Speech Database (Speech-DB) – An on-line platform for recording, storing, validating, and searching spoken language data

Jolene Poulin and **Daniel Dacanay** and **Antti Arppe**
Alberta Language Technology Lab (ALTLab)
Department of Linguistics, University of Alberta
{jcpoulin, dacanay, arppe}@ualberta.ca
https://altlab.ualberta.ca

Abstract

The Speech Database (Speech-DB: URL: https://speech-db.altlab.app) is an on-line platform for language documentation, written and spoken language validation, and speech exploration; its code-base is available as open source. In its current state, Speech-DB has expanded to contain content for several Indigenous languages spoken in Western Canada, having started with audio for the dialect of Plains Cree spoken in Maskwacîs, Alberta, Canada. Currently, it is used primarily for validation and storage. It can be accessed by anyone with an internet connection in six levels of access rights. What follows is the rationale for the development of speech-DB, an exploration of its features, and a description of usage scenarios, as well as initial user feedback on the application.

1 Introduction

The Speech Database (Speech-DB: https://speech-db.altlab.app) is an online platform of spoken language data intended for use in the preservation and documentation of less-resourced languages. With dual function as a searchable database for transcribed and translated audio data and as a validation interface for editing spoken dictionary entries, it is available online for anyone to use, and it is easily adaptable for use in various language pairs.

While Speech-DB has been used to store spoken data for multiple Indigenous languages and their dialects spoken in Western Canada, here we exemplify its use in language documentation primarily for Plains Cree (*nêhiyawêwin*, iso: crk), an Algonquian language spoken throughout Western Canada, specifically the dialect spoken in Maskwacîs, Alberta (Canada). This paper will review the objectives of Speech-DB, both at the time it was developed as well as how they evolved, the current and future features of the service, the means by which it was developed, and some key technical features. Furthermore, we discuss how the Speech-DB differs from other, similar services available online, as well as describe the practical usage of Speech-DB through a selection of qualitative user evaluations.

2 Background

The origins of Speech-DB may be traced to an earlier language documentation project; namely, the *Spoken Dictionary of Maskwacîs Cree / nêhiyawêwi-pîkiskwêwina maskwacîsihk* (Littlechild et al., 2018; Arppe et al., 2022a,b). This joint endeavor between Miyo Wahkohtowin Education (now part of the Maskwacîs Education Schools Commission (MESC: https://maskwacised.ca) and the Alberta Language Technology Lab (ALTLab: https://altlab.ualberta.ca), sought to achieve three primary goals: 1) to record audio for all entries ($n = 8996$) in an existing dictionary, the *Maskwacîs Dictionary of Cree Words / Nêhiyaw Pîkiskwêwinisa* (Maskwachees Cultural College, 2009), as spoken by multiple native speakers from Maskwacîs, Alberta (Canada); 2) to fill lexical gaps in the content of this dictionary; and 3) to elicit and record example sentences for as many of these entries as possible (Reule, 2018) This project resulted in the accumulation of 341 approximately 2-hour recording sessions, each of which involved two-to-four fluent native speakers of Cree and at least one linguist. These sessions were recorded at intervals in Maskwacîs between June 2014 and May 2018, and ultimately resulted in the elicitation of 20,299 Cree words and sentences, with anywhere between one and several tens of pronunciation tokens of the same entry by one or more speakers. In 2019-2020, these recording sessions were annotated by undergraduate students to isolate the Cree vocabulary items therein and align them with the transcriptions and English translations provided in the field elicitation sheets.

Proceedings of the Second Workshop on NLP Applications to Field Linguistics, pages 30–39

3 Objectives and their evolution

The original objective for the development of Speech-DB was to construct a centralized database for the Maskwacîs Cree audio entries (and their associated metadata) in a format that was easily accessible from other services, such as *itwêwina* (https://itwewina.altlab.app), an online, morphologically intelligent Plains Cree – English dictionary (Arppe et al., 2018, 2022c). The process of validating the recording quality, Cree transcriptions, English translations, and metadata (e.g. speaker ID codes) of the database's audio recordings was initially planned to take place in-person in Maskwacîs. However, when in-person activities became all but impossible in early 2020, a new approach was needed to enable this validation task to take place virtually. The Speech-DB was subsequently expanded to support this task.

