessible for 1) orthographic standardization and linguistic analysis by linguists, and 2) content-wise validation by Cree speakers. This was happening after the COVID-19 pandemic had started and therefore ruled out in-person meetings with the mostly elderly Cree speakers. We therefore employed a professional software developer to create a web interface to facilitate this process. In the end, this became a multipurpose Speech Database platform (Speech-DB: https://speech-db.altlab.app/, Poulin et al. 2023). This online application supported not only the validation tasks but also general access and search of its contents by, for example, language instructors and learners—indeed, this platform has turned out to be easily extendable to other Cree dialects and other Indigenous languages. We are still planning with MESC how exactly to publicly announce and raise awareness about this resource, once completed. Nevertheless, all the recordings are already generally accessible, and can be searched and listened to, through Speech-DB platform. In addition, the recordings are also accessible via the *itwêwina* online dictionary, when the recordings can be matched with the latter resource's lexical content, whether they be entry words or phrases, or their inflected forms in the paradigms. For now, all the recordings are backed up to long-term tape storage provided by Digital Research Alliance of Canada (formerly Compute Canada).

Standardizing the transcriptions and translations and validating the content and recordings. The linguistic standardization task was started by several ALTLab linguists in 2021 and has hitherto covered about four-fifths of the content, whereas the content validation tasks, involving at least one fluent Cree speaker (who is also a MESC employee), has reached just over halfway. Both tasks take place within the Speech Database.

The standardization process makes use of a spellchecker version of our Cree computational morphological model as well as the original recordings and aims to identify laxness in vowel length marking (using diacritics) or obvious errors in the original transcriptions. Such errors are not unusual, since transcriptions were done on the fly during the sessions by elicitors who were not fluent in the language and thus could not easily, for example, fill in segments that were reduced or deleted in fluent speech or recognize deep orthographic conventions. Where a word cannot be easily linked to an existing dictionary entry, the translation is used together with knowledge of Cree (especially derivational) morphology to ensure that the semantics and morphosyntax are consistent. More difficult cases involve errors in the original transcriptions, such as incorrect word boundaries, so the recordings and semantics are used to correct these. As for any project, the process has gotten faster with time and practice—we have become more familiar with the less frequent morphological patterns or less well-documented phonological variation.

The content validation, as it has been undertaken during COVID-19, has taken place in Zoom meetings, usually three per week, starting in April 2021. These meetings involve ALTLab linguists and one or two fluent Cree speakers who were also involved to at least some extent in the original recording sessions. Speech Database entries are checked by the speakers for accuracy in the translations of the transcriptions and the accurate matching of the transcriptions with the recordings, as well as the quality of the recordings. If the recordings are inaccurate or of poor recording quality, they can be re-recorded in real time in the database application. Figure 1 shows the *validation* (aka "language expert") view for entries, where (once logged in) the "Good" and "Bad" buttons can be used to mark the recordings for quality. "Wrong Speaker" and

"Wrong Word" are used when the automatic assignment of lemma and speaker codes did not apply correctly. New recordings can also be added from this main page. "More Options" brings us to the *linguist* view (also requiring logging in), where transcriptions, translations, and the analysis (following the morphological model) can be directly edited, as in Figure 2.

