Through their eyes

The online experience of three European indigenous language communities

Valentina de Vito

20th November 2019
Introduction

This study has been commissioned by EURid to support the UNESCO International Year of Indigenous Languages 2019, and as part of the long-term collaboration between the two organisations to enhance linguistic diversity in cyberspace.

The study highlights the challenges faced by speakers of indigenous languages in locating and creating online content in their own languages. It provides three case studies of European indigenous languages: Catalan, Saami and two variants of Italian Grecanic, Salento Griko and Calabrian Greko. This paper describes both the positive and negative impacts of current technology affecting indigenous language communities. It analyses the social and technical barriers faced by indigenous language communities when engaging in language revitalisation efforts. It identifies factors that contribute to the adoption, adaptation or abandonment of a language.

In principle, everyone has a right to access the Internet.\(^1\) In reality, there exists an enormous gap between the abundance of spoken languages in the offline world and the languages in cyberspace. The European Union is linguistically diverse, comprising three percent of world languages that are spoken offline. However, such variety is not reflected online. Globalisation has paved the way for the rise of English in the last few decades as a global lingua franca\(^2\) at the expense – particularly – of indigenous and minority languages. Speaking a dominant language is now essential for full civic participation and economic advancement. Such necessities have been prioritised over the efforts of preserving minority languages.

According to UNESCO, of the circa seven thousand languages spoken worldwide, 2680 are endangered.\(^3\) In Europe, there are 128 languages at risk of extinction – though with various degrees of endangerment.\(^4\)

The digital age represents both a challenge and an opportunity for indigenous and endangered languages and, while communication technology has been a huge force in the propagation of majority languages like English, the online environment can also provide opportunities to strengthen and raise awareness of minority languages.\(^5\)

UNESCO’s role in supporting linguistic diversity online

In 2003, with the adoption of the Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace, UNESCO encouraged its Member States to develop comprehensive language-related policies to allocate resources and use appropriate tools in order to promote and facilitate linguistic diversity and multilingualism, including the Internet and media. The Recommendation supports the inclusion of new

---

1 Resolution A/HRC/32/L.20
languages in the digital world and the creation and dissemination of content in local languages on the Internet, and encourages multilingual access to digital resources in cyberspace.⁶

Since 2005, UNESCO has been the international organisation responsible for reporting on implementation of WSIS action line C8 on Cultural Diversity and Identity, Linguistic Diversity and Local Content.⁷

The survival of Indigenous languages is essential for the preservation of humanity’s cultural heritage. The United Nations General Assembly has declared 2019 as the International Year of Indigenous Languages to raise awareness of and support the work of indigenous language communities.

A language is considered endangered when its speakers cease to use it, use it in fewer and fewer contexts, use fewer of its registers and speaking styles, and/or stop passing it on to the next generation. No single factor determines whether a language is endangered, but UNESCO experts have identified nine interdependent factors that impact the vitality of a language: intergenerational language transmission, absolute number of speakers, proportion of speakers within the total population, shifts in domains of language use, response to new domains and media, availability of materials for language education and literacy, governmental and institutional language attitudes and policies including official status and use, community members’ attitudes toward their own language, and the amount and quality of documentation.⁸

**UNESCO and EURid partnership - Internationalised Domain Names (IDNs)**

Since 2011, UNESCO and EURid have been collaborating to provide an evidence-based study on the impact of Internationalized Domain Names (IDNs)⁹ on online linguistic diversity, through the annual World Report on Internationalised Domain Names.

The evidence shows that IDNs help to enhance linguistic diversity in cyberspace.¹⁰ Despite ongoing technical challenges in their development and adoption (a lack of universal acceptance),¹¹ IDNs constitute a vital step towards achieving a multilingual Internet by enabling users to navigate the internet in their native language.¹²

IDNs are both a key to accessing content online¹³ and a proxy for multilingualism and language rights on the Internet¹⁴ but “the internationalisation of domain names is [...] just the top of the iceberg. It is not only the domain names that have to be internationalised but also all the necessary tools to take advantage of them, such as search engines, browsers, e-mail servers, office suites, operating systems and so on.”¹⁵

Thus, despite the general consensus concerning the positive role of technology in the revitalisation of indigenous and endangered languages, namely in preserving, analysing,
manipulating and transmitting languages, efforts made at local, regional, national and international level still need to be intensified and coordinated more efficiently in order to achieve a truly multilingual cyberspace.

