Barriers experienced by contributors to small language versions of Wikipedia

A Movement Strategy Implementation Grant Research Report by the Language Diversity Hub

WIKIMEDIA LANGUAGE DIVERSITY HUB

WIKIMEDIA
NORGE

Abstract

With many new editions of Wikipedia emerging, we have identified the need to understand the situation for their contributors, as they encounter different barriers than those experienced by contributors to well-established editions of Wikipedia. This research project aims to lay the foundation for the Language Diversity Hub to improve the working conditions for contributors on new Wikipedia editions. The work is part of the implementation of the Wikimedia 2030 Movement Strategy, Initiative 4 (Ensure Equity in Decision-Making), Point 25 (Regional and Thematic Hubs).

Thirteen language communities across Europe, Africa, Asia, and Latin America were interviewed for this research project. Whilst some of the participants of this research project were active contributors to smaller language Wikipedias in the Wikimedia Incubator, others were contributors to full-fledged Wikipedias. The interviews aimed to find challenges related to technology, education, economy, and/or social conditions, with the main emphasis on technology. The top challenges we found were: few contributors (10/13), language tech challenges (10/13), economic limitations (8/13), and the need for more training (7/13). More than half of the communities report that mobile phones are used as the main editing device. We found that all contributors that have edited in the Incubator report that editing became easier, and many of them report increased motivation after their Wikipedia edition was published. While some of the barriers are way beyond the scope of the work of the Language Diversity Hub (i.e. local infrastructure, lack of language standardization), others can be solved by putting existing Wikimedia resources to work (improving the Incubator, providing more training), and some would benefit from collaborations with other organizations (language tech projects, attracting more contributors).

Using the Language Diversity Hub to work with these issues provides a central unit that can facilitate collaboration with external institutions, and be a central point for launching projects. However, more work needs to be done to define what role the Language Diversity Hub should play and how the Hub can be governed in an inclusive and efficient way.

Content

Introduction

When Wikipedia was launched more than 20 years ago, the goal was to make knowledge available to all for free, online. And as history has shown, the impossible project of an encyclopedia written as a collaboration between volunteers is not only possible, but also wanted and celebrated. The importance of Wikipedia can be said to have outgrown its initial ambition; it is no longer *just* a place for free knowledge. It is a place for reclaiming and revitalizing languages, history, culture, and knowledge that are in danger of being lost. However, Wikipedia is biased towards contributors with good technical skills that write about topics that are well-documented in research and news outlets and in languages that are supported by technical tools. How is the situation perceived by Wikipedia contributors that fall outside one, more, or all of these categories?

As part of the Wikimedia Language Diversity Hub work, this research project's overarching theme is initiative 4 (Ensure Equity in Decision-Making), point 25 (Regional & thematic hubs). The goal with the research project is to collect and analyze data from a globally varied array of sources about what the challenges are for new and upcoming Wikipedia editions and their communities.

The goals of this research project are:

- Identifying barriers: This research project will be the first step in documenting such barriers on a global scale for this class of contributors. Since this is an early stage, we wish to be open to a large spectrum of possible barriers and challenges, to get a foundation for building the Hub and for further research.
- Building community: The research project will bring the indigenous and small language communities closer together, facilitate communication between them, and build a global community of language diversity.
- Exploring the Hub method: As the hubs and their purpose are not very well-defined yet, an important part of this research project is exploring how a thematic hub might work. The research project will be a small scale pilot for future work of a hub.

Method

Interviews were chosen as the main method to gather data for this research. We wanted to gain rich qualitative data that will serve as a broad foundation for further research and further work on building a language diversity hub. 10 language communities were initially chosen, and we ended up interviewing 13 communities.

The interviews were semi-structured and designed to cover five main categories of potential barriers. The interview guide contained open questions that encouraged descriptions, and suggested follow-up questions to help dig deeper into the challenges.

The interviews were held in person or online. We tried to find interviewers as close as possible to the communities, preferably speaking in their own language. Interviews were recorded, transcribed and, when needed, translated to English. The transcriptions of the interviews were collected and analyzed using the tool Quirkos, to code and organize the information into the predetermined categories of barriers.

Barriers:

- Technical
- Economic
- Education and knowledge
- Social

The method was chosen to give the contributors a way to voice their opinions, describe their situation, and also give them a chance to reflect on their challenges and the ways they have solved them. The interviews were effective in building a stronger relationship between the interviewees and the interviewer.

