

Revitalising Endangered Languages and Cultural Heritage through Language Technology: A Pilot Study for Dzardzongke

Hannah M. Claus, Songbo Hu, Emre Isik, Anna Korhonen,
Kitty Wenying Liu and Marieke Meelen

University of Cambridge

{hmc78, sh2091, sei31, alk23, wl399, mm986}@cam.ac.uk

Abstract

In this short paper, we present the first prototype of a mobile application to help preserve and revitalise the endangered language and cultural heritage of the speakers of Dzardzongke, a Tibetic language spoken in South Mustang, Nepal. With this pilot study, we provide a collaborative and highly accessible solution to revitalisation that has potential for any community interested in preserving their language and culture.

1 Introduction

Despite an increased awareness of linguistic diversity, dominant language technologies continue to privilege well-resourced languages (Blasi et al., 2022; Joshi et al., 2020), leaving many endangered language communities further marginalised in the digital age (Anastasopoulos et al., 2020). With English now comprising roughly half of the Internet’s content, speakers of underserved languages must often abandon their native languages to participate in the technological mainstream (Grenoble, 2011), which accelerates language shift and undermines intergenerational transmission of cultural heritage (Claus, 2024).

Recent advances in natural language processing (NLP) and mobile-assisted language learning (MALL) demonstrate that technology can support documentation and revitalisation (Rießler, 2013; Varlamov et al., 2020), but most tools remain inaccessible to communities whose languages lack standard orthographies or substantial digital data (Anastasopoulos et al., 2020; Zhang et al., 2022). Dzardzongke is one of those severely endangered languages, spoken by around 1,200 people in South Mustang, Nepal (Meelen et al., 2024a).¹ The lan-

¹Dzardzongke’ is one of the terms used by the local population to refer to their way of speaking, but in some documentation, e.g. Glottolog (bara: 1356) and the World Atlas of Linguistic Structures (WALS), it is referred to as ‘Baragaunle’, which is a Nepali term for the region.

guage is used alongside other languages, but absent from education and digital spaces (Kretschmar, 1995; O’Neill et al., 2023). With no written record and younger speakers moving away, switching to more dominant languages, the next generation is at high risk of language attrition and complete language loss (Meelen et al., 2024a).

In this paper, we therefore present the first prototype of a mobile application designed to support Dzardzongke language revitalisation and cultural preservation. The app adopts a participatory, community-driven design, developed in close collaboration with speakers across the local region and the international diaspora community (Perlin et al., 2021). This participatory approach reflects a broader call in the field for NLP practitioners working with oral societies to develop locally appropriate technologies that centre the speech community rather than treating language as data for machine exploitation (Bird and Yibarbuk, 2024). Building on prior documentation and archiving efforts (Meelen and Ramble, 2022), we present a case study of Dzardzongke, demonstrating how language technology can support the preservation and revitalisation of endangered languages and their associated cultural heritage.

2 Background Information

All native speakers of Dzardzongke are multilingual and spend their daily lives in other languages, highlighting its status as an endangered language (Meelen et al., 2024a). Dzardzongke is primarily an oral language, with no established written tradition and no digital textual resources. As a result, the language is extremely underserved, with only very limited written materials available, including one article (Drandul, 2024) and a short children’s book covering numbers, colours, and animals (Meelen et al., 2024b).

To address this gap, a standardised orthogra-

phy has recently been developed in collaboration with speakers, providing a first step towards enabling digital language use and supporting the development of language technologies for Dzardzongke (O’Neill et al., 2023). The orthography is based on a romanised script, reflecting community preferences for using Latin characters already familiar from English and commonly used in informal digital communication, even for Nepali. This orthography is also based on the standard range of the Latin alphabet, with just one addition of the acute accent for some high-tone syllables that would otherwise be homographs. This romanised orthography enables a straightforward mapping between spoken and written forms, facilitating both literacy development as well as NLP.

At the same time, increasing smartphone adoption in the region creates new opportunities for MALL, despite remaining connectivity limitations. Early interactions with community members further indicate a growing interest in tools that support literacy and language preservation, motivating the development of MALL and NLP applications for Dzardzongke.