**Methodology**

In order to understand the different status of indigenous and endangered languages in the European Union, this study selected for observation one large, one medium and one small sized community – the Catalan, Saami and Griko/Greko communities. Not all indigenous languages are also endangered. Some, such as Catalan (one of the languages in this study) are widely spoken. As previous studies have observed, even widely spoken languages such as Catalan or Basque are in a high-risk category in terms of their future sustainability. Languages with a higher degree of endangerment face even tougher challenges in the online environment.\(^\text{16}\)

This study has used a combination of qualitative interviews from indigenous language community representatives and experts and desk-based research to gain insights into languages identified by UNESCO as endangered. A total number of seven qualitative interviews were undertaken during the period of August-September 2019, using a combination of an online questionnaire and telephone conversations. Interviewees were identified through their representation of the three indigenous language communities identified for study and expertise in the online environment. One interviewee requested to remain anonymous, and will be referred to as “Representative for the Saami Language Community”. The other interviewees are identified.

**Case studies:**

**Catalan**

According to Ethnologue, Catalan is a statutory provincial language, with has nearly 9 million speakers.\(^\text{17}\) There is anecdotal evidence that up to half of those speak Catalan as a second language.\(^\text{18}\) Catalan speakers are mainly based in the eastern coast of Spain, largely in the Autonomous Communities of Catalonia, parts of the Valencian Community and the Balearic Islands, with smaller Catalan communities residing in the Autonomous Communities of Aragon and Murcia. Outside of Spain, Catalan is spoken in Northern Catalonia (corresponding to much of the French department of Pyrénées-Orientales), in the Sardinian city of L'Alguer (Italy) and in Andorra where it is the only official language.\(^\text{19, 20}\) Catalan is co-official alongside Spanish in the Balearic Islands, Valencia and Catalonia. It enjoys some recognition by both the Italian and the Sardinian government in L'Alguer, but it remains unrecognised everywhere else.\(^\text{21}\)

Catalan is widely spoken by all age groups in Catalonia and the Islands, but mostly by the elderly elsewhere (though it is quite vital still in rural areas of Aragon and Valencia). It is the main language of education in Catalonia and the Balearic Islands, and one of the languages in Valencia; in other regions where Catalan is spoken, majority languages such as Spanish, French or Italian are the language of education. Although the language is vibrant in Catalonia,

\(^\text{16}\) META-NET’s Strategic Research Agenda for Multilingual Europe 2020 (SRA) (2012).

\(^\text{17}\) https://www.ethnologue.com/language/cat

\(^\text{18}\) Guillem Belmar (2019), Email interview.


\(^\text{21}\) Guillem Belmar (2019), Email interview.
the Balearic Islands and Valencia, elsewhere Catalan is receding rapidly. In both Northern Catalonia and in L’Alguer the local variants are almost extinct and have been replaced by either French/Italian or Standardised Catalan.22

Despite not being considered as a minority language for several reasons, including above all its legal status and its demography, Catalan shares some of the characteristics of minority languages: “the practical inexistence of monolingual people and therefore, bilingual population, the fact that the territories of its linguistic region belong to far larger states where the majority speak a different language, or the lack of presence in some sectors of social life”.23

The many political borders dividing the Països Catalans – the geo-historical regions where Catalan is spoken – both at the state and local levels have had a strong impact on the development of standard varieties in the past and thus, to the creation of multiple standards which are mutually intelligible.24 However, academic institutions, notably the Institut d'Estudis Catalans (IEC), have undertaken research and study into the Catalan culture. The IEC’s principal work is in standardising the Catalan language.25

One of the identified issues for Catalan is a lack of written proficiency among those who did not learn the language during education and have then learned it as adults.26 In fact, the Franco regime (1939-1975) “prohibited the use of the Catalan language in education, in the publishing of books, newspapers and magazines, and the transmission of telegrams and telephone conversations in Catalan. […] Administrative, notarial, legal and business documentation produced in Catalan was considered legally null and void.”27 Thus, people who were born in that period did not have opportunities to learn the Catalan language, beyond voluntary and semi-clandestine initiatives.28

Catalan uses the Latin script with accents and diacritics such as à é í ò ó ú ü ç l·l. It appears among the 104 world languages supported by Google Translate, though translations occasionally contain errors. Most of the mainstream platforms have been localised in Catalan. This contributes to its extensive use in social media, such as Facebook and Twitter, where several groups and/or profiles related to hobbies, fandom, the language itself or variants of it are present. According to our Catalan interviewee, Catalans do not report any difficulties in finding online content in Catalan which is ranked 20th in Wikipedia’s ranking by number of articles.29

Although Catalan has its own keyboard, supporting its character set, settings of technological devices are generally not available in Catalan. This does not present a major drawback to members of this language community, as the special characters are also available in other language keyboards, such as Spanish, German or French.

