

Linguistic and Cultural Diversity

IN CYBERSPACE

Proceedings of the 2nd International Conference



United Nations
Educational, Scientific and
Cultural Organization



Information for All
Programme

Commission of the Russian Federation for UNESCO
Ministry of Culture of the Russian Federation
Russian Committee of the UNESCO Information for All Programme
Interregional Library Cooperation Centre

Linguistic and Cultural Diversity in Cyberspace

Proceedings of the 2nd International Conference
(Yakutsk, Russian Federation, 12-14 July, 2011)

Moscow
2012

Financial support for this publication is provided by the Ministry of Culture of the Russian Federation

Compilers: Evgeny Kuzmin and Anastasia Parshakova

Translators: Tatiana Butkova, Ekaterina Komarova and Anastasia Parshakova

English text edited by Anastasia Parshakova

Editorial board: Evgeny Kuzmin, Adama Samassekou, Daniel Prado, Evgenia Mikhailova, Sergey Bakeykin, Anastasia Parshakova, Tatiana Murovana, Nadezhda Zaikova, Liudmila Zaikova

Linguistic and Cultural Diversity in Cyberspace. Proceedings of the 2nd International Conference (Yakutsk, Russian Federation, 12-14 July, 2011). – Moscow: Interregional Library Cooperation Centre, 2012. – 312 p.

The book includes communications by the participants of the 2nd International Conference Linguistic and Cultural Diversity in Cyberspace (Yakutsk, Russian Federation, 12-14 July, 2012), which became a unique platform for discussing political, cultural, educational, ideological, philosophical, social, ethical, technological and other aspects of the activities aimed at supporting and preserving languages and cultures and promoting them in cyberspace.

The authors present linguistic situation in different countries, share international best practice in the policies, standards and instruments for the preservation and development of linguistic and cultural diversity.

The authors are responsible for the choice and presentation of facts and for the opinions expressed, which are not necessarily those of the compilers

ISBN 978-5-91515-048-7

© Interregional Library Cooperation Centre

Contents

Preface	7
Messages	9
Irina BOKOVA, UNESCO Director-General.....	9
Grigory ORDZHONIKIDZE, Executive Secretary of the Commission of the Russian Federation for UNESCO	11
Alexander AVDEYEV, Minister of Culture of the Russian Federation.....	13
Andrei FURSENKO, Minister of Education and Science of the Russian Federation,	15
Egor BORISOV, President of the Republic of Sakha.....	16
Abulfas KARAYEV, Minister of Culture and Tourism of the Azerbaijan Republic.....	17
Evgenia MIKHAILOVA, Rector of the North-Eastern Federal University	18
Plenary meetings	19
Adama SAMASSEKOU. Developing Institutional and Legal Instruments to Support Multilingualism Worldwide	19
Daniel PRADO. It Is Time to Place Multilingualism and Linguistic Diversity at the Heart of the International Debate.....	22
Evgenia MIKHAILOVA. Our Common Goal is to Preserve Not Only Our Language, but also Our Culture, Environment and People	25
Evgeny KUZMIN. Social Institutions Supporting Linguistic and Cultural Diversity in Cyberspace: Roles, Functions, Responsibilities	32
Yoshiki MIKAMI, Katsuko T. NAKAHIRA. Measuring Linguistic Diversity on the Web	46
Daniel PIMIENTA. DILINET: Building Indicators for Linguistic Diversity in the Internet.....	56
Dwayne BAILEY. Multilingual Moments of Truth: Moments that Define a Person’s Future Perception of Multilingualism	59
William W. MCLENDON, Jr. New Developments and Trends to Facilitate Multilingualism in Cyberspace.....	66

Daniel PRADO. Language in the Virtual World and in the Real World: Parallels to Take into Account in Language Policy	72
Sozinho F. MATSINHE. African Languages in Cyberspace: Challenges and Prospects	80
Gilvan MULLER DE OLIVEIRA. Portuguese and the CPLP Languages in Cyberspace.....	89
Javier LOPEZ SANCHEZ. Revitalization, Strengthening and Development of National Indigenous Languages of Mexico in Cyberspace	95
J. Prabhakara RAO. Multilingualism in India and Indian Languages in Cyberspace.....	100
Natalia GENDINA. The Importance of Information Literacy in the Development of Multilingualism in Cyberspace	107
Valery YEFIMOV. Preserving Linguistic and Cultural Diversity in the Republic of Sakha (Yakutia): Foresight Project	116
Section 1. Instruments for Language Preservation and Promotion in Cyberspace.....	123
Marcel DIKI-KIDIRI. Creating Contents for Less-Diffused Languages on a Web Site.....	123
Virach SORNLERLTLAMVANICH. Linguistic and Cultural Knowledge Co-Creation on Social Network	127
Yunseok RHEE. A Web-Based Participation Model to Preserve Linguistic and Cultural Diversity	131
Tunde ADEGBOLA. Multilingualism, Multimedia and Orature in the Information Age: Using Multimedia to Include Oral Cultures in Cyberspace	136
Nicholas THIEBERGER. Multilingualism in Cyberspace – Longevity for Documentation of Small Languages.....	141
Dietrich SCHÜLLER. Audio and Video Documents at Risk: Safeguarding the Documents Proper of Linguistic Diversity and Orally Transmitted Cultures.....	151
Adel EL ZAÏM. Cyberactivism in the Middle East and North Africa: Promoting Local Languages.....	158

Elmir YAKUBOV. Developing Digital Content in the Languages of Dagestan: A Means of Overcoming the Increasing Marginalization of Local Languages	165
Sergey BOBRYSHV, Emil YAKUPOV. Multilingual Fonts for Visual Communications.....	168
Emil ASANOV. Localization of Free Software as a Means to Support Multilingualism and Information Literacy Development of the Population of Kyrgyzstan.....	174
Dzhavdet SULEYMANOV. The Tatar Language and ICTs.....	177
Sergey VASILIEV. Yakut Epic in Cyberspace: Olonkho Information System and Olonkho.Info Multilingual Portal	185
Lyubov BADMAYEVA, Yulia ABAYEVA. Modern Linguistic Technology to Help Preserve Buryat Language: Corpora & Database Development	192
Anatoly ZHOZHNIKOV, Svetlana ZHOZHNIKOVA. Creating a Web Portal on Indigenous Cultures of North-Eastern Russia.....	195
Section 2. Institutes to Promote Linguistic and Cultural Diversity.....	199
Yan-Henry KESKITALO. University of the Arctic's Contribution to Maintaining Linguistic and Cultural Diversity in the Arctic: Improvements Through Networking and Partnership.....	199
Sergey BAKEIKIN. Promoting Linguistic and Cultural Diversity in Cyberspace: Activities by the Russian Committee of the UNESCO Information for All Programme and the Interregional Library Cooperation Centre.....	209
Claudia WANDERLEY. Multilingualism and Digital Libraries in Local Languages	214
Zhanna SHAIMUKHANBETOVA. The Role of Kazakhstan's Libraries in Shaping the Country's Linguistic Landscape	225
Irina DOBRYNINA. Creating Content in Minority Languages: Enhancing Users' Capacity	231
Anastasia PARSHAKOVA. Elaborating Practical Guidelines for the Stakeholders of Multilingualism Development.....	236