3 App Design & Development

3.1 Data Collection & Curation

We collected digital language data in Dzardzongke during fieldwork trips in 2022 and 2025, recording conversations and word lists. In total, we recorded two conversations and a word list comprising over 500 lexical items.² Examples 1 and 2 show an excerpt of one of the conversations in the newly-developed orthography that distinguishes between high (with acute accent) and low tone (no accent) in, for example, *ngá* ‘five’ vs *nga* ‘I, me’:

- (1) *Khangpa la mi gatsoe yoeta?*
house in people how.many are.PRES
‘How many people are there in your house?’
- (2) *Égi ngá yoe. Nga, ngi áwu, áni,*
we.EXCL five are me my father aunt
no cik, numu cik.
younger.brother one younger.sister one
‘We are five: me, my father, aunt, one younger brother and one younger sister.’

The data collection was inspired by existing pedagogical resources for underserved and endangered

²All materials, including the app-specific recordings, are archived at ELAR: <http://hdl.handle.net/2196/aa07e8d9-de4a-4820-af20-a34054068b91>.

languages, including *Ti Liv Kréyòl* (‘Little Book of Creole’) (Guillory-Chatman et al., 2020), which provides a user-friendly introduction to Louisiana Creole. Additional inspiration was drawn from open-access materials for the endangered Chatino language, developed by Hilaria Cruz and colleagues, including illustrated resources that were adapted for use in our application (see Figure 1) (Cruz, 2022).³

Finally, we also collected culturally-specific data, including descriptions of local villages and festivals, informed by anthropological research on the region (Ramble, 2008). Integrating elements of cultural heritage was a deliberate choice since these are more salient (and therefore deemed more worthy of preservation) to the local community than the language. Years of monolingual education in Nepali and a lack of acknowledgement of local languages without a written history mean awareness and appreciation of these languages is often low, because knowing official languages such as Nepali and English is associated with prestige and economic gain.

The recent re-appreciation for traditional festivals and rituals by the local community can, through this approach, be linked to the languages that are at the brink of extinction alongside these cultural traits. Dzardzongke is a prime example since to this day it still does not have its own ISO code, does not exist in the EQUATE Language AI Readiness Index (Occhini et al., 2026), and is not recognised on its own in the national context, being often confused with Loke, another Tibetic variety spoken in Upper Mustang, just north of the Dzardzongke area. The recently-built road through the main Dzardzongke valley towards the pilgrimage site of Muktinath has made it easier for migrants from the local community to return to their home villages for traditional festivals. Therefore, with rising enthusiasm to preserve these cultural elements, we raise awareness for language preservation as well.

3.2 Language Learning App Survey

In order to design the most appropriate educational activities, we conducted a structured review of 10 commercial language learning apps, which are listed in Appendix A.1. We systematically identified design patterns and functionalities relevant to a human-centred app for the Dzardzongke language

³More of Hilaria Cruz’s work on Chatino is found on <https://ir.library.louisville.edu/chatino/>.

and culture. The analysis focused on (i) onboarding and user journeys (e.g. What profile information is collected from users, and how are learning goals set?), (ii) vocabulary presentation and practice mechanisms (e.g. Which spaced repetition or drilling techniques are used, and how is new vocabulary introduced?), (iii) support for speaking, listening, and literacy (e.g. How are audio, image, and text combined to support multimodal learning?), and (iv) treatment of cultural content and community perspectives (e.g. Does the app provide cultural context alongside linguistic content, and does it facilitate interaction with other speakers?). Each app was evaluated against a feature checklist derived from these questions, and notes were taken on which design choices appeared particularly suitable or unsuitable for an endangered, primarily oral language setting.

The apps were selected from the most popular language learning apps in the iOS App Store to cover a range of pedagogical approaches (e.g. game-like micro-lessons, pronunciation training, social interaction, and vocabulary drilling). From the survey, we adopted (i) deck-based vocabulary organisation and spaced repetition from flashcard-oriented apps, (ii) short, focused quizzes to support self-assessment, and (iii) rich multimedia support (images, audio, and short dialogues) for contextualised learning. For our mobile application, we treat native Dzardzongke users as already orally proficient in their own language and focus the design on reading, writing, and cultural heritage to support literacy. Unlike the analysed apps, which introduce entirely new languages to beginners, our app adapts familiar interaction patterns (flashcards, quizzes, dialogues) specifically to map between spoken Dzardzongke and the new Latin-based script to enable heritage speakers to benefit from the app as well.