22 Guillem Belmar (2019), Email interview.
28 Guillem Belmar (2019), Email interview.
29 Guillem Belmar (2019), Email interview.
30 https://meta.wikimedia.org/wiki/List_of_Wikipedias
Saami

Saami languages are a group of Uralic languages spoken in Northern Europe. They are divided into two main groups: western and eastern, which are further divided into various subgroups and individual languages. Neighbouring Saami languages are mutually intelligible whereas speakers of more widely geographically separated areas may not be able to understand one another without exposure to the other language.

North Saami – the most spoken variant of Saami – has around 20,000 speakers in the Northernmost parts of Norway, Sweden and Finland, about two-thirds of the full Saami population. Saami speakers mostly also speak Norwegian, Swedish, Finnish, and English. According to UNESCO’s Language Vitality and Endangerment framework, North Saami has been classified as ‘definitely endangered’. This, in terms of intergenerational transmission, means that children no longer learn the language as a mother tongue in the home. In the words of Jin Sook Lee and Eva Oxelson, “losing proficiency in their heritage language is more than just a loss of a linguistic system; it is a separation from their roots, a denial of their ethnic identity, and a dismissal of their potential as a bilingual and bicultural member of society.”

The Saami population is becoming more scattered and influenced by increasing mobility and the concentration of the language speakers does not always follow the trajectories of the learners. Thus, the physical distance between Saami speakers, and of Saami natives from cultural centres and traditional settlement areas, is identified as one of the issues currently affecting the language community. The pedagogical approach to this proposed by Hanna Outakoski (2014) recognises the potential of 3D virtual environments as real-life-resembling indigenous learning spaces, whilst also identifying technical, financial and epistemological challenges.

Saami uses accents and diacritics (á ž č š đ ŧ), which are also present on keyboards. Settings of technological devices are generally available in Saami. The Saami character set works with all of the popular social media platforms, but only Facebook – which is the most used together with Instagram and Snapchat – has localised its menu and functions in Saami. Saami is not supported by Google Translate. Members of Saami community use all digital tools – email, websites, social media, chat apps, video conferencing – to communicate and coordinate. However, in order to reach a wider audience, “often posts in North Saami are supplied with translation into of the majority languages (Norwegian, Swedish, Finnish) or English”. Finding online content in Saami remains the major issue for Saami natives; in particular, Wikipedia only hosts around 7,500 articles in Northern Saami, which is ranked 156th in Wikipedia’s rankings by articles.

Griko (Calabria) / Greko (Salento), Southern Italy

32 https://www.omniglot.com/writing/saami.htm
37 Representative for the Saami Language Community (2019), Email interview.
38 Representative for the Saami Language Community (2019), Email interview.
39 https://meta.wikimedia.org/wiki/List_of_Wikipedias
The Greek community of Salento (Apulia) is made up of the following villages: Calimera, Castrignano dei Greci, Corigliano d'Otranto, Martano, Martignano, Sternatia, and Zollino. These villages together have a population of around 40,000. Calabrian Greek is mostly spoken in five villages on the Ionian coast of the metropolitan city of Reggio Calabria, Italy. The area is known as Area Grecanica, though it comprises many non-Grekophone municipalities.

UNESCO indicates that the number of speakers for Greko (Salento) is approximately 20,000; Calabrian Griko is spoken instead by circa 2,000 people. However, the lack of recent reliable data makes the actual estimate of the number of speakers very difficult.

Both variants of the language spoken in the South of Italy are classified as severely endangered by UNESCO, which means that the language is spoken by grandparents and older generations (older than 65). While the parent generation may understand it, they do not speak it to children or among themselves. Thus, “Greko lost its communicative value as language to be used in everyday interactions”.

The two variants of Italian Grekanic are mostly written in Latin script, though some people still write it in Greek script. The lack of standard orthography can pose problems and disputes from time to time when engaging online. Thus, since an alphabet has never been codified, speakers eventually end up writing as they see fit.

Italian Grekanic is not supported by Google Translate in any of its variations and Italian is used on mainstream platforms to communicate and coordinate. Settings of tech devices and menu systems of the major platforms are not available in Griko because of the highly difficult adaptation of the language to the changing reality. This does not seem to constitute a problem as both younger and adult speakers are able to use device settings in Italian.