In the planning and organization of the various subtasks within the validation work, we aimed to optimize the impact of, and minimize the time commitment for, our native speaker consultants, who (in Maskwacîs) were predominantly elderly individuals whose time was in high demand for various other language documentation and instruction tasks. Thus, we divided the validation tasks into activities which categorically required the participation of a native speaker of Cree and tasks which could be accomplished by a linguist knowledgeable in the language. Therefore, the native speakers (or 'Language Experts') would be categorically needed for 1) judging the accuracy of English translations for all the Cree entries in the database (and providing corrections to these), 2) judging the accuracy and naturalness of Cree sentences (word choice, word order), as well as 3) judging the quality of each individual spoken token, in particular their exemplariness. Consequently, the supporting linguists could undertake preparatory standardization work, such as 1) reviewing and fixing any apparent inconsistencies in the Cree transcriptions, which is coupled with 2) reviewing the linguistic analyses of the transcriptions (including the lemma, stem, and other lexical information). This workflow is described in detail in Section 5. In addition, given that there were roughly 150,000 individual unvalidated recording tokens at the beginning of the validation process, provisionally made available through *itwêwina*, we also made it possible for any person to flag recordings in Speech-DB that were in any respect problematic. (i.e. poor recording quality, unusual transcription or translation) for review by linguists or language experts.

As the validation process proceeded to take place online (both asynchronously and synchronously through teleconference with Cree elders and other speakers in Maskwacîs), new features were added to ease the workflow, in effect extending the use of Speech-DB to (new) language documentation. One such additional feature was the ability to record entries directly into the Speech-DB. Previously, recordings had to be done on a separate computer or with a separate software, annotated by a linguist to segment relevant snippets from larger recordings, and then uploaded to the Speech-DB using a custom script written by the software developer. With the addition of this feature, any authorized user (see Section 4.1) can add a new recording directly to the Speech-DB, so long as they know the transcription and translation of the entry. These recordings are then subject to review and approval by a linguist prior to being made available to the general public.

The database is also structured in such a way that new language groups can be added with minimal technical effort. All that is required for the addition of a new language on Speech-DB is for the site administrator to enter in the new language family; users can then immediately begin adding and viewing entries. This new language family is then presented on the introductory page as a new section of the Speech-DB. New sections can be instantiated with no recordings as an empty version of the Speech-DB; with recordings supplied by a linguist or community member in a format that can be parsed and uploaded by the software developer.

Alternatively, sections containing only prompts for future recordings can be created. In the case of the last option, these prompts may be taken from handwritten, gestalt lists of entries, or, more effectively, from an existing, codified semantic domain set, such as that used in the SIL Rapid Word Collection Method (Boerger and Stutzman, 2018), which would both provide an overall structure for entries and ensure a relatively balanced coverage of the lexicon. Consequently, besides audio for Plains Cree spoken in Maskwacîs, Speech-DB has expanded to incorporate content for another Cree dialect spoken in *môswacîhk*, Saskatchewan, as well as selected outputs from a Plains Cree synthesizer (Harrigan et al., 2019). Additionally, extensive audio exists for the Dene language Tsuut'ina, imported into Speech-DB, as well as for three areal variants of

the Siouan language Nakoda.

Speech-DB's search functionality was initially very basic, featuring only the option to search for entries matching a search string. However, as needs evolved, an advanced search feature was added, allowing users to search by a variety of attributes, including recording quality ('GOOD' or 'BAD'), speaker, morphological analysis, transcription, translation, and semantic classification. The last of these attributes, semantic classification, is based on the semantic domain assigned to the entry according to the aforementioned SIL Rapid Word Collection Methodology, which was used in the initial recordings sessions to collate similar vocabulary to be covered per each session. However, this semantic classification search functionality is not yet fully operational.

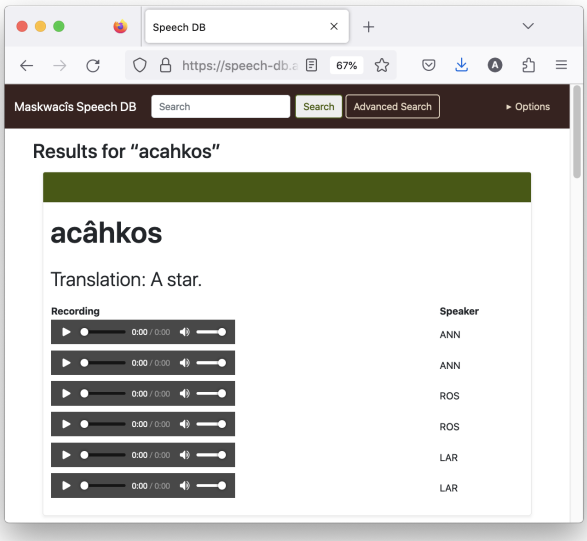

Figure 1: The entry "acâhkos" as viewed by a user who is not logged in.