3.3 Final User Journey

Starting the app, the user is prompted to log in or create an account to enable progress and other features. Once the user has created an account, the first page is the “Decks” screen, where new vocabulary is introduced through flash cards as shown in Figure 1. When choosing a specific topic, a card is presented showing an image or a word. When tapping on the flash card, the card turns around and shows the spelling of the Dzardzongke words and an example. By clicking “Got it”, the user advances to the next image and word. A flash card

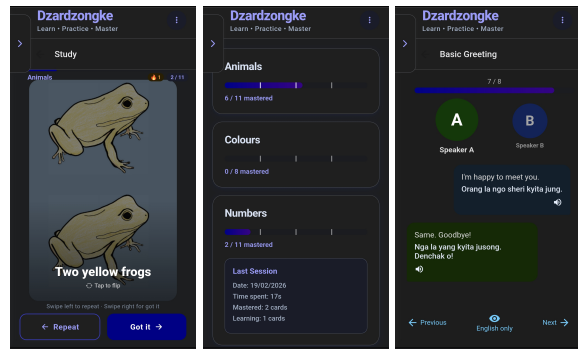


Figure 1: Flashcard decks, progress & conversations.

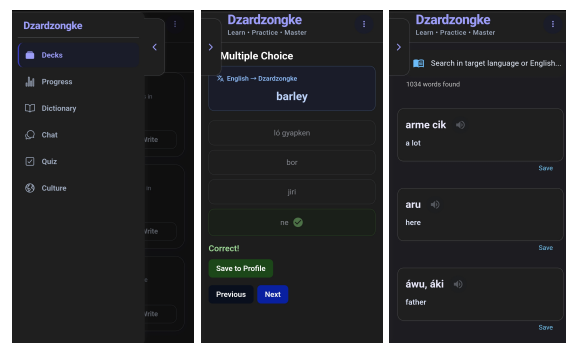


Figure 2: Menu, quiz & dictionary.

can also depict numerals and how they are spelt, as it is not just about the visual representation of them, i.e. “1”, but also about the spelling of the word *cik*, as in “one”, in Dzardzongke.

A searchable dictionary offers all the available words in alphabetical order, with their English translation and a sound file to practise listening and pronunciation, as shown in Figure 2. It also allows users to ‘Save’ items to their profile to facilitate easy custom-vocabulary retrieval. This approach to dictionary design for endangered language learners builds on prior work developing mobile dictionary interfaces specifically tailored to novice users of underserved languages, including approximate search to accommodate orthographic uncertainty (Littell et al., 2017). Furthermore, the user can be tested through multiple-choice quizzes. If the wrong word is picked, the user can save that word in their personal word list. Users can also see vocabulary in context through the chat feature, divided into themed categories of conversations. Each speech bubble in the chat has a sentence in English and Dzardzongke and is accompanied by audio files with the option of showing the conversation in English-only or Dzardzongke-only, so the

user can adjust it to their preference.

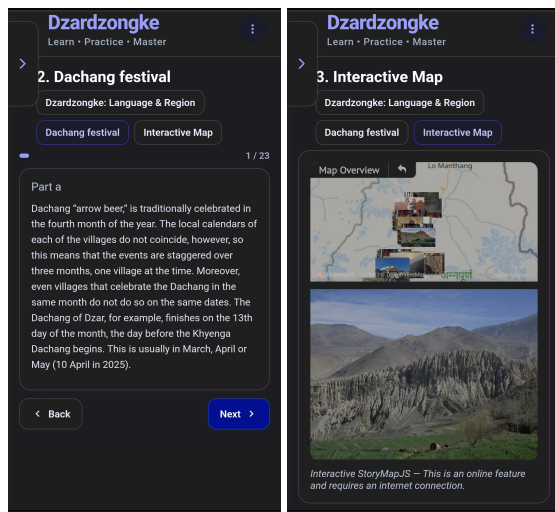


Figure 3: Samples of the culture section.