Our interviewees confirmed that the greatest challenge for Griko/Greko speakers is in finding online content. Less complicated is the possibility of building human connections online and to interact on online platforms. In fact, it is mainly through social media that members of the Griko/Greko communities coordinate their language revitalisation efforts. Griko/Greko natives try to keep their language of origin alive by sharing poems, traditional songs, story tales, sayings through social media, which are also the main platforms where new speakers and learners can practise the language.

Linguist Maria Olimpia Squillaci – who learnt Calabrian Greko at home from her father – analyses in a recent report positive steps taken to revitalise the Greko language through an indigenous language association. Attempts to revitalise the language since the 1960s brought about a significant change in attitude of the community towards its own language, making it a source of pride. However, those efforts were unable to secure an intergenerational transmission of Greko, and the number of speakers has further decreased. The organisation Jalò Tu Vua took the challenge back on recently by organizing intensive language courses combining teaching, cultural seminars and excursions as well as Tandem activities carried out

http://www.unesco.org/culture/en/endangeredlanguages/atlas
41 Maria Olimpia Squillaci (2019), Email interview.
http://www.unesco.org/culture/en/endangeredlanguages/atlas
43 Maria Olimpia Squillaci (2019), Email interview.
44 Maria Olimpia Squillaci (2019), Email interview.
45 Giuseppe De Pascalis (2019), Email interview.
46 Maria Olimpia Squillaci, Daniele Palma and Giuseppe De Pascalis (2019), Email interviews.
47 Maria Olimpia Squillaci, Daniele Palma and Giuseppe De Pascalis (2019), Email interviews.
through Skype. A WhatsApp group and the Facebook page To Ddomadi Greko were created to practise the written language and a dictionary smartphone application called ‘Grekopedia’ was developed. In less than a year Jalò Tu Vua has managed to build up a community of practice which also includes Calabrian diaspora and new learners and which is also provided active support of many older Greko speakers. In order to make such efforts sustainable over time, crowdsourcing campaign “If you speak me, I live - adopt Greko” was launched and it is still ongoing.48

Squillaci goes further in her report by sharing three main steps considered crucial when undertaking language revitalisation efforts. Whether people start learning a minority language to deepen their cultural heritage, out of curiosity or for research purposes, the elements that make a difference are the feeling of emotional connection to the language, the presence of a team in which motivation and practice can be mutually nurtured and the possibility to “foresee social and economic outcomes that can result from speaking the language and that can lead to concrete improvements for the community”.49

Key findings

- Technology – especially communication technology – is seen as helpful in the revitalisation of indigenous languages by all our interviewees. It offers speakers the chance to build online communities to meet up, learn about, and practise the language with their fellows,50 thus enhancing a sense of belonging to the community.

- “As most young people approaching their heritage language live away from home, communication technology is fundamental to help them maintain a constant contact with their origins” and to encourage intergenerational dialogue between young and their older relatives at home.51

- Public exposure of indigenous languages online may further its diffusion. In the words of the Saami language representative, “languages must be heard but also seen physically, to increase the motivation of learners and the acknowledgement of other communities”.52 However, two practices in particular may weaken any effort in this sense by creating confusion in learners: the use of Internet slang, especially on social media platforms and the lack of standardised orthography which characterises some minority languages.

- The frequent use of code-switching – i.e. loan words from other languages which are sometimes phonologised – especially in the field of technology, discourages the adaptation of a language to new contexts, which hasten the language’s obsolescence. Furthermore, the need to reach broader audiences, “may be pushing users to use a major language on the Internet to gain more followers”,53 even if they are not entirely fluent in that language.

50 Guillem Belmar (2019), Email interview.
51 Maria Olimpia Squillaci (2019) Email interview.
52 Representative for the Saami Language Community (2019) Email interview.
53 Guillem Belmar (2019), Email interview.
• Despite Google and other platforms’ goal to expand the number of supported languages, there still are challenges to do with inclusivity, particularly when many less-spoken languages remain in oral form only or without a standardised orthography, as is the case for Griko and other minority languages.

• Machine translation, voice recognition and language learning in different and emerging media forms – such as podcasts and interactive digital literacy platforms – have been identified by one of our interviewees as useful new technologies for the revitalisation of languages.54

• Next to these technologies, traditional forms of culture diffusion such as movies, TV shows and series “in indigenous languages, made by indigenous people and based on indigenous peoples’ culture, traditions and world views”,55 are believed by our interviewees to still be effective means through which a language – and its culture – can be kept alive.

