

The Endangered Languages Project (ELP): Collaborative Infrastructure and Knowledge-Sharing to Support Indigenous and Endangered Languages

Anna Belew

Endangered Languages Project
1890 East-West Rd., #569, Honolulu, HI 96822
anna@endangeredlanguages.com

Abstract

While global language loss is currently occurring at a rate higher than ever in known history, the prevalence of language documentation and revitalization work is also rising dramatically. However, many such initiatives are taking place in relative isolation, without clear pathways for collaboration and knowledge-sharing across nations and global regions. The Endangered Languages Project (ELP) addresses this need by providing infrastructure for the dissemination of knowledge about Indigenous and endangered languages, their vitality, and their revitalization efforts; sharing of digital resources related to Indigenous and endangered languages; and fostering international collaboration and networking between language workers and researchers.

Keywords: knowledge-sharing, infrastructure, collaboration

摘要

儘管現今全球的語言正在以史上最快的速度消失，語言保存和復振也在快速的進當中。然而，這些活動往往單獨的在進行，而且顯少有跨國或全球性的合作及知識途徑。瀕危語言計畫（Endangered Languages Project (ELP)）正是為了這個需求誕生，提供傳播原住民和瀕危語言知識、活力以及成果的設施，並與大眾分享原住民族和瀕危語言的數位資源，積極促進語言工作者和研究者之間的國際合作和人際網絡。

1. Global Language Endangerment and Revitalization

Of the roughly 7,000 languages spoken on Earth today (Eberhard, Simons and Fennig, 2019), nearly half—approximately 3,170 languages, or c. 45%—are currently at risk of falling silent (*Catalogue of Endangered Languages*, 2019). Another 183 languages, at least, are known to have fallen silent in the past half-century, and an additional 70 or more languages are currently being revived after having lost their last fluent speakers (*ibid.*). The majority of these are Indigenous and/or minoritized languages, and shift away from these languages often corresponds to other forms of socio-economic, political, and cultural marginalization and injustice.

While rates of global language loss are currently higher than at any point in known history, the incidence of language documentation and revitalization is also at an all-time high (Belew and Simpson, 2018). Indeed, Pérez Báez, Vogel, and Patolo (2019) report that, of the revitalization programs surveyed in their research, 65% were initiated after the year 2000, and 30% began in the current decade. At least 41 known revitalization programs have begun in the past ten years, and new initiatives are continually being launched (*ibid.*). Similarly, Thieberger (2017) reports that language documentation, if assessed in terms of grammars produced for small and endangered languages, has increased notably since the year 2000.

As programs to support linguistic diversity proliferate in all parts of the world, there is a need for digital infrastructure to support these initiatives with reliable data, technological tools, knowledge-sharing mechanisms, and avenues for collaboration. The following sections outline these needs, and how the Endangered Languages Project's digital services and resources can

support those working to sustain Indigenous and endangered languages.

2. Needs for Infrastructure to Support Language Documentation and Revitalization

While the number of initiatives to document and revitalize the world's languages is growing, many language workers today are still carrying out their work in relative isolation, without a clear path for connecting with people and programs in other parts of the world. Language workers may be discouraged by the feeling carrying out language work alone, if there is limited support from populations and institutions in a given context. For example, one language activist from South Africa, in conversation with the author during a conference, described being demoralized by feeling that they were the only person engaged in this type of work; they had few connections with other language activists, and limited knowledge of initiatives in other parts of the world. However, after being invited by ELP to attend an Indigenous language revitalization conference to network, collaborate, and share knowledge with language workers from around the world, and maintaining online contact with their new colleagues after they returned to their home countries, the language activist reported that their work was reinvigorated by new motivation, ideas, and methods. Knowledge-sharing and networking with fellow language workers is often key to sustainable, effective language work.