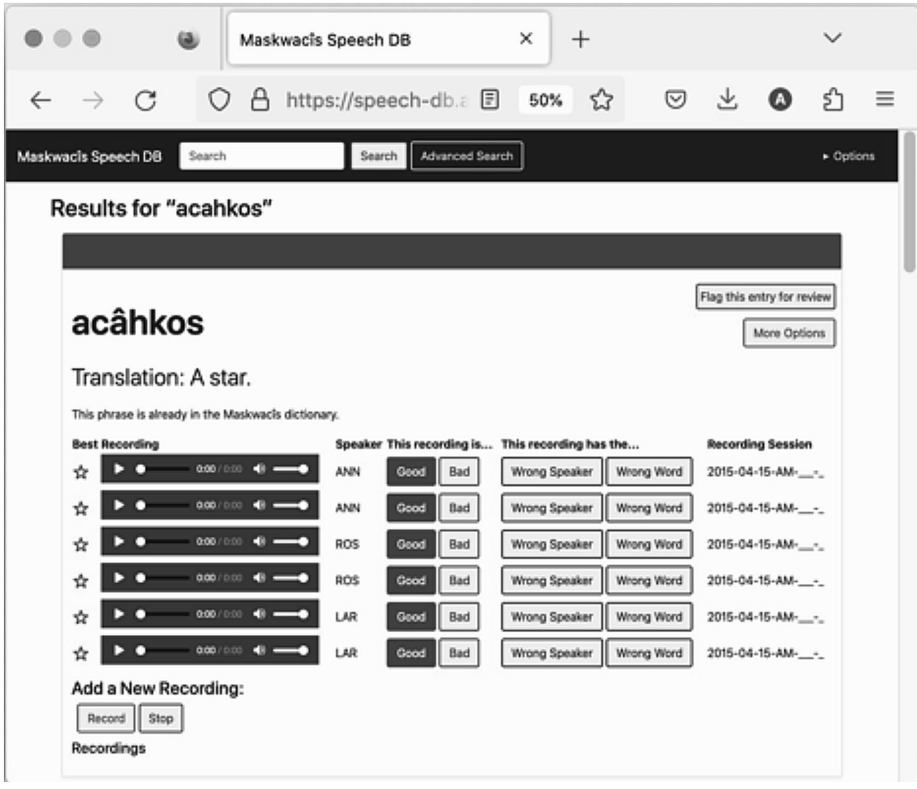

FIGURE 1 Validation ("language expert") view for the entry **acâhkos** in the Speech Database

Where a match between the database and existing dictionary entries can be established, recordings are also presented in *itwêwina*, the online dictionary, as in Figure 3. ALTLab has received substantial feedback from MESC instructors and other employees concerning this application, which has significantly influenced its design principles (Arppe et al. 2022). The primary guideline has been to keep the *itwêwina* user interface as simple as possible, so as to not overwhelm linguistically