Selected language communities

Language	ISO-code	Region	Language project status	Research stats
Angika	anp	South Asia	Incubator	complete
Dagbani	dag	Middle East and Africa	Wikipedia	complete
Wayuunaiki	guc	Latin America (LATAM) and the Caribbean	Incubator (published 2023)	complete
Gurene	gur	Middle East and Africa	Incubator (published 2023)	complete
Igbo	ig	Middle East and Africa	Wikipedia	missing transcript
Kashmiri	ks	South Asia	Wikipedia	missing transcript
Mon	mnw	East, Southeast Asia, and Pacific (EASAP)	Wikipedia	complete
Nahuatl	nah	Latin America (LATAM) and the Caribbean	Wikipedia	needs translation
Nigerian Pidgin	pcm	Middle East and Africa	Wikipedia	missing transcript
Paiwan	pwn	East, Southeast Asia, and Pacific (EASAP)	Wikipedia	complete
Quechua	qu	Latin America (LATAM) and the Caribbean	Wikipedia	needs translation
Tachelchit	shi	Middle East and Africa	Wikipedia	complete
Inari Sámi	smn	Northern and Western Europe	Wikipedia	complete

For Igbo and Nigerian Pidgin, data from the interviews is included in the quantitative parts based on the recorded interviews, because a transcript was not yet available at the time of analysis. For Nahuatl and Quechua, there is not yet an English transcript available, but efforts have been made to include them in parts of the quantitative summaries based on an automatic Spanish transcription of the interviews.

Findings

From the very first interviews, just by reaching out and sharing information, small waves of action happened. That way, one of the first findings is that active and personal outreach to smaller language communities and sharing information about the opportunities within the movement are on its own support. We did find barriers, though, and we will present them organized according to the four categories initially described: technical, social, economical, and educational.

Throughout the interviews, other different approaches and reflections that did not fit into the barriers surfaced. One prominent example is people describing their motivation to contribute. We have chosen to include some information about that at the end of this chapter.

Technical barriers

We set out on this project with an ambition to get clear pointers to where we could improve the platform technically to facilitate for the new language communities. The technical barriers we expected were not mentioned that much, instead the need for more training was a recurring need.

Technical barriers
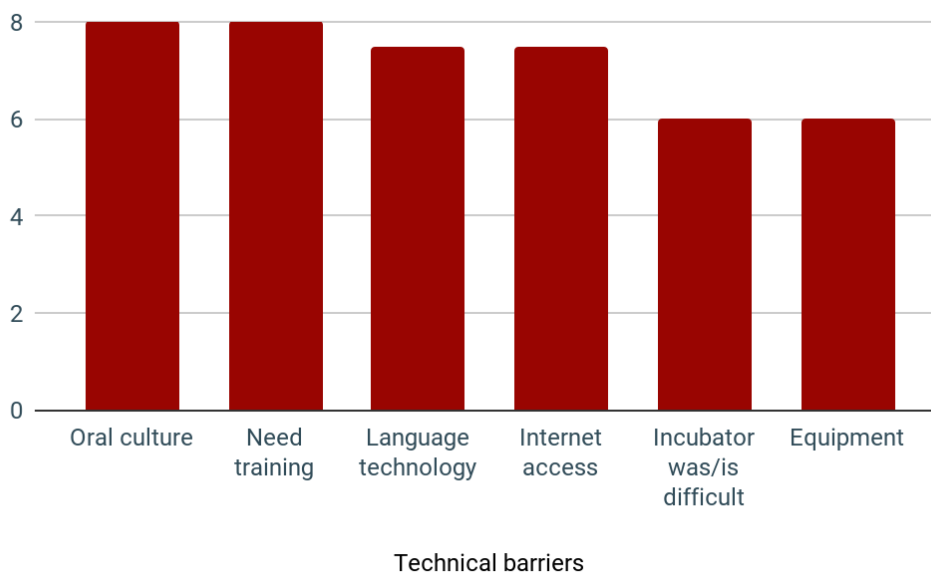
Oral culture

The oral culture of many indigenous cultures, and the challenges regarding oral sources on Wikipedia have been discussed for a long time already in some parts of the movement. It continues to be an important challenge to solve when working with underrepresented knowledge. The Paiwan community describes a

fast and efficient pipeline for recording and writing down oral culture, and then updating it to Wiki. The Wayuunaiki community specifically mentions *meeting the criteria of Wikipedia* as the most challenging part of dealing with an oral culture. New language versions of Wikipedia have their own power to establish their own requirements for their content.

Although many communities mentioned it as a challenge, no one pointed out oral culture as one of the top three barriers to contributing to Wikipedia in their language. It will be of value for smaller language communities to learn from each other in these questions, and their dialogue will also add value to larger language editions working to close knowledge gaps.