Finally, there is a “Culture” section with information on local traditions and festivals shown with images, videos, audio files, and multiple-choice quizzes to test the user’s knowledge. An interactive map that guides the user through the villages is also included, as can be seen in Figure 3. This can be used to boost local business and promote tourism as well.

The current version is a prototype, which will be shown to the first test users in Nepal and the diaspora in the next phase (see Limitations). Based on their feedback, we will amend the design to reflect the community’s interests and ideas. The updated version will then be made available free of charge for both iOS and Android devices.

3.4 Technical Implementation

The app is implemented using React Native with the Expo framework⁴ to support cross-platform development and deployment across Android, iOS, and web environments, ensuring compatibility with a broad range of devices. This cross-platform architecture enables researchers to modify or extend the system and redeploy it across platforms with minimal additional engineering effort (Claus et al., 2025).

Importantly, our implementation separates language-specific content from the core application framework. All linguistic resources are stored as structured static files (e.g. JSON) with associated multimedia assets, while the application layer handles interface rendering, state management, and

⁴<https://expo.dev>

user interaction logic. This content–framework decoupling enables new languages or additional resources to be incorporated without modifying the underlying system architecture. As a result, researchers can first of all easily add new language materials whenever they become available. Curating endangered language materials from documentation fieldwork typically takes time. Therefore, having the option to start out with some initial content that is more limited at first and expand it later provides the opportunity to quickly test features and allow for more interactive input from the local community in the development phase. Crucially, it also allows any other researchers to adapt the framework to develop similar applications for other languages without requiring any app development expertise.

Finally, all linguistic content and user data are stored locally on the device and in the current version, only the interactive map feature requires an internet connection. This offline-first design allows the application to function without continuous internet connectivity, which is particularly important in rural and underserved regions. Local storage further reduces infrastructure dependencies and enhances data privacy by avoiding cloud-based transmission.

3.5 Developing a Dzardzongke Keyboard

Because developing writing skills is one of the goals of the app, especially in heritage culture & language learning settings abroad, we also developed a keyboard for Dzardzongke to enable predictive text (Liu, 2025). Effectively, this allows speakers to have a specialised keyboard on their phone to use in- or outside of our app, with auto-correct and predictive texting based on the newly-created, standard Dzardzongke orthography. Prior work on endangered language keyboard design has demonstrated that community familiarity with existing typing conventions must be carefully considered alongside technical design choices (Santos and Harrigan, 2020). We created the keyboard using the Keyman Developer platform (Keyman, 2025). Keyman Developer allows users to modify a basic keyboard layout by changing the symbols that map to each key, and creating custom dead-keys and shortcuts for special characters. Once a custom keyboard is built, the Developer tool can read in a wordlist (with or without word frequency information) for the language that the keyboard is designed for, to provide orthographic informa-

tion to build the lexical model for autocorrect and predictive texting. Although symbols used in the romanised Dzardzongke orthography are readily type-able in existing keyboards,⁵ a specialised keyboard for Dzardzongke was useful in order to have the lexical model for prediction and autocorrection as well. This will enable users to quickly recognise how to spell words they use in their daily lives, even if they have not learnt how to read and/or write them before. The value of such predictive text tools for endangered language communities is supported by prior work, which has similarly developed text prediction capabilities as part of a broader suite of community-driven language technologies (Kuhn et al., 2020).

Finished Keyman keyboards are publicly available across a wide range of devices through the Keyman app, giving our Dzardzongke keyboard and lexical model a wide reach. Public Keyman keyboards can be easily integrated into Android keyboards by Google, which would allow for general use without the Keyman app as well. Overall, the availability of Dzardzongke autocorrect and predictive texting will increase user satisfaction with their experience on our app, and with typing Dzardzongke generally in other applications, e.g. through text messages, thereby furthering the long-term impact of this pilot project.

4 Conclusion

In this paper, we presented the first prototype of a mobile application and dedicated keyboard designed to support the preservation and revitalisation of Dzardzongke and its associated cultural heritage. Building on existing documentation, community consultations, and a survey of popular language learning apps, we adapted familiar interaction patterns to an endangered, primarily oral language context. Even the mere existence of such an app aids preservation, as it has already shown local communities, but also officials, that the language and culture is worth the effort.⁶ This is particularly pertinent for highly endangered languages like Dzardzongke that do not have any acknowledged

⁵Apart from the standard 26 letters in most English-based keyboards, Dzardzongke adds the option of an acute accent on vowels.