Similarly, many researchers working with Indigenous and endangered languages lack a place to make their research more widely accessible, and to connect with

language stakeholders outside of academia. Moreover, reliable data on endangered languages may be difficult to come by for many researchers and language workers—it may be locked behind a paywall, buried in dense academic volumes, or written in a language not known to the reader—and without accurate information about a language and its current context, it is difficult to devise an effective strategy for documentation or revitalization. Similarly, policymakers, funding agencies, activists, and governments are in need of accurate data about each of the world’s endangered languages, in order to make informed decisions about language policy, public outreach, and the allocation of limited resources (Hauk and Heaton 2018). As some prominent sources of language information are now moving to paid access models¹, and others are falling out of date after losing funding, there is an enormous need for online resources which provide free, up-to-date information about endangered and Indigenous languages.

In addition, helping stakeholders stay informed of global activity in the field of Indigenous and endangered languages is a significant need—as demonstrated by the spread of the highly successful language nest model from Aotearoa/New Zealand, to Hawai‘i, to North America and Europe (Hinton 2018), the dissemination of effective ideas and best practices across the globe is crucial to ensuring that Indigenous and endangered languages can thrive in the coming decades.

Finally, as technology tools—such as smartphone keyboards, software localization, and language learning apps—become an increasingly central part of language maintenance and revitalization, there is a growing need for language workers to have access to best practice guides, networks of fellow practitioners, and reliable information to support the development of these technologies.

3. The Endangered Languages Project (ELP)

The Endangered Languages Project (ELP) is an online resource (accessible at www.endangeredlanguages.com) designed to promote and facilitate the documentation and revitalization of at-risk languages around the world. ELP aims to bring together language workers, researchers, and the public to share knowledge, best practices, ideas, and news about the world’s Indigenous and endangered languages. The project serves both as a resource for information on the endangered languages of the world, and as an arena to collaborate with others working to document, revitalize, and promote Indigenous and endangered languages.

Launched in 2012, ELP was originally founded by four core partners: the First Peoples’ Cultural Council (FPCC), the University of Hawai‘i at Mānoa (UHM), The LINGUIST List at Eastern Michigan University, and Google.org. It is currently led by two founding partners

(FPCC and UHM), and a Governance Council of individuals from nine organizations and universities around the world². The website interface is currently available in seven languages (English, Spanish, French, German, Traditional Mandarin, Russian, and Brazilian Portuguese), and is currently in the process of being localized into additional languages, including Korean, Japanese, Italian, and Hindi. ELP plans to expand the site interface’s availability in other languages, particularly Indigenous languages, as the project continues. The site currently has over 20,000 registered users, and receives an average of three million annual pageviews.

The following sections outline the services offered by ELP, and the ways in which ELP’s infrastructure may be used and expanded to meet the needs of Indigenous and endangered languages online.

3.1 The Catalogue of Endangered Languages (ELCat)

The *Catalogue of Endangered Languages* (ELCat) is a database of information about the world’s endangered languages. ELCat is a cost-free resource for language workers, researchers, educators, language communities, policymakers, funding agencies, and the public to learn about global language endangerment and revitalization.

ELCat provides information about the vitality and context of each of the world’s endangered, dormant, and awakening³ languages, which at the time of writing numbered 3,426. The ELCat database aggregates information from all available, reputable sources, including journals, books, field notes, reports from Indigenous language organizations, and direct communications from individuals working directly with a given language. To ensure the reliability of ELCat data, this information is vetted by the project’s International Board of Directors, who are individuals with expertise in the languages of specific global regions; see (Campbell and Belew 2018) for more on ELCat’s data collection practices. ELCat uses an original metric for assessing language vitality, the Language Endangerment Index (LEI), and provides visualizations of the vitality of all of the world’s endangered languages in the form of a global map, as shown in Figure 1. For more on the LEI, see (Lee and Van Way 2016).



Figure 1: Partial view of ELP map of endangered languages, color-coded by vitality level.