> ""
>
> The other challenge refers to the orality of our language.
>
> Our primary sources come from the elders of our community and the collective memory. There we have the challenge of aligning ourselves to the criteria of the traditional Wikipedia.
>
> Source: *Wayuunaiki*

> ""
>
> The reading and writing of the old people of the tribe is almost zero, but they will listen and talk. and we can combine that with the community associations and groups. For example, we can record the stories of the elders, write them down, and upload it. Updating this to the wiki is pretty fast, and the pipeline is good. I think this combination is pretty good.
>
> Source: *paiwan*

Wish for more training

Almost all the communities' clearest wish is to receive more training. The ask for more training comes from different perspectives. On one end, people wish to learn more about templates or other wiki-related skills, on the other—communities are in need of basic computer skills.

It might be that for the contributors, it is unnatural to point out limitations to the technology. They might be more focused on increasing their own knowledge and skills to be able to use the platforms better. In those cases, it is the responsibility of the more advanced users and the owners of the technology to understand the behavior and the needs of the users, and adapt the platform accordingly.

> It is a process, and we can also talk about the challenges we have encountered, such as the lack of equipment and connectivity; it is difficult for volunteers not only to have the equipment, but also to learn how to operate the computer equipment. Many volunteers lack digital literacy before even thinking about wiki editing
>
> Source: *Wayuunaiki*

> Mostly when we have in-person workshops or training in the office, routers are made available for contributors or participants to connect to their phones in the other access the internet and also the institutions where the community have a hub in Tamale College of Education) always permit us to used their Wi Fi for our activities during our training sessions in their institutions
>
> Source: *Dagbani*

On this note it is also worth mentioning how young potential contributors to Wikipedia in general are feeling ill-equipped, or with too low technical skills to edit Wikipedia. Thus increasing universal design, and reducing the technical skills needed for editing will probably benefit the entire movement.

Many communities report a combination of wishes for more training, a lack of equipment, and challenging internet conditions. A way to solve that in some communities has been offering a space where people can gather, learn together, access the internet, and sometimes also desktops, as a good service to the community. This can be helpful for attracting more contributors in general.

> **"** What I want to say is I know some about technique, but I want to learn more about how to write a good quality article, what kind of source we can use and then how to edit. When we don't have training, we feel that we are not good enough and we don't have confidence over what we have done.
>
> Source: *Mon*

Language Technology

A limitation for the presence of many languages on the Internet in general, and by consequence also Wikipedia, is the availability of tools to write in that particular language. Appropriate keyboards and online dictionaries are unavailable for many of the communities we spoke with. This was mentioned as a clear challenge in seven of the language communities. There are reasons to believe that the tools for spelling and grammar checking are underdeveloped in most of the languages, but they might not consider it as a barrier for contributing to Wikipedia. Due to the context of that language, they might not even have the expectations to find those tools.

However, the lack of such tools is a barrier for using and producing digital, written content, and so this lack is a clear barrier for building a written encyclopedia.

> **"** our major challenge is how to type the special keys in Dagbani. Most of our contributors are using mobile phones whiles other are using laptops and with android phone it seems to be better because the Ghanaian keyboard is of help to us but with iPhones and laptops is not easy at all, with the iPhones, you need to pay before he or she can have access to this special keys whiles the laptops you either installed it at a fee or need to search for specialized person to teach you how to used the normal keyboard at a fee. Please we need a lot help with regards to this challenge.
>
> Source: *Dagbani*

Finding the right translations for some of the technical terms has also been a challenge in several of the communities. The Paiwan community describes meetings where language users from different regions come together and discuss words and expressions. Many of the technical terms have no equivalent in these languages, and good processes for establishing new terminology on the conditions of the language will be of value for the languages in total.

Inari Sámi can celebrate a successful revitalization, going from no children speaking the language 30 years ago, to having schools and language nests and a growing number of speakers today. They are fortunate enough to have linguists that are also native speakers. Within the community, they have solid knowledge to make good translations and create new terminology on the premises of the language. The Inari Sámi community uses digital tools developed by Giellatekno, center for language technology at the Arctic University of Tromsø.

> "
> To translate some of the technical terms on the main page was very difficult for example terms like file,folder,edit,edit source and others are not easy to translate in Dagbani. It has to consume so much of our deliberating on this terms and to come agreement was not easy at all
>
> Source: *Dagbani*

The Incubator

We have interviewed language communities in the Wikimedia Incubator and those that have graduated to a full-fledged Wikipedia. Everyone who has contributed to both versions expresses that the full-fledged version is much easier to contribute to, and also that the motivation to contribute increased knowing that people could now find and start reading the articles.