4 Description of the application

4.1 User types and functionality

The Speech-DB supports six distinct user types, implemented so as to segment permissions and authorizations. The first of these user types are Unauthorized users; that is, users who are not logged in to a Speech-DB account. Such users can see all publicly available language families and can view and listen to all recordings belonging to those families. They have zero permissions to provide feedback or make any changes to the database, and are shown minimal metadata information for each entry (Figure 1).

The second user type is designated as 'Learners'. These users, who must be logged into an account, have access to all features available to unauthorized users, with the addition of being able to flag entries for review. This allows Learner users, who are assumed to be neither fluent speakers nor linguists, to provide feedback on entries without making any direct changes to the database. An internal Issue is created for each flagged entry, storing the feedback from the user. Issues can then be reviewed and addressed by more advanced users. In addition, Learners can record new audio directly into the Speech-DB, subject to review by linguists.

The third user type is that of the 'Instructor'. Currently, Instructors have the same privileges as Learners. In the future, Instructors will receive access to specific layouts and displays intended for instructing the language, such as the option to view entries grouped by lesson type or complexity.

The fourth level of access is the 'Language Expert'. Language Experts, assumed to be native or otherwise fluent speakers, have more available options than Instructors, and are shown additional metadata for each entry (Figure 2). In addition to flagging entries for review and adding new entries to the database, Language Experts can validate the recording quality of existing entries. Validation is done through a series of steps, each involving its own button or pair of buttons on the entry. Firstly, the Language Expert can indicate if the transcription and translation are both spelled correctly and if the meanings are correct through the use of "Yes", "No", and "I don't know" buttons. The last option is provided so as not to oblige users to accept or reject entries with which they are not familiar. This option also informs the site administrators which entries require further review. Next, the Language Expert can listen to each recording for the entry, marking them as "Good" or "Bad" based both on recording quality and quality of pronunciation. These changes are directly reflected in the database. While listening to the recordings, Language Experts can note if the Cree word(s) in the recording do not match the transcription (but are otherwise valid), or if a recording is assigned to the incorrect speaker, using a series of buttons on each entry. These issues are logged as Issue items and can be reviewed by either Language Expert users or linguists.

'Linguists' constitute the fifth user group, and

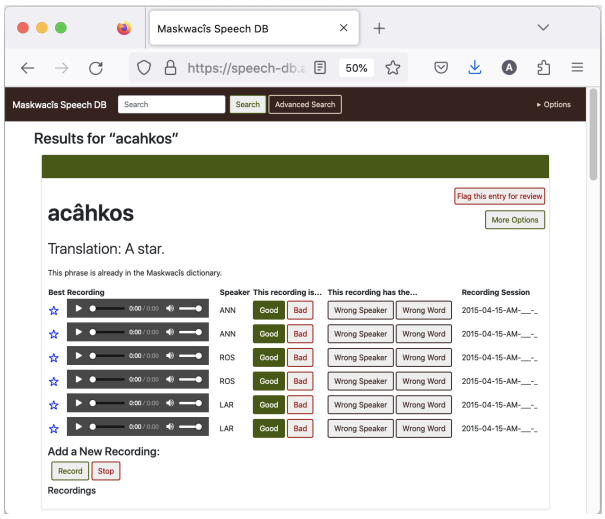

Figure 2: The entry “acâhkos” as viewed by a Language Expert level user.