Valentina SAMSONOVA. Preservation of Cultural Heritage of Indigenous Small-Numbered Peoples of the Russian North, Siberia and the Far East	240
Tjeerd DE GRAAF. The Foundation for Siberian Cultures and the Use of Information Technology for the Preservation and Teaching of Indigenous Languages in Siberia.....	246
Katsuko T. NAKAHIRA. Usage of Mother Tongue in Education in the Case of Multilingual Activities with ICT	251
Anatoly ZHOZHNIKOV. North-Eastern Federal University: Strategic Action Lines to Promote Cultural Diversity of Indigenous Communities of the Russian North in Cyberspace.....	261
Liudmila ZAIKOVA. Establishing the Centre to Advance Multilingualism in Cyberspace under the North-Eastern Federal University	265
Syrgylana IGNATYEVA. Arctic Arts & Culture Institute: Creating Online Resources on Russia’s Northern Indigenous Communities’ Culture	269
Section 3. Creating Positive Environment for the Promotion of Linguistic and Cultural Diversity	272
Dieter W. HALWACHS. Linguistic Diversity, Dominated Languages and the Internet: The Case of Romani.....	272
Fikret BABAYEV. Azerbaijani State Policy Toward the Development of Intercultural Dialogue and the Preservation of Cultural Diversity	282
Nikolay PAVLOV. Role of Modern ICT in Keeping Sakha Language Afloat	288
Viktoria GOLBTSEVA. Role of Modern Technology in Preserving Linguistic and Cultural Diversity of Indigenous Chukchi Peninsula Communities	294
Vyacheslav SHADRIN. Yukaghir Language and Culture in Cyberspace	300
Final Document. Yakutsk Call for Action: A Roadmap Towards the World Summit on Multilingualism (2017)	304
Lena Resolution	306

PREFACE

The 2nd International Conference on Linguistic and Cultural Diversity in Cyberspace took place on 12-14 July, 2011, in Yakutsk, Russian Federation. It became one of the key events within the framework of the Russian chairmanship in the UNESCO Information for All Programme.

The event was organized by the Russian Committee of the UNESCO Information for All Programme, the North-Eastern Federal University, the Interregional Library Cooperation Centre, the MAAYA World Network for Linguistic Diversity, and Latin Union with the support of the Ministry of Culture of the Russian Federation, the Commission of the Russian Federation for UNESCO, and UNESCO.

The conference gathered about 100 experts from about 30 countries of all continents, including Australia, Austria, Azerbaijan, Belarus, Belgium, Brazil, Cape Verde, Central African Republic, China, Dominican Republic, Egypt, France, India, Japan, Kazakhstan, Korea, Kyrgyzstan, Mali, Mexico, Moldova, Netherlands, Nigeria, Norway, Russia, South Africa, Thailand, UK, Ukraine, USA.

The First Conference on Linguistic and Cultural Diversity in Cyberspace, held in Yakutsk in 2008, was our first step in raising public awareness of the problems of multilingualism preservation and its development in cyberspace. It strengthened and developed professional relations and gave birth to continuous friendly contacts.

Three years after the First conference heads and leading experts of intergovernmental and international organizations; governmental authorities; institutions of culture, education, science, information and communication; representatives of business entities; civil society; media gathered again in Yakutsk to discuss political, cultural, educational, ideological, philosophical, social, ethical, technological and other aspects of the activities aimed at supporting and preserving languages and cultures and promoting them in cyberspace.

While the First Conference was organized by the Russian team only, the Second Conference was prepared with the active participation of the MAAYA and Latin Union. Thematic coverage of the Second Conference's professional programme was even broader than that of the First Conference and highlighted three decisive factors for the development of languages in cyberspace, namely: instruments for language promotion in cyberspace; institutions that are actively involved in the promotion of these instruments; and creation of favorable environment.

The Second International Conference on Linguistic and Cultural Diversity in Cyberspace was significant for the entire world and for Russia in particular because our country is one of the most multiethnic, multilingual, multicultural and multiconfessional countries of the world.

The conference final document – Yakutsk Call for Action: a Roadmap towards the World Summit on Multilingualism (2017) – was unanimously adopted at the closing session.

The conference was a real success thanks to the active support by the Commission of the Russian Federation for UNESCO and its leaders – Sergei Lavrov, Chairman of the Commission and Minister of Foreign Affairs of the Russian Federation, and Grigory Orjonikidze, Executive Secretary of the Commission; by the Ministry of Culture and especially Andrei Busygin, Deputy Minister, as well as Tatyana Manilova, Head of the Division of Libraries and Archives, and by the Government of the Republic of Sakha (Yakutia).

We are extremely grateful for the invaluable contribution to the preparation of this conference to a new partner of the Russian Committee of the UNESCO Information for All Programme, the North-Eastern Federal University (NEFU), which is being managed nowadays by our old friend Evgenia Mikhailova, former Vice-President of the Republic of Sakha (Yakutia). If not for her good will and patronage we would not have been able to hold our First Conference in 2008 either in Yakutsk or in any other city of Russia. Consequently, the Second conference could not have been convened either.

Special thanks go to Adama Samassekou and Daniel Prado, whose achievements, commitment to the honorable cause of language preservation, and global vision of the problem of multilingualism have been inspiring us over the recent years.

And, last but not least, I thank our friends from the NEFU – Vice-Rector Nadezhda Zaikova; a young Head of the recently established Centre to Advance Multilingualism in Cyberspace Liudmila Zhirkova, for whom participating in the Conference preparation turned out to be a real trial by fire; a new Director of the University Library Tatiana Maximova and all Yakut colleagues who contributed to the organization of this conference.

We believe that this collection of conference materials will be valuable for all those who face the necessity to tackle the essential problems of preserving linguistic and cultural diversity, and developing it in cyberspace at the contemporary level.