⁶As was the case, for example, by the publication of research by colleagues on the Gompa Gang temple in Chuksang, which led directly to a large investment of international charities to renovate and preserve the temple helping not only the local temple-going community, but also the general economy boosting tourism in the area.

official status yet. One of the strengths of our approach is the flexible set-up, which allows both language users and researchers on Dzardzongke, but also other languages, to create their own version. Future work will also involve iterative co-design and user studies with speakers in South Mustang and the diaspora to evaluate usability, learning outcomes, and community acceptance.

Ethical Considerations

Since the app contains audio-visual materials, some parts of the data cannot be anonymised, and enhanced ethical approval was sought and obtained for two separate fieldwork trips from the following universities: University of Cambridge and EPHE-PSL, Paris. The choice for starting this prototype for Dzardzongke first was guided by the recent renewed interest from this specific community (both locally in Nepal and in New York) for their cultural heritage.

Limitations

The choice to make this first version a prototype was a deliberate one. It has been tested by developers and linguists and preliminary features have been shared with Dzardzongke speakers to keep them involved throughout the development. It therefore forms a real contribution to endangered language documentation and revitalisation research. There is a major risk involved with rolling out an app for endangered language speakers that has not gone through comprehensive testing yet, as they may quickly lose interest if any errors are encountered. To maximise impact, we therefore plan a careful evaluation of the finished prototype that not only involves thorough checking of the data and general features, but also surveys ease of use in line with community wishes and preferences, both for local communities in Nepal as well as heritage speakers abroad. We will systematically test each section with users from different backgrounds and age groups to find out how easy it is to go through the exercises, which sections have their preference and why and how they think the app can generally be improved. Speakers who would like to participate in this survey have already been recruited, and the overall evaluation should be finalised in the next few months. The basic structure for the app can, in the meantime, be made available (Claus et al., 2025).

Acknowledgements

We would like to thank the Dzardzongke community in Nepal and abroad for their warm welcome and enthusiastic participation in this project, including the teachers and kids of the Lubrak school who shared their drawings and specifically the speakers who kindly helped with recordings: Kemi Tsewang, Palgen Bista, Gyaltzen Gurung and Charles Ramble.

This research was partially supported by the European Union (ERC, PaganTibet, 101097364), the Endangered Language Documentation Programme (ELDP SG 0716), the Cambridge Humanities Research Grant (CHRG), and the UK Research and Innovation (UKRI) Frontier Research Grant EP/Y031350/1 under the UK government's funding guarantee for ERC Advanced Grants for the project entitled 'Towards Globally Equitable Language Technologies (EQUATE)'. Hannah M. Claus is supported by Gates Cambridge Trust (Grant no. OPP1144 from the Bill & Melinda Gates Foundation). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