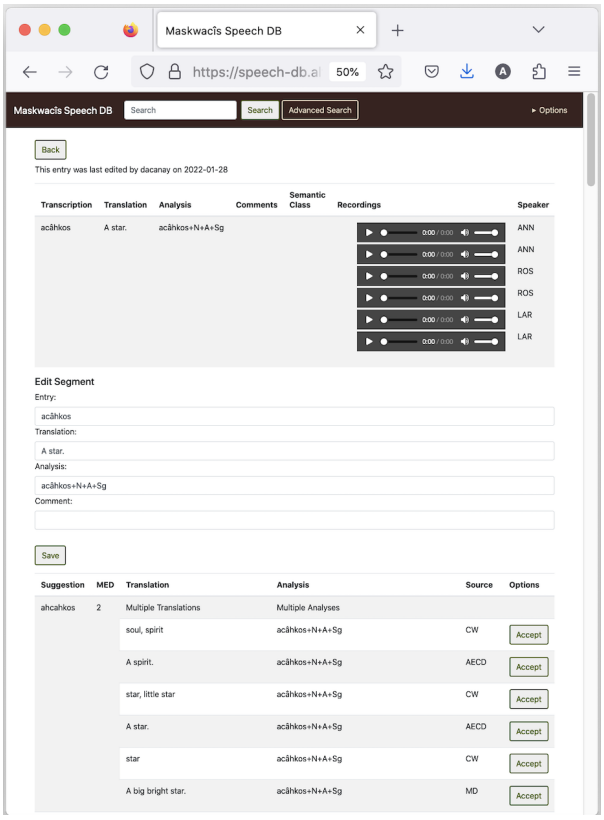

Figure 3: The top section of the “More Options” page, accessible only to Linguist level users, for the entry “acâhkos”. Features the entry metadata and recordings as well as the top items in the suggestions table.

have identical permissions as Language Experts, with the additional option to view and alter interlinear glosses, translations, and transcriptions directly using a button labeled “More Options” (Figure 3). This option allows ‘Linguist’ users to make changes directly to the database. When available, transcription, translation, and analysis suggestions are provided through the use of a finite-state morphological model (for Plains Cree, described in Snoek et al. (2014) and Harrigan et al. (2017)) and dictionary content from the sister application *itwêwina*. Suggestions are ranked by *Modified Edit Distance* (MED), which the service calculates itself. An entry’s MED is the number of changes needed for the suggestion to match the current input. The MED assigns a lesser penalty to some common inconsistencies in the spelling of Plains Cree words that we are aware of; for other spelling divergences the regular edit distance penalty is applied. For example, adding or removing an ‘i’ or an ‘h’ has a distance of 0.5, thus 0.5 is added to the total MED for every ‘i’ or ‘h’ that is added or removed from the original entry in order to match the suggested entry. Adding or removing a diacritic from a character has a cost of 0, whereas adding or removing any letter other than ‘i’ or ‘h’ has a cost of 1. All these changes are calculated and summed up to present the total MED between the current transcription and the suggested spelling. Lastly, this Linguist-specific view contains a table listing all previous changes made to an entry, when those changes were made, and by whom the changes were made (Figure 4). This table can then be used to revert an entry to a previous state in the case that it was incorrectly changed at some point in history.

The sixth and final user type is the ’Administrator’, a role reserved for one or two software-educated users who update the database in the backend and make changes to the service using Django’s Administrator interface. The role has no special privileges on the front-end and has total control over the backend, with the ability to change any and all aspects of any given entry.

As previously mentioned, many user-types have the ability to record new entries directly into the Speech-DB. This can be done either from the entry itself, which then adds a new provisional recording to the database containing the transcription and translation of that entry, or through the page directly intended for recording new entries. In the latter option, the transcription and translation are added in text fields before the user records as many entries as desired, saving only the ones that meet their standards of pronunciation and audio quality. If this user has recorded entries in the past, there will be a “speaker” object associated with the user

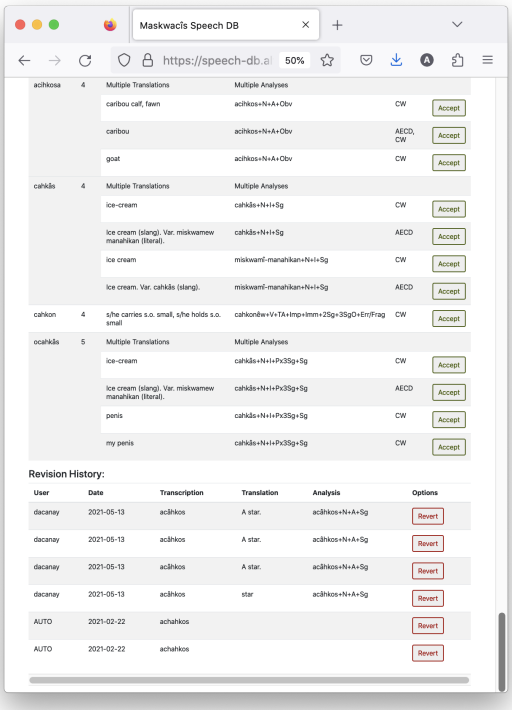

Figure 4: The bottom end of the "More Options" page, accessible only to Linguist level users, for the entry "acâhkos". Features the end of the suggestions table and the revisions table.

and this speaker object will be used. Otherwise, a new speaker object is made using the name and username provided when the account was created.