Evgeny Kuzmin

*Chair, Intergovernmental Council and Russian Committee,
UNESCO Information for All Programme,
President, International Library Cooperation Centre,
Co-Chair, Conference Organizing Committee*

MESSAGES

Message from Irina Bokova, UNESCO Director-General, on the occasion of the 2nd International Conference “Linguistic and Cultural Diversity in Cyberspace”

All languages are linked through their origins and borrowing, but each is a unique source of meaning for understanding and expressing reality. As wellsprings of knowledge, languages are essential for the transmission of knowledge and information, they are of vital importance to manage the cultural diversity of our world and achieve the internationally agreed development goals. This is even more important this year, in 2011, as we celebrate the 10th anniversary of UNESCO’s universal declaration on cultural diversity.

The digital revolution is providing us with new frontiers for innovation, creativity and development. Increased access to knowledge and information provides new possibilities for individuals and societies. Having the necessary literacies and means to participate in these digital spaces is key to improve the quality of our lives. Such possibilities must be effectively shared by all, in all languages.

The globalization process is very much facilitated by new technologies and the Internet. The success of the online edition of UNESCO’s *Atlas of the World’s Languages in Danger* shows the power of the Internet to track the state of languages and multilingualism, and to raise awareness with a global audience.

Globalization is also coupled with a tendency towards standardization, which jeopardizes the presence of many languages in cyberspace and weakens cultural diversity. In a few generations, more than half of the approximately 7,000 languages spoken today in the world could disappear.

UNESCO is committed to promoting multilingualism on the Internet. A plural linguistic cyberspace allows the wealth of diversity to put in common. By elaborating and implementing policies that address constraints to linguistic diversity, including in cyberspace, UNESCO contributes to fostering linguistic and cultural diversity and improve the conditions for promoting sustainable development and peace. These goals guide UNESCO in its work with the Internet Corporation for Assigned Names and Numbers. Strong initiatives aimed at reinforcing linguistic and cultural diversity on-line as well as off-line are being undertaken. UNESCO’s Member States adopted in 2003, the “Recommendation Concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace”, and the “Charter

on the Preservation of Digital Heritage”. Both instruments provide guidance on steps that are to be taken to advance multilingualism in cyberspace.

This second edition of the international conference “Linguistic and Cultural Diversity in Cyberspace”, is an important step in consolidating and advancing the progress to date. Your endeavours are vital to find appropriate solutions to pressing concerns such as: elaborating public policies on languages in cyberspace; identifying techniques to ensure the presence of absent and under-represented languages; supporting the implementation of UNESCO’s normative instruments. I greatly welcome these actions and encourage you in your efforts to facilitate cooperation and exchange at the national, regional and international levels. Let us all harness the power of progress to protect diverse visions of the world and to promote all sources of knowledge and forms of expression.

I wish you fruitful discussions and look forward with great interest to the outcomes of your conference.

Irina Bokova

**Message from Grigory Ordzhonikidze,
Executive Secretary of the Commission
of the Russian Federation for UNESCO,
to the participants of the 2nd International Conference
“Linguistic and Cultural Diversity in Cyberspace”**

I cordially greet the organizers and participants of the 2nd International Conference “Linguistic and Cultural Diversity in Cyberspace”.

The topic of the first conference held three years ago has proved its relevance: the rapid processes of building a global society, on the one hand, increase the unification of cultures, and on the other hand, offer opportunities for preserving and developing cultural diversity, in particular in such a universal, cross-border field as cyberspace.

I believe that a way to eliminate this dialectic contradiction can be found in preventing the negative effects of globalization, in the collaborative search for ways to achieve sustainable development for everybody, in harmonizing relations among nations and civilizations, encouraging cultural diversity and identity of the peoples of our planet. UNESCO as the world’s most influential organization can and should make a serious contribution in these activities. It is no coincidence that this organization has developed the concept of “World culture” built upon the idea of forming a new type of international relations, based on tolerance, non-violence, respect for human rights, mutual respect of cultures, traditions and religions.

The problem of linguistic and cultural diversity preservation is topical for all countries, particularly for such multinational ones as Russia populated by over 180 peoples speaking more than 100 languages. No wonder that in 2008 here, on the Yakut land, where the climate is harsh, but people are kind and responsive to current problems of mankind, an international conference which became a contribution to the International Year of Languages, adopted a document of great importance known as “Lena resolution.” What is extremely valuable about this appeal to the world is that it offers ways to implement some of the recommendations of the World Summit on the Information Society, and launches initiatives to provide universal access to information and knowledge, in particular the idea of holding a world summit on linguistic diversity in cyberspace under the aegis of UNESCO and MAAYA Network.

I am confident that this conference will further support these initiatives and give participants an excellent opportunity to exchange ideas and experiences, while facilitating the preservation and development of cultural and linguistic diversity.

I would like to thank the Government of the Republic of Sakha (Yakutia) and the Russian Committee of the UNESCO Information for All Programme for organizing such a representative and useful forum.

I wish you successful and fruitful work, vivid impressions of staying on the unique land of Yakutia and all the best.

Grigory Ordzhonikidze

**Message from Alexander Avdeyev,
Minister of Culture of the Russian Federation,
to the participants of the 2nd International Conference
“Linguistic and Cultural Diversity in Cyberspace”**

On behalf of Russia’s Culture Ministry, I extend a warm welcome to all delegates attending our second “Linguistic and Cultural Diversity in Cyberspace” conference.

In UNESCO’s 2001 Universal Declaration, cultural diversity is viewed as part of the humankind’s heritage. Linguistic diversity is what forms the basis of cultural diversity. Languages are crucial to progressing toward sustainable development, given their key role in providing quality education, spreading knowledge, and stimulating social integration and economic development. This is why the theme of this conference is of so much relevance to the world community as a whole and particularly to Russia, which is one of the most multi-ethnic, multi-faith and multilingual countries, with as many as 180 languages spoken here and nearly 40 indigenous languages enjoying official language status. The Russian Constitution proclaims that the languages spoken by the country’s constituent communities are all part of the national cultural heritage.

This forum is being held within the framework of Russia’s chairmanship in the Intergovernmental Council of the UNESCO Information for All Programme.

I would like to thank all of the event’s organizers, notably the government of the Republic of Sakha (Yakutia), the North-Eastern Federal University, the Russian Committee of the UNESCO Intergovernmental Information for All Programme (IFAP), the Interregional Library Cooperation Centre, the National Commission of the Russian Federation for UNESCO, the MAAYA World Network for Linguistic Diversity, the Latin Union, and, of course, UNESCO itself.