• Technology and social media can be a means of language revitalisation but they alone cannot solve the erosion of cultural heritage through loss of endangered languages.

• “The main problem with technology for language revitalisation is the access to technology some communities have (or lack thereof)”.56 Unlike speakers of European endangered languages, “many other indigenous groups around the globe, whose language may also be endangered, still lack easy access to technology, which makes all its benefits useless to them”.57

• Our study found a general lack of awareness about the existence of IDNs, and of domain names with accents and different scripts among our interviewees.

Analysis and recommendations:

Steps to be taken towards a multilingual cyberspace

Our concluding comments and recommendations are set out below, and build on the comments and insights shared by our interviewees.

• Urgent, coordinated action by multiple stakeholders is necessary to reverse the decline in indigenous languages. New broader multifaceted long-term strategies must be discussed and developed, with the close involvement of local, national and international stakeholders interested in revitalising a particular language.

• National legislation should include indigenous peoples’ rights and mechanisms for their implementation and enforcement based on good practices already present within language communities in Europe and abroad.

• Member States should be encouraged at an international level to produce national strategic plans for the promotion of endangered languages and to make continuous

54 Eddie Avila (2019), Email interview.
55 Representative for the Saami Language Community (2019) Email interview.
56 Guillem Belmar (2019), Email interview.
57 Guillem Belmar (2019), Email interview.
funding available for endangered language communities in order to support them in their revitalisation efforts.

- Endangered languages have to be used for recreational activities both in their community contexts as well as online, in order to persuade young people that their languages are useful, relevant and desirable, since young people’s involvement constitutes the basis for the intergenerational transmission of languages.

- In the wake of the Facebook initiative called ‘Community Translations’, which enables any language community to add their language to Facebook and make it publicly available, other social media platforms should allow their users to do so. In fact, thanks to community work, Facebook supports nine EU endangered languages: Corsican, Faroese, Irish, Welsh, Frisian, Belarusian, Basque, Breton, and Sardinian – and six more are in progress. “More than 40,000 people are accessing Facebook in endangered languages, catapulting these languages into the 21st century and making them ‘cool’ again.”

- An overall shift in attitudes towards the importance of maintaining language is fundamental. Pride, a sense of belonging to a minority and the awareness of the value and cultural wealth that this constitutes should be stimulated by informed political elites. Technology can support a shift in attitude in this sense as captivating and well-designed online promotion campaigns made at national and local level can stimulate users’ interest in language courses or cultural activities.

- Endangered languages must be incorporated into the daily language repertoire, including online spaces. All responsible actors are encouraged to make technological tools available to users which can support the use of indigenous languages online.

- Awareness of IDNs must be raised by all stakeholders and urgent steps are required to improve their functionality and to achieve universal acceptance. Often times, it does not seem to be “a matter of creating new technologies, rather a matter of making what we have equally accessible for everyone.”

To conclude, language profoundly affects people’s experience on the Internet. To say it in Ludwig Wittgenstein’s words re-adapted to today’s reality, “The limits of my language online mean the limits of my world”.

---

59 Guillem Belmar (2019), Email interview.
Interviewees

- Carson Kiburo, Co-Chair UN Inter-Agency Network for Youth Development (Kenya).
- Daniele Palma, Native Speaker and Contact Person for Kaliglossa (http://www.kaliglossa.it/) (Griko, Salento - Italy)
- Eddie Avila, Director at Global Voices Online (https://globalvoices.org/) and Rising Voices (https://rising.globalvoices.org/).
- Guillem Belmar, Socio-linguist and Lecturer, University of Groningen (Catalan - Spain)
- Giuseppe De Pascalis, Native Speaker (Griko, Salento - Italy)
- Maria Olimpia Squillaci, Linguist and Activist, Jalò Tu Vua (https://www.jalotuvua.com/) (Greko, Calabria - Italy)
- Representative for the Saami Language Community (Asked to remain anonymous)

About the author

Valentina de Vito is attending her last year of the joint master’s degree in International Security Studies at Sant’Anna School of Advanced Studies (Pisa) and the School of International Studies (University of Trento). She holds a bachelor’s in Political Science and International Studies (University of Florence) and a BA (Hons) in Journalism (LCC, London). Through her studies, she has developed a strong research interest in cybersecurity, cyber warfare and Internet governance. Her previous work experience is in the field of communication and research.

Acknowledgements

This report has been fully funded by EURid for inclusion in the World Report for Internationalised Domain Names 2019, as part of EURid’s support for UNESCO’s International Year of Indigenous Languages.