4.2 Technical specifications

The entire Speech-DB code-base is available as open source in: https://github.com/UAlbertaALTLab/recording-validation-interface. The service is a Django monoserver, in that it uses Python's Django as both the backend and frontend components. The backend handles all the logic of importing new recordings, storing newly-recorded entries, and handling user input. The frontend displays all the information and options to the user using the Django framework and templates. All the information is stored on a server, which serves the site to the public using uwsgi and nginx, and the data themselves are stored in a sqlite3 database. This server is housed on a server provided by Digital Research Alliance of Canada, running Ubuntu and serving the sites to nginx using Docker. Audio information is kept on the server in its original .wav quality format, but it is also converted into .mp4 format at the time it is added to the database as this format is smaller and easier to serve over the web.

The database itself contains seven tables with an additional eight relational tables to store all the information. The seven main tables are for storing Issues, as discussed above, language variants or language families for each new language pair that is supported by the Speech-DB, phrases, recordings, recording sessions, semantic classes, and speakers. Speakers are either users who have recorded an entry, or manually entered names of people who have contributed to the database. Adding a new language pair is as simple as adding a new entry to the language variant table, which takes maximally five minutes. Entries and speakers are then associated with this new language family and only presented to users when viewing the entries for that language family.

When another service, such as *itwêwina*, requests a recording from Speech-DB, it makes a GET request to the bulk recording API built into the back-end of Speech-DB. This API endpoint can accept up to 30 query terms and returns a JSON object containing the terms that were found in the database along with a separate list of terms that were not found in the database. For every entry found in the Speech-DB, a list of the corresponding recordings is returned with the name of the word. When searching for words, each instantiation of *itwêwina* contains the community code, which is found in the URL of any language family's main content (e.g., the code for Maskwacîs is "maskwacis", a URL-safe version of the name), and each of those URLs are queried for the term. In the case of Plains Cree, the Speech-DB needs to account for potential spelling variations, mainly using macrons, <ē>, instead of circumflexes, <ê>, or in some cases neither diacritic, <e>. To accommodate any such spellings, each query is done with each set of characters and then the entry associated with the recording is correctly assigned back to the initial query term by undoing the changes done to the accent marker.

This exchange between Speech-DB and another service allows for the presentation of spoken forms for individual words, or their collections as organized into "spoken paradigms", both types exemplified in Figure 5 (for the entry "nipâw"[1]).

[1] https://itwewina.altlab.app/word/nipÃćw/?paradigm-size=full

Figure 5: The presentation of spoken recordings for both the individual Cree search term “nipâw” and a collection of audio organized into a “spoken paradigm”, fetched from Speech-DB to another on-line application, namely *itwêwina*. The recordings can be played by toggling the speaker icons, which, when paired with a human face icon, indicate a human recording (in contrast to synthesized audio).

4.3 New features

In addition to its current features, the Speech-DB will soon see numerous expansions in the form of new features. The first of such features is the completion of the possibility to search for entries based on their semantic classifications, which is currently only partially implemented. This feature will be expanded to include classifications made using WordNet (Miller et al., 1993; Dacanay et al., 2021) as well as more accurate RapidWords classifications. The ability to include hypernyms and hyponyms for WordNet classifications will also be added to this search functionality.

Next, the ability to start a new database with RapidWords prompts will be added. While this is currently theoretically possible, in order to accomplish it, the software developer must add all these entries by running a script on the database. This new feature would allow users to select a flag when creating a new language family that automatically populates the database with prompts from RapidWords, or indeed any other written and codified semantic domain set. Subsequently, the previously mentioned table of spelling and analysis suggestions accompanied by a MED only works on single word phrases. We would hope to expand this functionality into multi-word phrases and have suggested spellings and analyses for each component of the phrase.

Lastly, users have requested the ability to bulk change a particular speaker code for a particular session. This is necessary due to occasional errors in the initial recording sessions, in which certain speaker codes were erroneously miscoded. When such errors occur, users must change the speaker code of every incorrectly-coded recording individually. However, users will eventually have the ability to change speaker codes in bulk, shifting speaker A for speaker B for an entire session.