We are happy about Russian IFAP Committee working in partnership with the Ministry of Culture and the Federal Agency for Press and Mass Communications to enhance the programme and further raise its efficiency.

We are also glad that our joint efforts to promote linguistic and cultural diversity – both in the real world and on the Web – keep gaining momentum. The first conference “Linguistic and Cultural Diversity in Cyberspace”, held in Yakutsk in 2008, was attended by politicians and experts representing 15 countries from all around the globe. The lineup of this year’s gathering

features as many as 33 nations. The agenda is rich and interesting, with a whole array of topical issues to be discussed. We hope the forthcoming discussions will effect meaningful change, enabling us to find new efficient ways of preserving endangered languages and promoting language diversity in cyberspace with the help of innovative information technology.

I wish you all an enjoyable and productive forum.

Alexander Avdeyev

**Message from Andrei Fursenko,
Minister of Education and Science of the Russian Federation,
to the participants of the 2nd International Conference
“Linguistic and Cultural Diversity in Cyberspace”**

Esteemed colleagues, dear friends!

I am pleased to greet the participants of the 2nd International conference “Linguistic and Cultural Diversity in Cyberspace”!

Native language is a great treasure for each nation. Learning it allows entering a unique world of moral and spiritual values, embodied in literature, arts and science. The importance of multilingualism development and promotion cannot be exaggerated, as languages are storehouses of the entire knowledge of mankind. Languages are a means of socialization, of representing and sharing cultural traditions of their speakers.

In this context preserving languages and supporting minority languages and cultures is of vital importance.

I feel sure that the conference will facilitate elaborating and implementing strategies of knowledge development in a world, where information and communication technologies are actively used.

I wish you all successful, exciting and fruitful work!

Andrei Fursenko

**Message from Egor Borisov,
President of the Republic of Sakha,
to the participants of the 2nd International Conference
“Linguistic and Cultural Diversity in Cyberspace”**

Dear conference participants and guests!

On behalf of the Government of the Republic of Sakha (Yakutia), I congratulate you on the opening of the 2nd International Conference “Linguistic and Cultural Diversity in Cyberspace!”

To date, the Republic of Sakha (Yakutia), one of the most outstanding and original Russian entities, is home to over 120 nationalities. Preservation and development of the mother tongue and other national languages, cultures, customs and traditions are top priorities of the language policy implemented in the region. Great efforts are made to implement a consistent language policy based on the consolidation and development of the Yakut and Russian languages, as well as preservation of the Even, Evenki and Yukaghir languages – native languages of the republic’s indigenous peoples.

Preservation of linguistic and cultural diversity is impossible without the use of information and communication technologies. It is essential to preserve information in different languages, in particular in cyberspace. Expanding free access to cyberspace in the native languages of all peoples of the Republic of Sakha (Yakutia) is a prerequisite for it.

I am certain that the conference will represent significant cultural initiatives in the field of multilingualism preservation in cyberspace; lay the foundation and open prospects for further activities in this sphere in Russia; provide an opportunity to discuss various aspects of the preservation and development of linguistic diversity in the information space.

I wish you creative achievements, success in all undertakings, inexhaustible energy and new ideas!

Egor Borisov

**Message from Abulfas Karayev,
Minister of Culture and Tourism of the Azerbaijan Republic,
to the participants of the 2nd International Conference
“Linguistic and Cultural Diversity in Cyberspace”**

Esteemed conference participants, dear friends,

On behalf of Azerbaijan’s Ministry of Culture and Tourism and my fellow countrymen working in culture and the arts, I would like to greet all the organizers, participants and guests of the 2nd conference “Linguistic and Cultural Diversity in Cyberspace”.

Computer information systems have become part and parcel of modern life while cultural and linguistic diversity, which makes our life so much richer, is a top priority on today’s agenda. Linguistic and cultural diversity on the Web gives us a better understanding of the problems and interests of the present-day world.

Our mission is to preserve countries’ national identity along with their cultural and language diversity, crucial to ensuring the sustainable development of society and intercommunal harmony.

We seek to attain modern cultural standards based on universal human values while also remaining loyal to our own distinctive traditions. We should evolve as a community by enhancing our spiritual capacity and social cohesion.

I wish you all productive work and fruitful discussions!

Abulfas Karayev

Welcoming address by Evgenia Mikhailova, Rector of the North-Eastern Federal University

Ladies and Gentlemen, dear colleagues,

To the Yakut people, 3 is a magic number, symbolizing the past, the present and the future. Symbolically, our second conference “Linguistic and Cultural Diversity in Cyberspace” is taking place here, in Yakutia, precisely three years after the first such forum.

This conference provides a unique opportunity for expanding intercultural dialogue, establishing closer ties between communities across the globe and promoting solidarity, friendship, mutual understanding and support.

As a representative of the Sakha people, I can say that we take pride in our cultural heritage, just like other peoples and nations feel proud of theirs.

According to a Yakut saying, “the Sakha are a hospitable people; friends are their true wealth” (“саха ыалдьытымсах, кини доҕотторунан баай”). And indeed, the more friends one makes, the richer and stronger he or she becomes.

We are thankful to all of the forum’s delegates for having brought us the joy of new encounters and new friendships.

We appreciate your sincere interest and admire your commitment to working for the benefit of the world community.

Evgenia Mikhailova

PLENARY MEETINGS

Adama SAMASSÉKOU,

*President, MAAYA Network;
President of the International Council
for Philosophy and Humanistic Studies
(Bamako, Mali)*

Developing Institutional and Legal Instruments to Support Multilingualism Worldwide

The international context is currently characterized by the globalization of markets, which leads to losing the connection with the Other, the non acceptance of the Other, with more and more exclusion and violence in the relation to the Other.

The world crisis today which, far from being financial or economic, is rather a societal one, a values crisis, leads to the fall of the economic model and system linked to a profit making culture, the culture of consumption and gain.

Our world basically needs more humanness. Our world needs to develop another culture, the one of human being that is able to guarantee more humanness in people's relations and less mercantilism! There is a saying in Mali: *Mogotigiya ka fisa nin fentigiya ye!* ("Human relations are more valuable than money"). We find the same in the Russian tradition: *Не имей сто рублей, а имей сто друзей!* ("Better to have 100 friends than 100 roubles!")