In South-America they have launched a Indigenous language portal, to facilitate contribution from mobile phones, and provide safer steps with introduction on different tasks. Currently only available in Spanish and Wayuunaiki, this solution has potential for more language communities.

> "
> In relation to the incubator, the technology becomes much more complex for people who are not digital natives or people who are already on the web. We are grateful for the indigenous language portal that we are using, which has been very friendly to these new editors, to make the incubator more flexible when it comes to editing.
>
> Source: *Wayuunaiki*

Access to equipment

Computer editing
45,5%

Mobile editing
54,5%
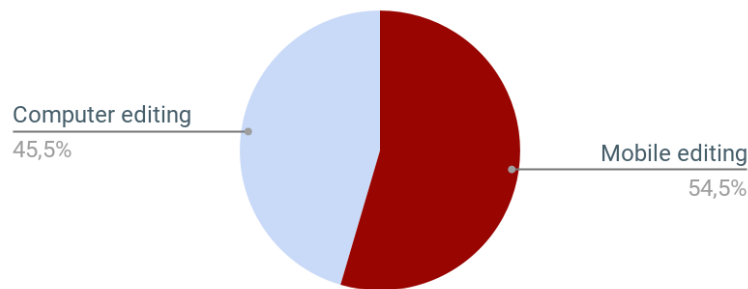
Instead of finding challenges that can be solved by improving Wikipedia, we have again been reminded that many of the marginalized language communities have very different challenges:

- lack of equipment
- unstable or expensive Internet connections
- insufficient technical skills
- relying on mobile devices

Editing from a mobile device is rarely an active choice, but for many, it is the only available option.

> "
>
> The mobile devices do not help members as we mostly use the English Wikipedia articles as a reference to edit in the incubator, and one has to be switching tabs in between. Also, distractions from WhatsApp and other social media platforms are a challenge
>
> Source: *Gurene*

Better mobile solutions would benefit all mobile editors, and for many language communities, that together with good systems for offline editing may be the most efficient action to attract new editors and create more content.

> **"**
> I use a computer. I can't write and upload by phone. Phone screen is too small for me. I always write on the computer. I think some people use phones when they write something. For me I use only the computer when I write articles. I think most of those who contribute to Wikipedia use computers. A few of them may use their phones.
>
> If there is also easier to contribute from the phone, I hope some more people will join it. But we must think about what kind of contribution there will be. I don't think that people can write a good article from their phone. It is not for their eyes.
>
> Source: *Mon*

Economic barriers

Not limited
25,0%

Limited by
75,0%
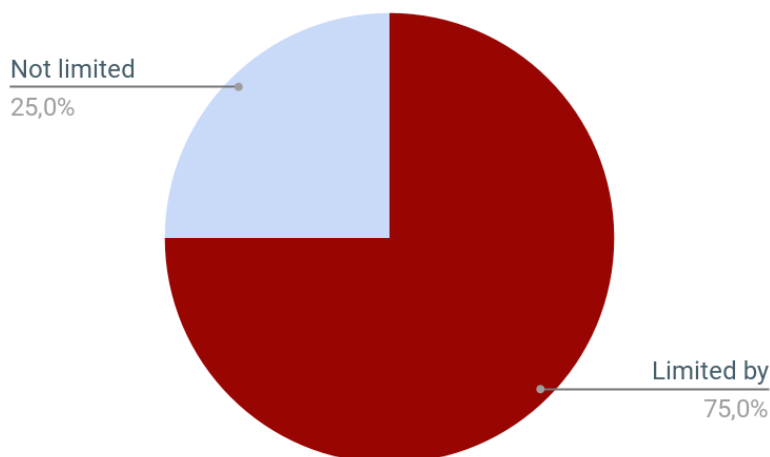
A majority of the targeted groups report that a lack of economic resources is a limiting factor for contributing to Wikipedia. Several reasons have been mentioned; need to work on paid jobs, a lack of culture for volunteering, or simply a lack of funds to buy necessary equipment or data bundles.

Communities that have previously accessed WMF grants describe how it has helped them to grow the community by providing physical spaces with access to the internet and equipment to facilitate contributions.

> ❞
> Yes, community members are not able to edit because of data bandwidth. This includes users of both mobile devices and laptop
>
> Source: *Gurene*

> ❞   I contributed for some time and had to halt it due partly to lack of money to purchase internet data. The economy is a limiting factor for me when contributing to Wikipedia. Contributing to Wikipedia requires internet connectivity data, and that has to do with money. I spend money on purchasing internet connectivity data in order to contribute to Wikipedia
>
> Source: *Dagbani*

Time is precious

The fact that so many communities mentioned time as a limiting factor to their contributions is not a surprise – Wikipedia is endless work for contributors in larger communities, too. However, with 3–4 contributors in total, the sheer workload on each one is enough to demotivate any new contributor to join.