¹ <https://linguistlist.org/issues/30/30-4081.html>

² <http://endangeredlanguages.com/about/>

³ “Dormant” refers to languages which have lost their last fluent speaker in roughly the past 50 years, while the “Awakening” label is applied to languages which at some point had no fluent speakers, but are currently undergoing concerted revival efforts. See (Belew and Simpson 2018) for more on these languages.

3.2 Dissemination of Digital Multimedia Resources

The ELP website also offers a platform for users to disseminate digital multimedia resources—videos, audio files, images, documents, and web links—related to endangered and Indigenous languages. In addition, there is a discussion function for users to comment on, share information about, and make recommendations for resources similar to those included in the site. ELP allows all site users to upload multimedia resources, encouraging the free exchange of ideas and media on the topic of endangered languages. Resources uploaded to the ELP site are lightly moderated by the project’s staff to ensure that inappropriate or irrelevant content is removed.

Users can browse multimedia resources by language—each language page features a “Resources” tab, displaying all relevant materials—as well as by category and keyword. Resource categories include topics such as language education, language and technology, and language advocacy and awareness; these categories aggregate topically similar materials from many different languages. For example, a user interested in language nests could browse all materials tagged “language nest” within the “language education” category, and access videos, documents, and news articles related to language nests worldwide. They can also view the profiles of other users who have submitted (or commented on) resources related to language nests, and connect with one another for collaboration and networking.

There are currently over 7,000 multimedia resources available on the ELP site, with new items being uploaded daily by users. Such a platform provides a useful avenue for resource discovery and dissemination, a key need in supporting the presence and use of endangered languages online.

3.3 Digital Training in Language Documentation

ELP seeks to address the need for free training and information related to language documentation, particularly among speakers of endangered and under-documented languages who do not have access to formal training programs, such as university courses or workshops, ELP has partnered with the Language Documentation Training Center (LDTC) at the University of Hawai‘i to offer digital training in language documentation. These eight-week webinars, held weekly via Facebook Live, provide training in basic facets of language documentation (audio/video recording, descriptive phonology, orthography development, etc.), as well as practice exercises and further reading. In addition to watching the livestreamed lessons, asking questions of the workshop leaders, and discussing the material with other learners, participants are also paired with “mentors” who have advanced skills in language documentation for more personalized guidance.

The first webinar, held in early 2018, drew 465 participants from 62 countries. Participants were particularly enthusiastic about the opportunity to network with language workers and researchers in other parts of the world; as discussed in §2, a sense of isolation in the work

of language documentation, revitalization, and advocacy can be demoralizing, while collaboration and interaction can renew motivation and sustainability of documentation efforts. This type of interactive webinar is one method for both increasing access to language documentation training, and for facilitating networking and knowledge-sharing among speakers of Indigenous and endangered languages.

3.4 Revitalization Directory and Helpdesk

3.4.1 Revitalization Directory

ELP is currently developing a directory of revitalization initiatives around the world, based on the Global Survey of Revitalization Efforts developed by Pérez Báez, Vogel, and Patolo (2019). ELP’s revitalization directory will build upon the results of Pérez Báez et al. and provide an expanded listing of language revitalization programs. Users will be able to submit information about revitalization programs they are involved with, such that the directory will constantly be expanded and updated.

The directory will allow ELP users to network with other revitalization workers who are working with similar methods, languages, or contexts. For example, a teacher at a Hawaiian-immersion high school could search for other secondary-education revitalization programs in order to find lesson planning ideas, or learn about successful revitalization methods for languages in similar sociopolitical contexts. In addition, the directory will highlight the growing prevalence and diversity of language revitalization work around the world; it is our hope that this will provide encouragement and a sense of global solidarity to language revitalization workers, and demonstrate the resilience of Indigenous and endangered languages in the face of pessimistic discourses of language loss. This directory will provide online infrastructure for knowledge-sharing and collaboration across global regions, while also serving as a reliable source of information on global trends in language revitalization, paralleling how ELCat provides data on global trends in language vitality (see Campbell and Okura 2018).