That's why there is an urgent need to preserve and promote world cultures that put human being in the centre of their concerns, to promote those societies characterised by a vision of the world based on the permanent search of harmony between human beings and nature, and friendly relations as the cornerstone of our human existence. That's what we mean when we refer to linguistic and cultural diversity: we confront the rampant process of uniformizing cultures, the development of global common thinking and utopian monolingualism. As a matter of fact, linguistic diversity is as necessary for humankind as biodiversity is for nature: it is the breath that guarantees its vitality. Nature without biodiversity is nothing more than a *nature morte* and to suppress the diversity of the society is the same as to create soulless robots. 'The beauty of a carpet lies in its different colours', said Amadou Hampâté Bâ, a philosopher, writer and wise man.

Dear friends, this advocacy for diversity is well known; even more, being militants for social transformation you know quite well that multilingualism is to culture what multilateralism is to politics: the frame that guarantees an equitable relation to others and the equilibrium of powers. Our purpose here is not to convince those present, but rather to turn from advocacy to action, that

is to walk the talk. We should offer the world the means and ways of this action. This is the *raison d'être* of our MAAYA network!

Our concern today is to be all committed to the development of institutional and legal instruments, that would enable preserving and promoting multilingualism worldwide.

Although multilingualism is in the world a norm rather than an exception, very few countries are now using institutional instruments to support it. Certainly, it is partially due to the historical process of a monolingual Nation State development, but, what is more important, it is also a consequence of certain nations being dominated by others, in particular in different regions of the world, at different levels of the evolution of human societies.

The experience of multilingual countries and regions (except for those where linguistic conflicts have separated peoples) shows that numerous languages can co-exist together, in harmony, in the same space, national or regional.

How can we learn from these positive experiences today and develop instruments necessary to strengthen multilingual practice?

From the very beginning we should realize that this is first and foremost a political question, linked to the philosophical conviction in the necessity of supporting diversity. It presupposes being mentally open to the Other and respect for fundamental human rights.

Our approach should be implemented at three levels.

National Level

The first institution to support multilingualism is the school (in the broadest sense, i.e. the educational system), which should be refounded in all countries in order to develop a mother tongue-based multilingual education (mother tongue understood as the most familiar language). The development of multilingualism is the best guarantee for the development of a society that is open to others and that respects linguistic and cultural diversity!

Besides, it is vital to create, where they do not exist, institutions of research and promotion of the country's various languages. And those countries where such institutions are already established, should develop and reorient them towards the enhancement of multilingualism.

An appropriate legal framework, with legislative and regulatory instruments should be elaborated, in the context of an explicit language policy aimed at promoting multilingualism and supporting every language of the country, both in public life and in the private sphere, according to the strategic approach we used to call «convivial functional multilingualism».

All the institutions promoting book, reading and literary creativity (libraries, publishers, and associations of writers) should contribute to

supporting multilingualism, particularly through proper usage of ICTs, including multimedia.

Regional Level

At this level there are two complementary trends.

On the one hand, at the level of each region of the world (according to the definition of the United Nations), it is necessary to encourage the creation of a structure similar to what African Academy of Languages (ACALAN) is for Africa, that is to say an intergovernmental scientific institution responsible for the enhancement and promotion of the languages of the continent, in the context of a continental language policy encouraging multilingualism and convivial partnership between all the existing languages.

This institutional innovation could be reinforced with a more or less restricting legal instrument, inspired by the «European Charter for Regional or Minority Languages», yet with a permanent concern for equity between all the languages.

On the other hand, at the level of each organization engaged in the promotion of one or several languages spoken in several regions in the world (OIF, Commonwealth, Latin Union, Arab League, CPLP, etc.), it is necessary to encourage them to leave the logic of defending a specific language or a group of languages in order to resolutely commit themselves to the promotion of languages co-existence and multilingualism.

International Level

The MAAYA World Network for Linguistic Diversity was established to constitute a kind of international platform for a multistakeholder partnership in order to safeguard linguistic diversity and promote multilingualism. The existence of such organizations should make it possible to federate the initiatives taken at the international level, both from the point of implementing all the various declarations, resolutions and conventions dealing with multilingualism and cultural and linguistic diversity (UNESCO, WSIS, UN, various regional and interregional organizations, Lena Resolution, Bamako International Forum on Multilingualism, etc.), and from the point of establishing a kind of monitor structure, a world observatory for multilingualism.

The MAAYA Network and its founding organizations initiated the process of the preparation for the World Summit on Multilingualism and this second Yakutsk conference is seen as a preparatory step for it. The Summit should permit to make concrete propositions at the highest level of the international community, in order to promote multilingualism in the world both from institutional and legal points of view, thus contributing to the global realization of the Millennium Development Goals (MDG) and to the event of a World of Peace based on the dialogue of cultures and civilizations.

Daniel PRADO

*Executive Secretary, MAAYA Network;
Director, Terminology and Language Industries, Latin Union
(Paris, France)*

It Is Time to Place Multilingualism and Linguistic Diversity at the Heart of the International Debate

Our world is not the same as it was after the Second World War, when two different ways of seeing the world faced up to each other, when international relations were based on military, political and economic criteria and no thought at all was given to valuing culture as the basis of society. It is not the same either as the world we knew from the late 1990s, with a single dominant culture becoming more hegemonic.

By 2011 the world has become multipolar, supranational alliances based on respect for culture have become stronger, alliances between regions of different nations sharing a language or a culture are being born every day. And although political and economic issues are still present in international relations, cultural aspects are taking a larger place. Religion, ethnicity, customs and language now play a part in the formulation of international policies, and even though they can be, unfortunately, a factor of discord, culture is increasingly perceived as a major vector in sustainable development and fair growth facilitating harmony among peoples and respect for their dignity.

The adoption of the Convention on the Protection and Promotion of the Diversity of Cultural Expressions was the very symbol of culture's changing place in international relations. It emphasized that culture should be seen not as a mere economic value but above all as an essential condition of human beings' existence and the best motor for development that respects the future of the planet.

The Convention is an opening that activists working for multilingualism and respect for linguistic diversity have seized upon as it is high time to consider language as closely associated to free expression and self-development, equal opportunities and promotion of understanding among peoples on fair and balanced foundations.

Language does have an implicit place in the Convention but implicit is not obvious to everyone. The Millennium Development Goals, having omitted to include culture as a goal in its own right, also failed to refer to the language of individuals. The Convention for the Safeguarding of the Intangible Heritage does make a careful reference and, then, both the Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace, and the World Summit on the Information Society gave body

to the idea of respect for linguistic diversity and multilingualism, but these instruments must be really followed by practical achievements. And we know that we are nowhere near being able to give all individuals the opportunity to develop freely in their own language.