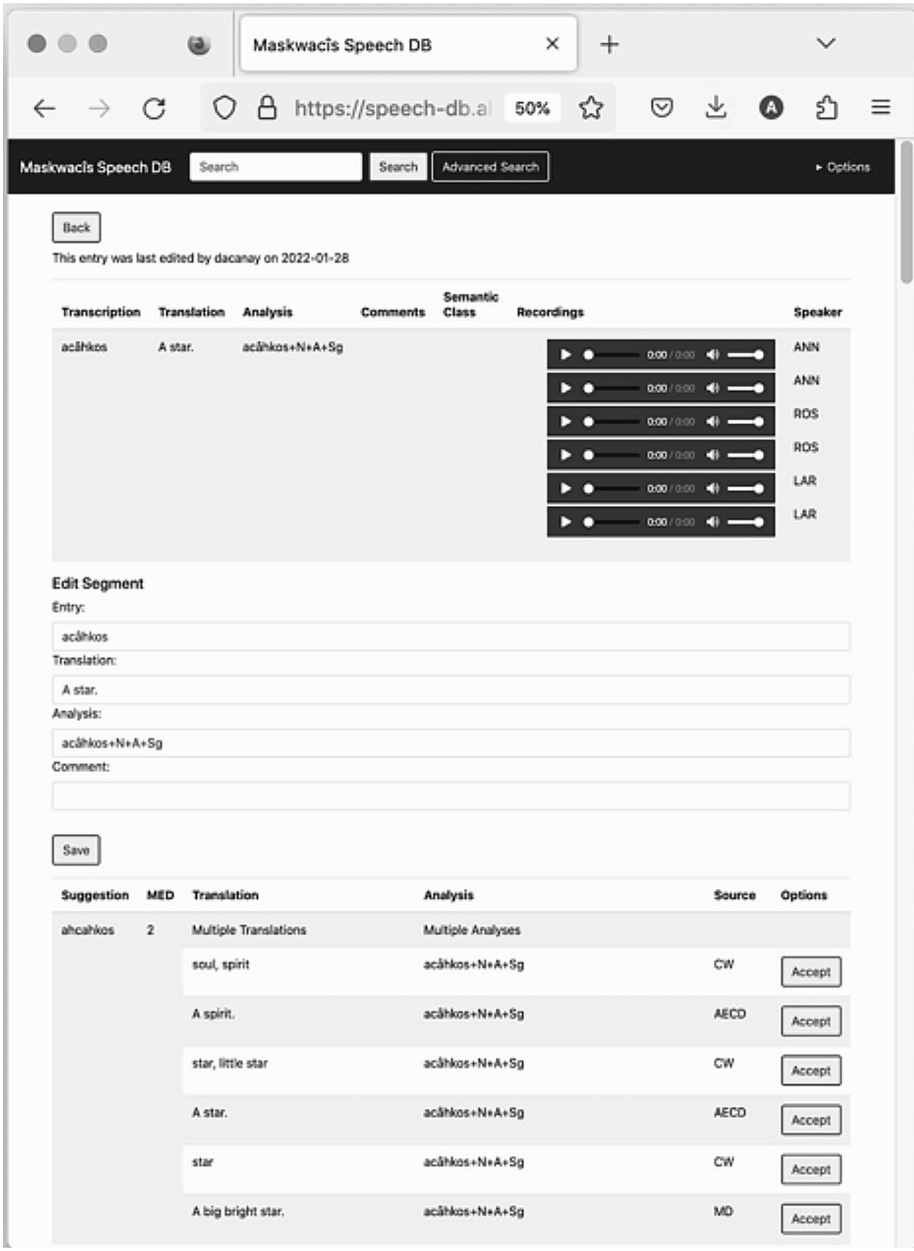

FIGURE 2 Linguist view for the entry **acâhkos** in the Speech Database

untrained users, but to still make it extendible, so that all the information that a linguist might wish to have is easily available. Thus, only essential, core information about a dictionary entry is shown first, but all lexicographic details are accessible with only a single click, as

a pop-up or hover window. We present all the linguistic information by default using plain, non-academic English, though the corresponding linguistic terms can be accessed via pop-ups. Furthermore, the user need not indicate whether they are searching with Cree or English terms; the underlying backend figures this out with some heuristics.