Some interviewees described situations where people had to have two jobs just to make enough to live on. Under those circumstances, editing Wikipedia can never be a priority.

> ❞
> The economic aspect is very important, we are talking about an indigenous community on the Colombian-Venezuelan border, and we see people who are constantly migrating from rural areas to urban areas in order to prepare themselves academically, so when they leave rural areas, the indigenous person has to look for a way to survive.
>
> Source: *Wayuunaiki*

However, not all communities report money as a limiting factor to their contributions. In Inari Sámi the biggest limiting factor is the amount of language speakers that have time and capacity to write, while the Mon community describe a lack of interest for knowledge and learning in the community.

The tradition of volunteering might also be less established in some of the communities, such as described in the Angika community. For some, the skills learned when editing Wikipedia might be considered reward enough, but the region where Angika is spoken is also among the poorest regions of India. The ones contributing to Wikipedia in Angika seem to have higher education, and be in a financially more independent position than most Angika speakers.

> No, I don't think that because of the economy people can't join Wikipedia. In our community older and younger generations there are very few who want to study and learn. That is why we have very few people who are sourcing sources and writing in Wikipedia. In my personal view I don't think that people didn't join Wikipedia because of their economy. People who want to learn something they source and learn. People who don't want to learn don't. They get facebook, it is enough for them.
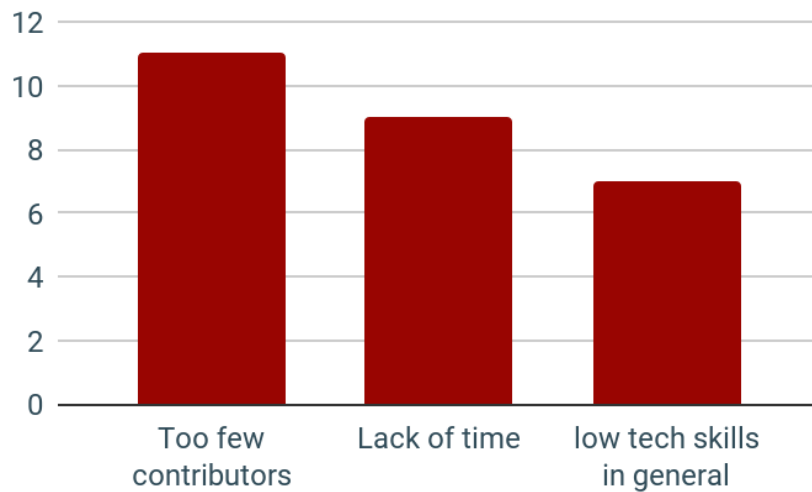>
> Source: *Mon*

> To be very honest, people usually whenever they do something they want something back. Those who are into open knowledge, and are inherently passionate about spreading knowledge maybe not. But in general, people would have this question that: if we do contribute, then what do we get in return?
>
> Source: *Angika*

Social barriers

With social barriers, we wanted to understand the social context in which the contributors operate. Some of the communities are very small because there are few language speakers, such as Inari Sámi. But many of the other languages have many speakers, but they are still struggling to build a significant community of contributors.

Two of the communities report an overweight of female contributors; the rest report that male contributors dominate. Most of them would like a better gender balance. However, with so few contributors in general, the gender issue is not a major preoccupation. The priority was getting more contributors in general.

In many of the language communities, the contributors have never even met, they only collaborate on Wikipedia. For some, there are long distances to travel to be able to meet. The challenge of motivating other people to contribute was mentioned in this phase of the interview.

"

Most of them are men. Women are not so active in our community, even if they are educated. Some people they don't have time to do. But I think, people don't want to use their ability to do.
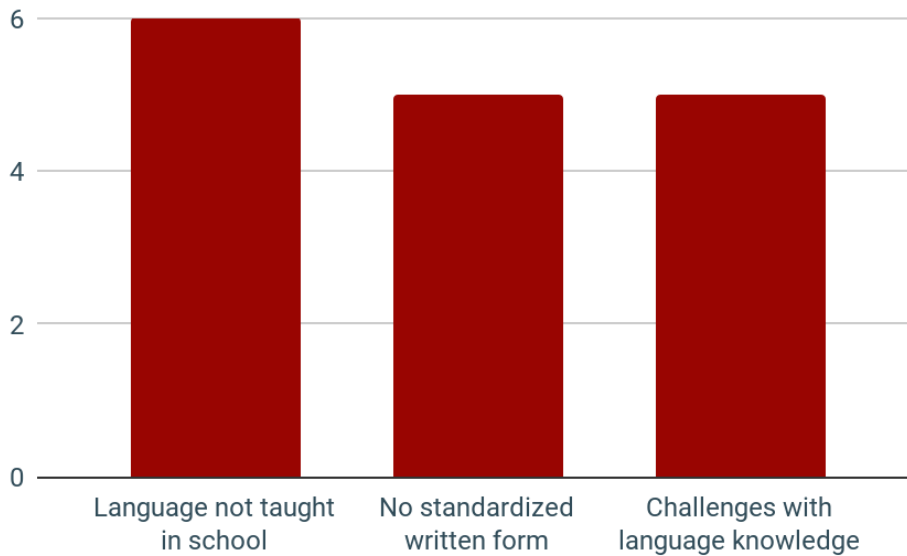
Source: *Mon*