Since the beginning of the millennium, and in a surprising turnaround after decades of linguistic hegemony, language has become more present in both political and commercial issues. The main languages of communication that are official languages alongside English in international forums (French, Spanish, Russian, Chinese and Arabic at the United Nations; 22 other languages in the European Union; Portuguese, French, Arabic and Swahili in the African Union; French, Spanish and Portuguese in the organizations on the American continent, etc.) demand respect for balance in the way they are treated.

Other languages which are not official (Portuguese at the United Nations, regional languages in the European Union, Guarani in South America, etc.) demand the right to be official too, and many initiatives (Council of Europe, Linguapax, UNESCO, etc.) are giving an increasingly significant role to languages that do not have an official national or regional status.

Given the absence of international conventions concerning language, however, given the lack of reliable indicators on its impact on global development, given the probable death of almost half the world's languages, given the (still too high) level of injustice owing to the fact of speaking an unrecognized or a marginalized language, a large number of measures are called for in the field of indicators, policies and promotion, without forgetting the legal instruments.

With such a wide-ranging debate, I shall emphasize only what the World Network for Linguistic Diversity MAAYA can propose in terms of promoting and enhancing (sometimes re-enhancing) languages, in particular in cyberspace.

Although it is true to say that scarcely one individual in three today has access to the Internet, we can see constant movement towards the universal spread of the phenomenon. We also know that cyberspace and related technologies are tending to gradually replace our old ways of communicating, expressing ourselves, transmitting information and sharing knowledge. If languages cannot ensure the circulation of this information, knowledge or dialogue, they are in danger of losing value in the eyes of their speakers, since migration and urbanization and universal access generate confrontations between languages in which only those which are held in high esteem by their speakers can survive.

Accordingly, MAAYA endeavours to inform those who can take decisions about policies to vitalize or revitalize languages of the full implications of cyberspace.

MAAYA's meetings like the World Congress on Specialized Translation (Havana, 2008), the Bamako International Forum on Multilingualism

(Bamako, 2009), and the first and second international seminars on multilingualism in cyberspace (Barcelona, 2009 and Brasilia, 2011) allow to compare opinions of the main experts, disseminate best practices and make decisions to improve the role of languages in knowledge.

Promotion and research initiatives came out of these meetings, such as *Net.Lang*, Dilinet and Voices and Texts.

Net.Lang, a book that is due for publication in October 2011 in English and French (and other languages later on), brings together the most eminent people active in promoting linguistic diversity in cyberspace so as to provide all possible ways forward in giving languages the best instruments for their inclusion in cyberspace. Some 30 articles and texts by politicians and researchers should help raise awareness of all those who can have a decision-making role in the dynamic of languages. The hard copy of the book, and various digital formats free of charge and with universal access on the Internet (ePub, PDF, etc.), will be distributed worldwide.

Dilinet is a research project that will facilitate the elaboration of indicators for the presence of (all) languages in cyberspace. The indicators, as we know, are the instruments that enable us to direct policies. And we feel a crying lack of them, as regards languages. The main laboratories and entities dedicated to this research, together with serious institutional partners and private companies should enable us, through Dilinet, to monitor closely the evolution of languages in cyberspace.

Voices and Texts, an upcoming project, will aim at applying voice recognition, machine translation and voice synthesis technologies to certain languages, giving them access to a series of software packages and many applications using these languages.

However, promotion and research are not enough. MAAYA is present in international forums trying to regulate and encourage the flourishing of all languages in the *shared knowledge society*, particularly those that came out of the **World Summit on the Information Society**, as sub-moderator for Action Line C8 (Cultural diversity and identity, linguistic diversity and local content) and lead moderator for the Dynamic Coalition for Linguistic Diversity in the context of the Internet Governance Forum.

MAAYA will be present in these major undertakings ahead of us so as to ensure that every of the world's languages occupies its rightful place and guarantee the right of its speakers to use it fully.

We are also glad to participate in the 2nd International Conference on Linguistic and Cultural Diversity in Cyberspace in Yakutsk. We hope that any resolutions adopted here will help us make a reality of the idea of a world that is fairer, better balanced and more harmonious.

Evgenia MIKHAILOVA

*Rector, North-Eastern Federal University
(Yakutsk, Russian Federation)*

Our Common Goal is to Preserve Not Only Our Language, but also Our Culture, Environment and People

The First International Conference on Linguistic and Cultural Diversity held three years ago in Yakutsk was extremely informative and enlightening for many of us. Earlier, I have already tried to take real actions within the scope of my powers to create conditions for studying native languages at preschool educational institutions and at schools in order to preserve and promote cultural and linguistic diversity in the republic, expand the social base of native languages and the scope of their application and to publish works of writers and poets representing all the peoples living in Yakutia. After the conference in 2008, the preservation and promotion of cultural and linguistic diversity of our peoples became my highest priority.

Our republic is multinational. The Yakuts (Sakha) are the most northern of stock-raising peoples, they have a particular economic structure, material and spiritual culture, as well as unique methods of education of their children and youth. People of Sakha familiarize their children with monuments of national culture from an early age. Conditions for the formation of linguistic competence and development of individuals, capable of using the system of global communication and familiar with information technology, have been created in educational institutions of the republic. In Yakutia, great importance is attached to the linguistic background, because it is viewed as a basis for cross-cultural communication that shapes the consciousness of a person, determines their world views, promotes readiness for dialogue, respect for their own culture and traditions, tolerance of other languages and cultures. In 1996-1997, which were declared by M. Nikolaev, President of the Republic of Sakha (Yakutia), the years of youth and education, content priorities of the courses of Russian, native and foreign languages were revised. The main focus was on the communicative aspect of language education, on free speech activity and communicational culture. Familiarizing students with the nation's cultural heritage and culture of modern society and teaching them to use their native language freely in all public areas of its application were announced priority tasks of language education.

In the academic year 1996/1997, it was made possible in schools to study and to get education in six languages, namely in Russian, Yakut, Evenk, Even, Yukaghir and Chukchi languages. Following results were achieved during those years: 1,099 Evens (47.7% of all Even children), 907 Evenks (27.1%),

96 (58.5%) Yukaghirs and 95 Chukchis (78,5%) were studying their native language. These figures were due to the lack of trained teachers, as well as parents' reluctance to educate their children in their native language. At our first conference in 2008, however, Feodosia Gabysheva, Minister of Education of the Republic of Sakha (Yakutia), came up with quite different figures: the Even language was studied as native language by 2275 people; the Evenk language by 3283, the Yukagir language by 168; the Chukchi language by 145; the Dolgan language by 252. The number of people exercising the right to study their native language is growing each year.