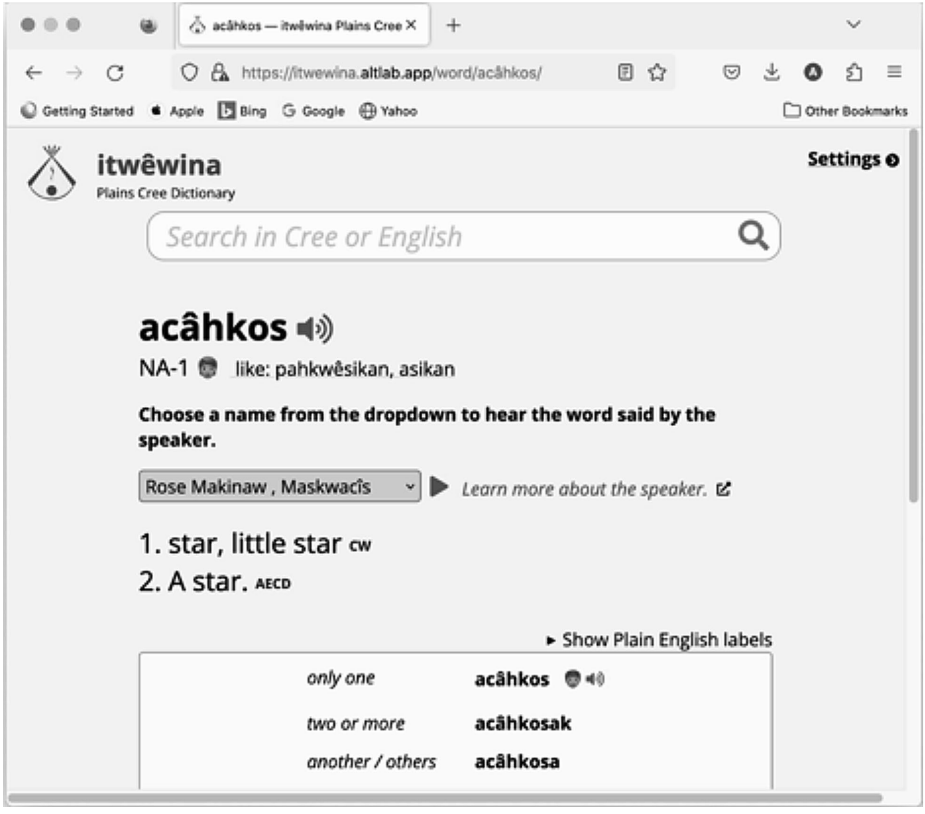

FIGURE 3 Recordings on the word details page for **acâhkos** in the online dictionary *itwêwina*

LESSONS LEARNED AND CONCLUSIONS

Like for many language documentation, stabilization, or revitalization projects, especially when working with Indigenous or other minority languages, we would be remiss to not stress the value of developing long-term collaborative relationships with speech communities. The relationships between the elicitors and community members were a considerable factor in the success of the project. When everyone could

enjoy the sessions and have fun talking about and sharing language, the tasks seemed less tedious. Furthermore, the long-term nature of the project allowed our academic elicitors consistent exposure to Cree from our collaborators, which created a much more effective environment for language learning than a semester in a classroom.

While the completion of this project has not yet been reached, the end is finally within sight within a year or so, and certain general experiences can be conveyed. Importantly, language community expectations can be far less technological than what the academics are capable of or inclined towards. Nevertheless, starting by addressing such community needs can and will result in novel linguistic data that will eventually allow for academically geared research and development, though not necessarily of the sort that was originally planned and expected. That has certainly been the situation in our case.

ACKNOWLEDGMENTS

This work has been funded jointly by Miyo Wahkotowin Education and Maskwacîs Education Schools Commission as well as by a Partnership Development Grant (#611-2016-0207), a Connections Outreach Grant (#890-2013-0047), and a Partnership Grant (#895-2019-1012), as well as personal doctoral fellowships (Harrigan, Schmirler) from the Social Sciences and Humanities Research Council (SSHRC) of Canada and a KIAS Research Cluster Grant from the University of Alberta.

REFERENCES

Arppe, Antti, Jordan Lachler, Trond Trosterud, Lene Antonsen, and Sjur N. Moshagen 2016. Basic language resource kits for endangered languages: A case study of Plains Cree. In *CCURL 2016: Collaboration and Computing for Under-Resourced Languages: Towards an Alliance for Digital Language Diversity* (LREC 2016 Workshop), Portoroz, Slovenia: European Language Resource Association. http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-CCURL2016_Proceedings.pdf (accessed October 16, 2023).

Arppe, Antti, Jolene Poulin, Eddie Antonio Santos, Andrew Neitsch, Atticus Harrigan, Katherine Schmirler, Daniel Hieber, Ansh Dubey, and Arok Wolvengrey. 2022. Towards a morphologically intelligent and user-friendly

on-line dictionary of Plains Cree—next next round. Presentation at the 54th Algonquian Conference, University of Colorado, Boulder, October 20–23, 2022. https://altlab.ualberta.ca/wp-content/uploads/2023/03/itwewina_AC54_Oct2022.pptx.pdf (accessed October 16, 2023).

Boerger, Brenda H., and Verna Stutzman. 2018. Single-event Rapid Word Collection workshops: Efficient, effective, empowering. *Language Documentation & Conservation* 12: 147–93.

Bontogon, Megan A. 2016. Evaluating *nêhiyawêtân*: A computer assisted language learning (CALL) application for Plains Cree. Honors thesis, University of Alberta. https://doi.org/10.7939/R3VD6P81C.

ELAN (Version 6.6) [Computer software]. (2023). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. https://archive.mpi.nl/tla/elan (accessed October 16, 2023).