5 Work processes by current users

As mentioned, the basic standardization and validation processes for which Speech-DB was designed may be exemplified in the undertaking of these processes for the Maskwacîs Cree audio, which (at the time that it was initially uploaded to Speech-DB) was aligned with transcriptions and translations taken from the field elicitation sheets and notes produced by the linguists who collected the audio. These elicitation notes were organized under semantic groupings following Rapid Words, and combined semantically classified content from the Maskwacîs Cree Dictionary (which did not adhere to Standard Roman Orthography (Okimāsis and Wolvengrey, 2008), as well as "new" Cree words and sentences in response to prompt questions and words throughout the elicitation session. As the multiple linguists recording the sessions were not fluent speakers of Plains Cree, the written records represented their best approximation of the phonological form of what they heard, rather than the orthographically standard form. The resultant transcriptions therefore required comprehensive orthographic standardization. Furthermore, the English translations varied, either following conventions in the *Maskwacîs Cree Dictionary* (Maskwachees Cultural College, 2009) or the larger *Cree: Words* dictionary (Wolvengrey, 2011), or some hybrid of both; these too were to be standardized.

To facilitate this standardization, the Maskwacîs

Cree audio clips were uploaded to the Speech-DB and grouped by the elicitation session in which they were collected (which would concern words mainly from related semantic domains). These sets of recordings were then manually reviewed by a linguist with knowledge of Plains Cree morphosyntax and orthographic conventions (initially the third author, and then primarily the second author). For each entry, the linguist would, using the provided recordings, verify that the word or sentence spoken in the audio was the word or sentence provided in the gloss. The linguist would subsequently standardize the spelling of the Cree words in the 'Transcription' field to SRO conventions, render the definition in the 'Translation' field to a format closely resembling that used in the largest currently existing Plains Cree dictionary (Wolvengrey, 2011), and provide an interlinear gloss detailing the inflectional characteristics of the word(s) present in the 'Analysis' field, making use of the computationally generated suggestions when suitable. A fourth field, the 'Comments' field, was used in instances in which the entry in question was notable or unusual in some respect; typically, in the process of standardization, this was reserved for alternative spellings, derivational breakdowns of semantically non-compositional terms, and morphosyntactic irregularities. However, this 'Comments' field (which was added to the site by the request of linguists working with Speech-DB) was also used as a miscellaneous repository for additional information on entries.

After being manually standardized and interlinearized by a linguist, the quality of the recordings and translations for these entries were also validated by Rose Makinaw, an L1 Cree-speaking elder from Maskwacîs, in collaboration with linguists (second and third author) and the software developer (first author). Across 162 validation sessions, totalling 262 hours, these audio validations have covered 50% of the total contents of the Speech-DB, as well as having provided 500 novel words to the database with multiple recordings of each.

6 Feedback from current users

In total, using Speech-DB as an editing interface, the second author has been able to standardize roughly 63% of the 20,299 entries of Maskwacîs Cree over the course of 21 months of sporadic work. He has noted no significant structural deficits with Speech-DB as a platform (with the exception of occasional server errors), and deemed the general layout as "intuitive" and as "not requiring a great deal of training to use".

Furthermore, the aforementioned native Cree speaker (who has no formal training in linguistics and a self-professed lack of tact in the use of computers and digital interfaces) reported no complaints regarding the practical usage of the site, and commented that she was "comfortable with it" after having been exposed to it for a time. When asked about how she would explain the interface to a new user, she commented that it would be sufficient to have them "sit beside me" during a validation session, and described her own experience of learning to use the site as "not that bad". When asked what skills a potential validation annotator using Speech-DB would need to begin their work, she mentioned only literacy in the Cree Standard Roman Orthography and for the annotator to be "fluent enough to know when ... the speakers [on the database] are saying it wrong"; no mention of specialized computational or linguistic knowledge was mentioned.

The software developer (first author) has participated in a large proportion of the validation sessions, from their beginning in March 2021 until the time of writing, in order to directly observe any erroneous or otherwise undesirable functionality, and consequently to resolve such issues as swiftly as possible. Several of the linguists involved in the initial recordings have also participated, and have consistently judged that the validation and associated standardization activities currently undertaken in Speech-DB are being accomplished as efficiently as can be reasonably expected while still giving each and every recording, transcription, and translation a sufficient amount of attention for proper quality assurance. Indeed, while the very first 10 validation sessions involved a learning process and covered on average 13 entries per hour, at the end of that period the rate had already increased to 25 entries/hour, having now doubled to 60 entries/hour. As for the standardization work, that has always progressed faster than validation, and has now reached a rate of 110 entries/hour.

7 Comparison with other relevant similar applications

Although other applications similar to the Speech-DB exist, none of the ones we are aware of are able to fill all of the aforementioned usage roles.