Barriers experienced by contributors to small language versions of Wikipedia

> ”
>
> The majority of contributors are actually women. not very common in the wikipedia world.
>
> Source: *Inari Sámi*

Education and knowledge barriers

Common challenges mentioned by several groups include:

- The language is not taught in schools, or is only taught to a limited degree
- There is no standardized written version of the language
- Challenges with language knowledge in the community

Some of the small language contributors report that their language is not taught in schools, or even if it is, it is not encouraged to be used and is not considered a literate language. For some communities, the language is only used in the first few years of school.

> Tachelicht and the variations of Tamazight was only oral languages. They are written forms, but the written form, the official form at least, came when the language became official in 2011 in a form called Tifinagh. Most people don't know how to write in their own language because they have not learned it in school. Only young children, and only some of them, learn in school. So that is also a challenge, you know how to spell and you can write it in Arabic, Tifinagh or Latin letters, but there is no standard, so people will argue about how to write.
>
> Source: Tachelicht

However, many of the smaller languages report on collaborations with schools, and in some of the groups, teachers or teacher-students are driving forces for the use of Wikipedia in that language.

> ❞   We recorded over 4,000 Dagbani words and successfully uploaded them to Wikimedia Commons. We plan to use these words to create a Dagbani dictionary. This, when created, will be printed and distributed to students for free. We, again, intend to create a pool of references to augment the few sources of references that we have on the internet that relate to Dagbani by documenting our culture and traditions that are usually passed to other generations in the oral form.
>
> Source: *Dagbani*

Five language communities contribute in a language that has no standardized written form. Some languages have developed strategies for negotiating issues of dialect difference, spelling divergence, and the lack of an official language standardization guide (e.g., Scots).

Connected to the lack of standardization, is the limited knowledge of the language among the contributors. This might lead to challenges when translating terminology, as well as with the general quality of the content. A linguist in the network describes a situation where many languages borrow terms directly from English instead of finding their own terms. That can be considered a lost opportunity to develop technological terminology for those languages.

> ❞   I have participated in the discussion of vocabulary translation. We invited other teachers of different dialects to discuss with us. We can use the local language of each place. Everyone has different opinions. Everyone will explain the meaning of this vocabulary and Where did that meaning come from. I think this is quite rich and interesting. I think this activity is very meaningful. I was so impressed
>
> Source: *paiwan*

Strengths of the small language communities

Research based on finding problems and struggles might easily overlook the strengths and opportunities in the populations we are trying to understand. To be able to properly support and encourage the small language communities, it is important to understand what drives them. This is a topic that should be explored further.

Motivation

In Wikipedia in English and probably other larger languages, contributors are motivated by their interest in a certain topic or in helping to bring knowledge to the world. Contributors to smaller language communities, however, are motivated by things like:
- Bringing knowledge about the language and culture to the next generation and to the rest of the world.
- Conserving the knowledge about culture, history and traditions.
- Revitalizing their language.
- Doing something for their culture and their people.

> ” Within all of this, there is also a sense of ownership, and the idea of feeling part of a project that leaves a footprint that institutionalizes a historical memory is also important

> ”
>
> Wikipedia has become a huge name. So knowing that if you look for something and Wikipedia will appear in your language, that validates you, right? Or your language. So that is what most people are looking forward to, I'd say.
>
> **Source:** *Angika*

What role can and should the Language Diversity Hub play?

The barriers we have documented in this work are not new to this movement, and numerous efforts are being made to grow communities, to improve technology, to increase the capacity and the competence of contributors. One of the goals of this research was to get some experience on how the Hub can work together, to become a useful and functional structure within the Movement. The question we keep in mind at each step is: How can the Hub support this challenge, and are there other units within the Movement that can do this better than the Hub?