25 years ago, when I was heading the Local Department of Public Education of the Yaroslavsky District of Yakutsk and the Urban Department of Public Education of Yakutsk, we have introduced a rule to create Yakut-speaking groups alongside Russian-speaking ones in every newly built kindergarten. We started create conditions for the children of indigenous peoples living in urban areas to study their native language and to get education in it.

As Minister of Education, I ensured in 1997 that Even, Evenk, Yukaghir, Chukchi and Dolgan children were no longer obliged to study Yakut language (according to the then-existing curriculum, these children had to be polyglots, as they had to study two official languages, Russian and Yakut, their native language and a foreign one, e.g. English). Being in charge of education, I saw my task in helping to create the best conditions for the development of each language and gradual formation of the need to use both official languages in full compliance with rights of a citizen to choose the language of study and understanding that nothing should be mandatory when it comes to human rights.

We clearly understood that the state is obliged to create conditions for efficient pre-primary education and in 1997 the Ministry of Education of the Republic of Sakha (Yakutia) obliged pre-school educational institutions to admit children from the age of 5. Within a limited time, all preschool teachers were trained to teach reading (from the age of 5). Already in 1999, 93% of all five-year old children in the republic (100% in some areas) were attending pre-school educational institutions. And in 2000, children who started their first grade were already able to read, and, most importantly, had the ability to learn in a group and live in the society.

In 2000, through the personal efforts by Vyacheslav Shtyrov, President of OJSC ALROSA, computer equipment worth \$2 million was donated by ALROSA to schools of the republic. Sakhatelecom company, then led by Nikolai Nikolaev, granted schools with a 3 year long "guest" Internet access. Such support helped to make a breakthrough in the education system of Yakutia.

One of the major indications of the progress achieved by all school teachers and officials of the Ministry of Education was the fact that in 2003 Yakutia

ranked third (after Moscow and St. Petersburg) in the federal competition "The best region of Russia in terms of ICT" within the programme of "Electronic Russia" for the use of ICT in education. The republic has made such a breakthrough in providing educational institutions with computers, because it recognized the priority of computerization and distance learning, given the remoteness of its educational institutions not only from Russian and foreign but even from Yakutsk research and methodological centers, and their significant information isolation by force of circumstances.

Today, thanks to the implementation of National Priority Projects, initiated in 2006 by Russian President Vladimir Putin, all educational institutions of the republic have access to broadband Internet. All secondary school graduates speak Russian, native, foreign languages and are active users of ICT. The generations born after the advent of computers and information technology, are familiar with modern technologies.

Today, the teacher is not the only source of information and knowledge. Instant communication technologies provide access to all relevant data. Nowadays through Internet people can always get the most effective courses by the best teachers. Modern school and college students no longer believe the myths that used to exist before the Internet:

1. School is the best place for learning.
2. Intelligence is unchangeable.
3. The level of education is the result of teaching.
4. We all learn in the same way.

Even now, people sometimes confuse education with schooling, their health with disease treatment and hospitals, law with lawyers. Earlier, some people profited from this situation, as it helped to make people less self-reliant and less eager to make decisions about their life. But the society is undergoing huge changes, the transition from the old system to the new one is under way. Parents wonder why children are being prepared for life in the past, and not for the life in the future. Recent polls in Yakutia confirm a positive attitude of people of the republic towards changes in its educational system, particularly in the sphere of higher professional education. According to the federal educational policy of the country, establishment and development of federal universities is regarded as an instrument of social and economic advancement of the regions within federal districts and development of an innovative economy in Russia in general.

For the period up to 2030, the Government of the Russian Federation expects Yakutia and its surrounding areas to meet a number of challenging social and economic objectives requiring a new quality of economic growth and

exploration of new ways of development based on modern approaches. Major investment projects in the Far East provide for large-scale transformations aimed at accelerating economic development of the north-eastern part of Russia, increasing concentration of available resources and developing more effective management. This is precisely why Egor Borisov, President of the Republic, and the Yakut government place so much emphasis on creation of a new talent pool able to generate ideas and bring them into life. In this regard, the North-Eastern Federal University named after Maxim Ammosov (NEFU) becomes a major resource for the development of the republic and the north-east of the country.

The NEFU has quite recently celebrated the first year of its new status in a new organizational form, the form of an autonomous institution. Systemic changes has started in all the spheres of the university life. Actions taken in accordance with the development priorities approved by the Government of the Russian Federation define and shape the establishment and development of the leading scientific, educational, methodological and cultural centre in the north-east of Russia.

The implementation of the NEFU Development Programme provides for the establishment of a strong research and innovative entrepreneurial university that will contribute to the development of regional technologically innovative economy. The NEFU doesn't seek to become the largest regional university, its goal is to be the best, high-demand and useful for the region, to support new technologies' development and promotion, combine efforts to improve cross-cultural ties and corporate services in partnership with the public, businesses, academic institutions and state authorities.

The university drives the expansion of the domestic market. It creates and ensures the development of human capital, which will contribute to the viability and competitiveness of the regional economy.

In comparison with the programmes of other federal universities, that of the NEFU has a stronger humanitarian component. The innovative project, titled "Preservation and Development of the Languages and Cultures of the Peoples of the North-East of Russia", was launched immediately after the Decree had been signed by Russian President Dmitry Medvedev and the programme had been approved.

The NEFU has established its branch in the Chukotka Autonomous District (in the city of Anadyr). The following institutions were also created:

- the Institute of Alexei Kulakovskiy, founder of Yakut literature;
- the Olonkho Research Institute (Olohkho was declared by the UNESCO to be a Masterpiece of the Oral and Intangible Heritage of Humanity);

-
- the Institute of Languages and Culture of the Peoples of the North-East of the Russian Federation;
 - the Institute of Foreign Philology and Regional Studies.

The NEFU attaches great importance to language education and study of language processes. There are three philological divisions in the University.

As far back as 1991, Yakutsk State University established a department of northern philology within the Philological Faculty. It was initially lead by the unforgettable Vasily Robbek, Academician of the Academy of Sciences of the Republic of Sakha (Yakutia). Over a relatively short period the department has significantly strengthened the intellectual elite of Even, Evenki and Yukaghir peoples. Before the establishment of the department there were only 8 Even schools, 2 Evenki schools and 1 Yukaghir school. Now the graduates of the department work in 20 regions of the republic as teachers of Even, Evenki, Yukaghir languages and literature, school psychologists, kindergarten teacher, some of them are employed by the National Broadcasting Company Sakha, the Institute of Humanitarian Research and Indigenous Peoples of the North of the Siberian Branch of Russian Academy of Sciences. In total, 269 highly qualified linguists – representatives of the Even, Evenki and Yukaghir peoples have graduated from the department.