- Antonios Anastasopoulos, Christopher Cox, Graham Neubig, and Hilaria Cruz. 2020. Endangered Languages meet Modern NLP. In *Proceedings of COLING 2020: Tutorial Abstracts*, Barcelona, Spain. International Committee on Computational Linguistics.
- Babbel. 2025. Babbel: Learn Spanish, French and Other Languages Online. <https://uk.babbel.com>. Accessed August 25, 2025.
- Steven Bird and Dean Yibarbuk. 2024. *Centering the Speech Community*. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 826–839, St. Julian's, Malta. Association for Computational Linguistics.
- Damian Blasi, Antonios Anastasopoulos, and Graham Neubig. 2022. *Systematic Inequalities in Language Technology Performance across the World's Languages*. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5486–5505, Dublin, Ireland. Association for Computational Linguistics.
- Busuu. 2025. Busuu - Learn Languages Online: Start for Free. <https://www.busuu.com>. Accessed August 25, 2025.
- Hannah M. Claus. 2024. Now you are speaking my language: Why minoritized LLMs matter. Ada Lovelace Institute. <https://www.adalovelaceinstitute.org/blog/why-minoritised-llms-matter/>. Accessed March 25, 2026.
- Hannah M. Claus, Songbo Hu, Emre Isik, Anna Korhonen, Kitty Wenyang Liu, and Marieke Meelen. 2025. *Language Learning App for Dzardzongke*. Accessed May 8, 2026.
- Hilaria Cruz. 2022. Chatino Tonal Books Project. <https://ir.library.louisville.edu/chatino/>. Accessed March 30, 2026.
- Nyima Drandul. 2024. *Mustang ki mithok zhi ki lungbi amchiyak ki kyiduk dang dzedzeta kor - Reflections on the Lives and Works of Four Generations of Village Amchis from Mustang*. Lumbini International Research Institute. Retrieved May 6, 2026.
- Duolingo. 2025. Duolingo. <https://www.duolingo.com>. Accessed August 25, 2025.
- EF Education First. 2025. EF Hello. <https://hello.ef.com>. Accessed August 25, 2025.
- ELSA Speak. 2025. ELSA Speak - English Accent Coach. <https://elsaspeak.com/en>. Accessed August 25, 2025.
- EWA. 2025. EWA - Learn English, Spanish, French Online | Language Learning. <https://appewa.com>. Accessed August 25, 2025.
- Lenore A. Grenoble. 2011. *Language ecology and endangerment*, page 27–44. Cambridge Handbooks in Language and Linguistics. Cambridge University Press.
- Adrien Guillory-Chatman, Oliver Mayeux, Nathan Wendte, and Herbert J. Wiltz. 2020. *Ti liv Kréyòl: A Learner's Guide to Louisiana Creole*. New Orleans: TSHOK. Retrieved May 6, 2026.
- HelloTalk. 2025. HelloTalk - Language Exchange - Learn Languages for Free. <https://www.hellotalk.com>. Accessed August 25, 2025.
- Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. 2020. *The State and Fate of Linguistic Diversity and Inclusion in the NLP World*. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6282–6293, Online. Association for Computational Linguistics.
- Keyman. 2025. Keyman keyboard layouts. <https://keyman.com>. Accessed August 25, 2025.
- Monika Kretschmar. 1995. *Erzählungen und Dialekt aus Südmustang: Wörterbuch zum Südmustang-Dialekt*. International Institute for Tibetan and Buddhist Studies. Retrieved May 6, 2026.