3.4.2 Revitalization Helpdesk

The ELP Revitalization Helpdesk will complement the directory of revitalization initiatives by providing case studies, best practices, and personalized human assistance for revitalization workers. It can be intimidating to attempt to begin a revitalization effort, particularly in the absence of connections with established programs or language workers. There are a myriad of questions to address—what methods exist, and what goals do they serve? What types of activities have proven effective in similar contexts? What kinds of resources are needed? etc.—and seeking answers in the academic literature can be daunting for laypeople.

The Revitalization Helpdesk will thus provide resources for strategic language revitalization planning, such as “roadmaps” and tools for assessment of a language’s vitality and revitalization goals, drawing upon the extensive experience of ELP’s founding partners at the First Peoples’ Cultural Council. The Revitalization Helpdesk will also provide best-practice guides, profiles of

successful revitalization efforts, discussion areas to connect with other users engaged in language revitalization work, and a “helpdesk” which will connect language workers with volunteer experts in particular revitalization methods, language families, or geographic areas.

The directory and helpdesk are currently in development, and will launch in 2020.

4. Conclusion

The Endangered Languages Project provides infrastructure to support key aspects of language documentation and revitalization globally, including a free, reliable database of information about endangered languages and revitalization initiatives; online dissemination of multimedia resources; internet-based training in language documentation; and a “helpdesk” for language revitalization workers to find guidance, ideas, best practices, and collaboration with other revitalization initiatives around the world.

9. Acknowledgements

The *Catalogue of Endangered Languages* and Endangered Languages Project have received support from the U.S. National Science Foundation (grants BCS-1058096 and BCS-1057725), the Henry Luce Foundation, Google.org, the University of Hawai‘i at Mānoa, and the First Peoples’ Cultural Foundation. Translation of this paper’s abstract into Mandarin was provided by Khái-în Lâm.

5. Bibliographical References

- Belew, A. and Simpson, S. (2018). Language extinction then and now. In Lyle Campbell and Anna Belew, editors, *Cataloguing the World’s Endangered Languages*, pages 49–65. Routledge, New York, 1st edition.
- Campbell, L. and Belew, A. (2018). *Cataloguing the world’s endangered languages*. Routledge, New York, 1st edition.
- Campbell, L. and Okura, E. (2018). New knowledge produced by the *Catalogue of Endangered Languages*. In Lyle Campbell and Anna Belew, editors, *Cataloguing the World’s Endangered Languages*, pages 79–84. Routledge, New York, 1st edition.
- Catalogue of Endangered Languages (ELCat)*. (2019). University of Hawai‘i at Mānoa, available online: <http://www.endangeredlanguages.com>.
- Eberhard, D.M., Simons, G., and Fennig, C. (2019). *Ethnologue: Languages of the world*. 22nd edition. SIL International, Dallas, available online: <http://www.ethnologue.com>.
- Hauk, B. and Heaton, R. (2018). Triage: Setting priorities for endangered language research. In Lyle Campbell and Anna Belew, editors, *Cataloguing the World’s Endangered Languages*, pages 259–304. Routledge, New York, 1st edition.
- Hinton, L. (2018). Approaches to and strategies for language revitalization. In Kenneth L. Rehg and Lyle Campbell, editors, *The Oxford Handbook of Endangered Languages*, pages 443–465. Oxford University Press, Oxford, 1st edition.
- Lee, N.H. and Van Way, J. (2016). Assessing levels of endangerment in the *Catalogue of Endangered Languages* (ELCat) using the Language Endangerment Index (LEI). *Language in Society*, 45:271–292.
- Pérez Báez, G., Vogel, R., and Patolo, U. (2019). Global Survey of Revitalization Efforts: A mixed methods approach to understanding language revitalization practices. *Language Documentation and Conservation*, 13(1):446–513.
- Thieberger, N. (2017). LD&C possibilities for the next decade. *Language Documentation and Conservation*, 11:1–4.