Garrett, Andrew. 2018. Online dictionaries for language revitalization. In *The Routledge Handbook of Language Revitalization*, edited by Leanne Hinton, Leena Huss, and Gerald Roche, 197–206. New York: Routledge.

Harrigan, Atticus, Antti Arppe, and Timothy Mills. 2019. A preliminary Plains Cree speech synthesizer. In *Proceedings of the 3rd Workshop on the Use of Computational Methods in the Study of Endangered Languages Volume 1*, 64–73, Honolulu: Association for Computational Linguistics. https://aclanthology.org/W19-6009 (accessed October 16, 2023).

Harrigan, Atticus, Katherine Schmirler, Antti Arppe, Lene Antonsen, Sjur N. Moshagen, Trond Trosterud, and Arok Wolvengrey. 2017. Learning from the computational modeling of Plains Cree verbs. *Morphology* 27(4): 565–98.

Maskwachees Cultural College. 2009. *Maskwacîs Dictionary of Cree Words / Nehiyaw Pîkiskwewinisa*. Maskwacîs, Alberta. https://www.altlab.dev/maskwacis/dictionary.html (accessed October 16, 2023).

Moe, Ronald. 2003. Compiling dictionaries using semantic domains. *Lexikos* 13: 215–23.

nêhiyawi-pîkiskwêwina maskwacîsihk: Spoken Dictionary of Maskwacîs Cree. Spoken by Cree speakers and Elders in Maskwacîs: Mary Jean Littlechild, Louise Wildcat, Jerry Roasting, Harley Simon, Annette Lee, Arlene Makinaw, Rosie Roan, Rose Makinaw, Kîsikâw Kiseyiniw, Betty Simon, Brian Lightning, Brian Lee, Linda Oldpan, Mariam Buffalo, Debora Young, Ivy Raine, Agnes Fraynn, Paula Mackinaw, Norma Linda Saddleback, Jack Mackinaw, Renee Makinaw, Linda Whitebear, Linda Wolfe, Ronald Littlechild, Lawrence Wildcat, and Marvin Littlechild; under the management of Rose Makinaw, Patricia

Johnson and Brian Wildcat. Edited by Antti Arppe, Atticus Harrigan, Katherine Schmirler, Daniel Dacanay, Dustin Bowers, Megan Bontogon, Lex Giesbrecht, Jolene Poulin, Timothy Mills, and Jordan Lachler, with the help of undergraduate and graduate research assistants and volunteers at the University of Alberta. https://speech-db.altlab.app/maskwacis/entries

Okimāsis, Jean, and Arok Wolvengrey. 2008. *How to Spell it in Cree: The Standard Roman Orthography*. Regina, SK: Miywâsin Ink.

Poulin, Jolene, Daniel Dacanay, and Antti Arppe. 2023. Speech Database (Speech-DB): An on-line platform for recording, storing, validating, and searching spoken language data. In *Proceedings of the Second Workshop on NLP Applications to Field Linguistics*, 30–39. Dubrovnik: Association for Computational Linguistics. https://aclanthology.org/2023.fieldmatters-1.4 (accessed October 16, 2023).

Reule, Tanzi. 2018. Elicitation and speech acts in the Maskwacîs Spoken Cree Dictionary Project. Honors thesis, University of Alberta. https://doi.org/10.7939/r3-tt3s-4655

Snoek, Conor, Dorothy Thunder, Kaidi Lõo, Antti Arppe, Jordan Lachler, Sjur Moshagen, and Trond Trosterud. 2014. Modeling the noun morphology of Plains Cree. In *ComputEL: Workshop on the Use of Computational Methods in the Study of Endangered Languages, 52nd Annual Meeting of the Association for Computational Linguistics*, 34–42. Baltimore: Association for Computational Linguistics. https://aclanthology.org/W14-2205/

Trosterud, Trond. 2006. Grammatically based language technology for minority languages. In *Lesser-Known Languages of South Asia*, edited by Anju Saxena and Lars Borin, 293–316. The Hague: Mouton de Gruyter.

Wolvengrey, Arok. 2011. *Cree: Words / nêhiyawêwin : itwêwina*. Regina, SK: University of Regina Press.