- Roland Kuhn, Fineen Davis, Alain Désilets, Eric Joannis, Anna Kazantseva, Rebecca Knowles, Patrick Littell, Delaney Lothian, Aidan Pine, Caroline Running Wolf, Eddie Santos, Darlene Stewart, Gilles Boulianne, Vishwa Gupta, Brian Maracle Owenatékha, Akwiratékha’ Martin, Christopher Cox, Marie-Odile Junker, Olivia Sammons, and 7 others. 2020. [The Indigenous Languages Technology project at NRC Canada: An empowerment-oriented approach to developing language software](#). In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 5866–5878, Barcelona, Spain (Online). International Committee on Computational Linguistics.
- Language Drops. 2025. Language Drops - Learn Languages. <https://languagedrops.com>. Accessed August 25, 2025.
- Lingvist. 2025. Lingvist: Learn a New Language Smarter and Faster Online. <https://lingvist.com>. Accessed August 25, 2025.
- Patrick Littell, Aidan Pine, and Henry Davis. 2017. Waldayu and Waldayu Mobile: Modern digital dictionary interfaces for endangered languages. In *Proceedings of the 2nd Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pages 141–150.
- Kitty Wenying Liu. 2025. Dzardzongke keyboard. <https://github.com/keymanapp/keyboards/tree/master/release/d/dzardzongke> Accessed May 5, 2026. Developed using Keyman Developer.
- Marieke Meelen, Alexander O’Neill, and Rolando Coto-Solano. 2024a. End-to-end speech recognition for endangered languages of Nepal. In *Proceedings of the Seventh Workshop on the Use of Computational Methods in the Study of Endangered Languages*, pages 83–93.
- Marieke Meelen and Charles Ramble. 2022. [An Audio-Visual Archive of Dzardzongke \(South Mustang Tibetan\)](#). Endangered Languages Archive. Retrieved July 11, 2025.
- Marieke Meelen, Charles Ramble, and Kemi Tsewang. 2024b. *Ngi Dzardzongke ki choe gongma [My first Dzardzongke book: early reader to learn the language of South Mustang, Nepal]*.
- Memrise. 2025. Memrise. <https://www.memrise.com>. Accessed August 25, 2025.
- Giulia Occhini, Kumiko Tanaka-Ishii, Anna Barford, Refael Tikochinski, Songbo Hu, Roi Reichart, Yijie Zhou, Hannah Claus, Ulla Petti, Ivan Vulić, Ramit Debnath, and Anna Korhonen. 2026. [Artificial intelligence is creating a new global linguistic hierarchy](#). Preprint, arXiv:2602.12018.
- Alexander O’Neill, Marieke Meelen, Rolando Coto-Solano, Sonam Phuntsog, and Charles Ramble. 2023. [Language Preservation through ASR](#). Poster at the Cambridge Language Sciences Annual Symposium 2023.
- Ross Perlin, Daniel Kaufman, Mark Turin, Maya Dau-rio, Sienna Craig, and Jason Lampel. 2021. Mapping urban linguistic diversity in New York City: motives, methods, tools, and outcomes. *Language Documentation & Conservation*, 15:458–490.
- Charles Ramble. 2008. *Tibetan Sources for a Social History of Mustang, Nepal*. International Institute for Tibetan and Buddhist Studies.
- Michael Riebler. 2013. Towards a digital infrastructure for Kildin Saami. In *Sustaining indigenous knowledge. Learning tools and community initiatives on preserving endangered languages and local cultural heritage*, pages 195–218. SEC Publications.
- Eddie Antonio Santos and Atticus Harrigan. 2020. [Design and evaluation of a smartphone keyboard for Plains Cree syllabics](#). In *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)*, pages 88–96, Marseille, France. European Language Resources association.
- Aleksandr Varlamov, Galina Keptuke, and Alexandra Lavrillier. 2020. [Electronic devices for safeguarding Indigenous languages and cultures \(Eastern Siberia\)](#). In Timo Koivurova, Else Grete Broderstad, Dorothée Cambou, Dalee Dorough, and Florian Stammler, editors, *Routledge Handbook of Indigenous Peoples in the Arctic*, 1 edition, pages 58–75. Routledge.
- Shiyue Zhang, Ben Frey, and Mohit Bansal. 2022. [How can NLP Help Revitalize Endangered Languages? A Case Study and Roadmap for the Cherokee Language](#). In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1529–1541, Dublin, Ireland. Association for Computational Linguistics.

A Appendix

A.1 List of apps for the language learning app survey

The following language learning apps were analysed: [Babbel \(2025\)](#), [Duolingo \(2025\)](#), [Memrise \(2025\)](#), [Busuu \(2025\)](#), [HelloTalk \(2025\)](#), [EF Education First \(2025\)](#), [Lingvist \(2025\)](#), [EWA \(2025\)](#), [ELSA Speak \(2025\)](#), [Language Drops \(2025\)](#).

A.2 Technical Implementation Details

The app is written primarily in TypeScript (96.2%), with a small JavaScript component (3.8%), using the React Native Expo framework for cross-platform deployment on Android, iOS, and web. The web version is publicly accessible, extending the app’s reach beyond mobile devices ([Claus et al., 2025](#)). Android APK builds are handled via EAS (Expo Application Services), enabling straightforward distribution during the pro-

prototype phase without requiring app store submission. All linguistic content is stored in structured JSON files: the dictionary is contained in a single file (`dzardzongke.dict.json`) with fields for the Dzardzongke word, English translation, example sentences in both languages, and an optional linked audio file; flashcard decks follow an analogous structure stored in `/assets/decks/`. The dictionary features fuzzy search, which is particularly important given that users may be uncertain about spelling in the newly standardised orthography. To lower the barrier for non-technical contributors, the app supports a Google Sheets-based content management workflow, whereby community members or researchers can add and edit dictionary entries, flashcard decks, culture content, and quiz material directly in a spreadsheet, then sync changes to the app with a single command (`npm run export-content`), with no coding required.