Feature / Application	Speech-DB	DGD2 (Schmidt, 2014)	Talk-Bank (MacWhinney, 2019)	Library of Congress (1986/2023)
Add new recordings	+	–	+	–
Validate existing recordings	+	–	–	–
Authenticate users	+	+	–	–
Add linguistic analyses to entries	+	–	?	–
Publicly view entries	+	+	+	+
Easily access recordings from other services	+	–	?	?
Search for recordings	+	+	–	+
Intended for language preservation, documentation, and exploration	+	–	–	–

Table 1: A comparison of the Speech-DB with other similar services.

Table 1 shows a comparison of the Speech-DB with three other similar services.

While these other services all existed at the time the Speech-DB was created, they differ in several respects. Foremost among these are the intentions of the service. The Speech-DB was custom designed based on a set of criteria aimed at documenting and preserving the language, which none of the other services have as their aim, nor do they offer some of the key elements the Speech-DB does provide, such as the ability to validate entries and access them from other services on the Internet.

8 Conclusion

The Speech-DB is an online platform for spoken language data available to the public in varying degrees of access, depending on the user's familiarity with the language. It serves as a service for documentation, exploration, and validation, with its functionalities having expanded over time to accommodate user needs. The primary users of Speech-DB regard it as easy to use and generally have no complaints about how it operates. The Speech-DB differs from other similar platforms primarily in its ability to grow and adapt with the language, easily add new language families, and easily add new recordings.

Limitations

Although Speech-DB can be used as a standalone exploratory tool for language learners, using it for extensive, rich documentation (of the kind outlined for Plains Cree) does require some degree of linguistic understanding, in that such an extent of analysis of the data necessitates the establishment or implementation of some form of coding convention for the linguistic features apparent in the entries, and/or the existence of a computational model/parser that can suggest such analyses. As such, although language community members can act largely independently in creating and populating a Speech-DB for their own language, the contribution of linguists may be needed for more advanced linguistic analysis. Furthermore, Speech-DB has been primarily used for analyzing and validating pre-existing recordings, which had been collected and processed separately, rather than solely recording the audio using Speech-DB; instead, Speech-DB was used afterwards for recording individual additional audio, when considered necessary. For more extensive recording projects using solely Speech-DB, the application would yet benefit from stream-lining the recording process to better support the recording of larger batches of vocabulary in a convenient and efficient fashion. Additionally, while Speech-DB provides the framework to allow users to search by categories such

as semantic domain, such categories do require the provision of additional information when entries are initially added or recorded.

Ethics Statement

The collection of audio which is stored and made available in Speech-DB is covered by an ethics review and approval at the University of Alberta (Study ID: Pro00023436). The platform described in this manuscript has been developed in order to support the explicit objectives of the language communities in question to record how their language is spoken in their communities and make that available for their next generations.

Acknowledgements

Creating a useful application that supports the revitalization of Plains Cree (*nêhiyawêwin*) is a task that relies on the knowledge, time, and goodwill of many people. We thank the Social Sciences and Humanities Research Council (SSHRC) for their Partnership Development Grant (#890-2013-0047), Connections Grant (#611-2016-0207), and Partnership Grant (#895-2019-1012) for supporting this project over the last decade, and are equally grateful for the Research Cluster Grant from the Kule Institute for Advanced Study (KIAS), University of Alberta. Earlier stages of the software development were supported by Eddie Antonio Santos and Andrew Neitsch. We would especially like to thank the staff at Miyo Wahkohtowin Education, now part of Maskwacîs Education Schools Community (MESC), for their wonderful enthusiasm, and for welcoming us into their community. In particular we want to note Brian Wildcat, Patricia Johnson, and Rose Makinaw, who were instrumental in ensuring the recording project would be carried out. The actual recordings were faciliated by Atticus Harrigan, Katherine Schmirler, Dustin Bowers, Megan Bontogon, and a number of other students at the University of Alberta. We would also like to acknowledge the crucial advice, attention, and effort of Jean Okimâsis and Arok Wolvengrey from First Nations University of Canada. Last but by no means least, we are indebted to all the Elders and native speakers of Plains Cree in Maskwacîs, enumerated in Littlechild et al., 2018. and in `https://speech-db.altlab.app/maskwacis/speakers/`, for the countless hours of their time that they contributed to creating the content of the Spoken Dictionary of Maskwacîs Cree, which we are today fortunate to be able to make available through the Speech-DB.

References

Antti Arppe, Atticus Harrigan, Katherine Schmirler, and Arok Wolvengrey. 2018. A morphologically intelligent online dictionary for Plains Cree. In *Stabilizing Indigenous Languages Symposium (SILS2018)*, University of Lethbridge, Alberta, June 7-9, 2018.