The language diversity community has to be involved in prioritizing how to direct the next steps of the Language Diversity Hub work. We have therefore suggested next steps under different headings based on the challenges we have documented, and conversations in the Steering Committee. It is a pressing issue for the Hub to create governance and build a democratic and inclusive structure. However, many of the barriers described here can be solved by others than the Hub. This can be considered an open invitation to the whole Wikimedia Movement to think on how you can include support for language diversity into your everyday work.

Training and community building

A part of the goal of this research was community building among the small language communities, to create bridges where they can reach out and support each other. However, almost all the language communities reported that they also need more community building inside their own communities.

They want
- more contributors
- more people using Wikipedia in their language
- more people that understand why the Wikimedia projects are important resources for their language and culture.
- receiving more training

There are several initiatives in the movement for growing and maintaining strong and vibrant communities, as well as providing training sessions, edit-a-thons and so on. It is usually best that the ones offering support are affiliates close to the communities, for practical, sustainable and economical reasons.

The Language Diversity Hub can support the emerging communities via the local affiliates working on community building in the area. If there are no local affiliates,

the Language Diversity Hub can look for ways to support the language communities directly.

Suggested next step
We suggest doing a needs assessment among affiliates in regions of marginalized language communities, to understand their relation with those communities, and gain an understanding of what kind of support they would benefit or need from The Language Diversity Hub to be able to include more language communities in their local affiliate.

The Language Diversity Hub does not need to be directly involved in organizing events, or working actively in local communities. However, giving support for grant applications or planning of such events can be within the scope of the Hub.

Language technology – keyboards, spell checking, and beyond

A recurring issue among all the smaller language communities is that tools for supporting their language online are lacking; they have insufficient keyboards for typing in their own language, or there might be few or no online dictionaries or spell checking software.

There are some on-wiki solutions for keyboards for some languages, and a few dedicated people in the movement have put a lot of time and effort into supporting those languages. However, some of the solutions only work on wiki platforms, and they require maintenance. Besides, these languages deserve to have keyboards available everywhere, on all devices and platforms.

The lack of digital tools is not only a barrier for Wikimedia activities. It is also a barrier for digital presence in general, and it is just one of the existing structures that keep marginalized languages in that status.

Giellatekno[1] (Center for Language Technology) at the University of Tromsø, Norway, has been working for years to build infrastructure to make technical tools for Sámi and other minority languages in Norway. They are interested in sharing this infrastructure, so that even more languages can use it. This would benefit the digital opportunities for those language communities in general, and also facilitate contributors to Wikipedia.

The collaboration has started with building contact between the Giellatekno team and some language communities. However, a systematic approach to making

---

[1] The head of Giellatekno, Trond Trosterud has been an active Wikipedian and serving on the Wikimedia Norge board for years.

this a service offered jointly by the Wikimedia Language Diversity Hub and Giellatekno is interesting to explore.

Suggested next step
Following up on selected languages going through the Giellatekno-path, documenting the process, identifying and supporting other language communities wanting these tools.

Network of institutions and linguists

Offering support to marginalized languages will at times require consultation and support from linguists and other language professionals. For example, when a language is to graduate from the Incubator, it needs to be confirmed by a linguist with expertise on that particular language. Or, when establishing new terminology for a language, linguists and experts on terminology can ensure better processes and end-results.

Many of the groups we spoke with had little digital exposure of their language, or the exposure was mostly in content (such as social media posts) but not in interface, which is needed to get a complete language experience. They also struggled with translating terminology.

Currently, there is an ongoing project to translate the whole of MediaWiki into Northern Sámi. The project is funded by external grants from two different organizations, the Sámi parliament and NUUG-foundation. Both Wikimedians and non-Wikimedians are involved in the translations.

Suggested next steps
There are already many linguists active in the Wikimedia movement, in different roles and capacities. The Hub wishes to build a network of linguists within and outside the movement, as well as institutions with expertise on indigenous and marginalized languages. This network will be a resource for the language communities and the affiliates that are supporting them.

The model of seeking external funding for translation MediaWiki and the Wikipedia application can also be replicated in other languages. The Language Diversity Hub can provide support in the grant application processes, and share experiences from the Northern Sámi experience.

Improving the Incubator

For many contributors to small language versions, the Wikimedia Incubator is the first platform they get to know. All the contributors that have tried both the Incubator and the full-fledged version describe the Incubator as being more difficult to use than the normal Wikipedia. We are essentially welcoming the new language communities to a more difficult platform, without providing much organized and systematic support.

Suggested next steps
We are suggesting two approaches related to the Incubator:

1. Improving the Incubator technically.
   a. Conduct Wishlist survey
   b. Applying for resources to do technical improvements
   c. Following up on Phabricator tasks regarding the Incubator that are not being prioritized
2. Improving the routines of the Incubator
   a. Creating routines for actively welcoming and following up the creators of new editions on the Incubator.
   b. Defining responsibilities between the Language Committee and the Language Diversity Hub

Today, the Language Committee has responsibilities in approving language versions. They also provide support for the new language versions. This can put them in a double role where they both are the helpers and the judges. The [Language Committee](#) are volunteers, while within the Wikimedia movement the [Language Team](#) are employees working on language related technology. Today participants from the Language Team and the Language Committee are also part of the Steering Committee. While improving the incubator, which is something that anyone with the skills can embark on, this project would be a good opportunity to explore further what role the Hub can and should play in relation to these other units.

Expanding on this work

This project has been a chance to get to know contributors, and for contributors to get to know how the movement works. Many smaller communities might experience problems where there already are solutions, they just don't know about them. Through the outreach we were able to provide immediate support, or connections that could be useful for the future and bridges were built.

This has also provided us with a baseline to later evaluate if the situation has improved for the 13 language communities in this group, and also to evaluate how

the situation has changed on a global scale for the new and smaller language communities in the future.

Suggested next step

As already described under [training and community building](), support for the marginalized communities should be offered by an organization close to them. The responsibilities of replicating this work could be with the local affiliate, while the Language Diversity Hub can provide support before the interviews and in following up on the challenges that are surfaced.

Many of the marginalized languages use other Wikimedia projects, such as Wikisource or Wikimedia Commons. Replicating interviews to focus on the use of these platforms can provide more insight into how Wikimedia projects beyond Wikipedia can be used in projects that aim to strengthen the digital presence of or revitalizing marginalized languages.

Building Hub governance

The first Steering Committee was designed by invitation of people that have done important work for language communities in the past, and consisted of 10 people. Since then, the Steering Committee has been open for interested parties to join as observers. Ten more persons have been included as observers, either because they wanted the Hub to be a part of their project, or because they wanted to support the Hub with their knowledge and skills. This means the Hub has been led by people who already belong to the international part of the movement. For the future the Hub needs to better represent the marginalized communities.

So far, all the decisions have been made by consensus reached at the meetings, or in the Telegram group. Until now this has worked well, there is a common sentiment that the Hub is needed, and that any work to promote the new and the smaller language communities is important. However, for the future development of the Hub, it will need a Steering Committee with a stronger leadership that takes hard discussions, and has the mandate to set a direction and priorities for the work of the Hub. As we have seen in this work, there are many challenges the Hub can take a part in solving, but not all at the same time.

The Steering committee has been composed of people with a strong personal and professional interest in the field of smaller languages. Knowing that economy is a barrier for some to engage in the movement, it will be important that there are allocated resources to make sure people representing the new language communities have a real possibility to work with the Hub.

A challenge that has surfaced is how to organize a fair selection process of the new Steering Committee. Many of the groups that need the support of the Hub the most might not even have a User Group, and there might be few contributors

from that language. Time, capacity and language skills might all be barriers for some to take part in the governance of the Hub. We wish to work on how to create a governance structure that ensures that the smaller communities can be involved and especially that their challenges are heard, but also that they get access to share their experiences and be a part of the global movement without draining on their already sparse resources.

Suggested next step
Setting down a group working on how to create governance of the hub:
- Defining the role of the hub in relation to the other language groups within the movement.
- Exploring how to compose a inclusive and representative steering committee

Concluding remarks

Is a Language Diversity Hub needed in the Wikimedia movement? With a Language team dealing with the technical parts of language diversity, and a Language Committee dealing with onboarding and implementation of new language editions of Wikimedia projects, what is left for the Hub? I would like to come back to the words of Delphine Ménard, describing the hubs as something that fills the cracks and spaces between the existing structures in the movement. There is a gap in this movement related to the responsibility for supporting emerging communities on a movement-wide level, and that is the main gap the Hub is addressing.

We have learned that some of the more challenging barriers might be the ones outside the Wikimedia universe. We believe that the Hub can contribute to collaborations with institutions and organizations outside the movement to overcome these too. The language communities can benefit from these collaborations on and outside Wikimedia projects.

This movement is one of remarkable people, and we find many of them in the global language diversity community. Motivated by the fear of losing knowledge about the culture, history, traditions and even language of their people, the sense of urgency and importance among the contributors in the emerging communities is remarkable. Together the language diversity community provides quite a few of the missing pieces of the puzzle to reach the sum of all human knowledge available.