Antti Arppe, Jolene Poulin, Atticus Harrigan, Katherine Schmirler, Daniel Dacanay, and Rose Makinaw. 2022a. *êkosi ê-nêhiyawi-pîkiskwêcik maskwacîsihk* – Towards a spoken dictionary of Maskwacis Cree. In *Fifty-Fourth Algonquian Conference (PAC54)*, Boulder, Colorado, October 22, 2022.

Antti Arppe, Jolene Poulin, Atticus Harrigan, Katherine Schmirler, Daniel Dacanay, and Rose Makinaw. 2022b. *êkosi ê-nêhiyawi-pîkiskwêcik maskwacîsihk* – Towards a spoken dictionary of Maskwacis Cree. In *Partnerships in Practice, Special Session at the 5th Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL-5)*, Dublin, Ireland, May 27, 2022.

Antti Arppe, Jolene Poulin, Eddie Antonio Santos, Andrew Neitsch, Atticus Harrigan, Katherine Schmirler, Daniel Hieber, Ansh Dubey, and Arok Wolvengrey. 2022c. *itwêwina* – Towards a morphologically intelligent and user-friendly on-line dictionary of Plains Cree – next next round. In *Fifty-Fourth Algonquian Conference (PAC54)*, Boulder, Colorado, October 22, 2022.

Brenda H. Boerger and Verna Stutzman. 2018. Single-event rapid word collection workshops: Efficient, Effective, Empowering. *Language Documentation and Conservation*, 12:147–193.

Daniel Dacanay, Atticus G. Harrigan, and Antti Arppe. 2021. Computational analysis versus human intuition: A critical comparison of vector semantics with manual semantic classification in the context of Plains Cree. In *Proceedings of the 4th Workshop on Computational Methods for Endangered Languages*, pages 33–43.

Atticus Harrigan, Antti Arppe, and Timothy Mills. 2019. Preliminary Plains Cree speech synthesizer. In *Proceedings of the 3rd Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL-3)*, pages 64–73, Stroudsburgh, Pennsylvania. Association of Computational Linguistics.

Atticus Harrigan, 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*, 27:565–598.

Library of Congress. 1986/2023. *American English Dialect Recordings: The Center for Applied Linguistics Collection (AFC 1986/022)*.

Mary Jean Littlechild, Louise Wildcat, Jerry Roasting, Harley Simon, Annette Lee, Arlene Makinaw, Rosie Rowan, Rose Makinaw, Kisikaw, Betty Simon, Brian Lightning, Brian Lee, Linda Oldpan, Miriam Buffalo, Debora Young, Ivy Raine, Paula Mackinaw, Norma Linda Saddleback, Renee Makinaw, Atticus Harrigan, Katherine Schmirler, Dustin Bowers, Megan Bontogon, Sarah Giesbrecht, Patricia Johnson, Timothy Mills, Jordan Lachler, and Antti Arppe. 2018. Towards a spoken dictionary of Maskwacîs Cree. In *25th Stabilizing Indigenous Languages Symposium (SILS2018)*, University of Lethbridge, Alberta, June 7-9, 2018.

Brian MacWhinney. 2019. Understanding spoken language through TalkBank. *Behaviour Research Methods*, 51:1919–1927.

Maskwachees Cultural College. 2009. *Maskwacîs Dictionary of Cree Words / Nêhiyaw Pîkiskwêwinisa*. Maskwacîs, Alberta.

George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller. 1993. Introduction to WordNet: an on-line lexical database. *International Journal of Lexicography*, 3:235–244.

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

Tanzi Reule. 2018. *Elicitation and Speech Acts in the Maskwacîs Spoken Cree Dictionary Project (Honors thesis)*. Department of Linguistics, University of Alberta.

Thomas Schmidt. 2014. The database for spoken German – DGD2. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 1451–1457, Reykjavik, Iceland. European Language Resources Association (ELRA).

Conor Snoek, Dorothy Thunder, Kaidi Lõo, Antti Arppe, Jordan Lachler, Juhani Järvikivi, Timothy Mills, Sjur N. Moshagen, and Trond Trosterud. 2014. Modeling the noun morphology of Plains Cree. In *Proceedings of the 2014 Workshop on the Use of Computational Methods in the Study of Endangered Languages (ComputEL)*, pages 34–42, Baltimore, Maryland, USA. Association of Computational Linguistics.

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