The Institute of languages and cultures of the NEFU has initiated an experimental study of the effectiveness of educating children in their native language. Experts collect materials related to the Yakut language and the languages of the peoples of the North-East of the Russian Federation and create an acoustic archive. Special attention is paid to the teaching of languages, culture and literature, and to the translation from Russian and foreign languages into indigenous languages of the Russian North-East.

In July 2010, the Centre to Advance Multilingualism in Cyberspace was established in the NEFU, which contributed greatly to organizing the Second Conference on Linguistic and Cultural Diversity in Cyberspace. Problems and prospects of the Centre were discussed at the seminar held at the NEFU with participation of Evgeny Kuzmin, Chairman of the Intergovernmental Council and Russian Committee of the UNESCO Information for All Programme.

In November 2010, Yakutsk hosted an offsite meeting of the Committee of the Federation Council of the Russian Federation on the North and Indigenous Peoples "On the use of modern information technologies for preservation and development of the languages, culture and spirituality of the peoples of the North (the case of the Republic of Sakha (Yakutia))." The NEFU played an active part in the event and organized a special round-table discussion on November 17, 2010. All suggestions made by the NEFU were supported and

included in the recommendations of the offsite meeting, and eventually an official policy document on the subject was adopted.

Together with the Siberian Federal University, we launched a «Foresight» project («Foresight» – Study of the Republic of Sakha (Yakutia)). The Yakut people are the main participants of this study aimed at defining a possible future of the peoples of the republic on the basis of expert assessments of social and economic spheres, utilizing the method of long-term regional development forecasting. Its main peculiarity is the focus on the development of practical measures for rapid achievement of the selected strategic objectives that guide the medium and long term development of the republic.

A programme for the preservation and development of the Yukaghir language and culture in digital media and cyberspace for 2011-2014 was launched. The results achieved during the first six months of 2011 include five educational DVD-based courses on the Yukaghir language, a basis for the web-portal of the indigenous small peoples of the North, www.arctic-megapedia.ru, where materials on the language and culture of Yukaghirs are already available.

Measures are being taken to further develop the ways for operational communication, interaction and bringing people together.

Modern information technology shortens the distance, helps to overcome the language barrier and does indeed change the world. Within the NEFU Development Programme, a number of steps are being taken to advance information and communication technologies. The main goal is to build an IT system to support research and education process and achieve full automatization of university work.

A Wi-Fi network provides a wireless Internet access to students within the University buildings and will soon be extended to cover the entire campus. The biggest university building of the Natural Sciences Department offers free Wi-Fi mobile working areas which are also helpful to our guests – foreign students from Norway, Finland, Sweden, who come here in the framework of the North-North exchange programmes, and South Korean students, who participate in the NEFU Summer School.

The University works to integrate cutting-edge technologies. There is a well-known expert estimate of ten information technologies that will change the world. Cloud Computing technology comes first among them: it enables small computers to process information using the potential of big datacenters located across the world.

In this area, we explore ways to use cloud computing to build an innovative platform for a virtual electronic university – “the Yakutsk INTER-University.” A student may enter a service cloud anytime and anywhere provided that he has an Internet connection. Any lecture room or classroom can turn into a laboratory or a computer classroom.

The NEFU Development Programme includes elaborating a hardware and software system in 2011 that will be used to provide both high-performance computing and secure storage of large volumes of information resources and their prompt accessibility. It will also serve as a basis to create a NEFU repository of information and educational digital resources in the languages of small peoples of the North-East of Russia, including digitized copies of books and documents, electronic manuals and teaching aids, cinema and video films, audio recordings, etc. It will also allow free downloading of digitized copies of text and multimedia materials for those owners who would like to provide a free access to their resources but do not have sufficient technical capacities.

We plan to put in place a free and open repository of information and educational digital resources in the languages of small peoples of the North-East of Russia and the Arctic regions transmitted by their authors and owners for secure storage and non-commercial use for educational purposes; to form an integrated catalogue of Internet resources; and to organize off-campus education projects with the help of the integrated catalogue and repository to boost interest in and use of Internet resources in the languages of small peoples of the North-East of Russia among speakers themselves, particularly among the youth.

On June 9, 2011 UNESCO and the NEFU signed a treaty to establish the University-based UNESCO Chair “Adaptation of Society and Man in Arctic Regions in the Context of Climate Change and Globalization.”

The Chair is inter-departmental and aims at addressing issues of adaptation of both the society and man in Arctic regions by establishing international scientific and educational collaboration, raising awareness of the population by means of modern information and telecommunications technologies and forming an Internet information environment on the issues of Arctic regions. The Chair will promote cooperation between the outstanding scientists of the NEFU and other universities of the Republic of Sakha (Yakutia) and the Russian Federation, as well as foreign countries.

The Constitution of the Russian Federation begins with the following words: “We, the multinational people of the Russian Federation...”, and we must constantly bear in mind that Russia is a poly-ethnic State. According to the 2002 population census there are 180 ethnic communities living in Russia. Over the last decade, focus has been on the language policy. The Republic of Sakha (Yakutia) could be considered polylingual. We understand that, with all the importance of bilingual motivation, it is linguo-ecological motivation that has a crucial role to play, which implies preserving and developing languages of all peoples, expanding their application and use.

I am convinced that our conference will contribute to the further development of linguistic and cultural diversity in cyberspace.

Evgeny KUZMIN

*Chairman, Intergovernmental Council and Russian Committee,
UNESCO Information for All Programme;
President, Interregional Library Cooperation Centre;
Co-Chair, Conference Organizing Committee
(Moscow, Russian Federation)*

Social Institutions Supporting Linguistic and Cultural Diversity in Cyberspace: Roles, Functions, Responsibilities

Most nations in the contemporary world enjoy neither statehood nor sovereignty. Their languages are not state languages because a majority of countries are multiethnic and multilingual. Even in the best possible scenarios, when governments and dominant ethnic groups are rigorously protective of ethnic and linguistic minorities, most of their languages are still marginalised to varying extents. They exist and develop (or decline) in the shadow of the country's dominant language, which is used in all spheres of influence—political, economic, educational, cultural, scientific, etc.

Globalisation, various possibilities for migration in a context of high mobility, and the rapid pace of urbanisation have made many ethnic minorities undervalue their native language. Learning native language to talk with fewer and fewer people on a decreasing number of topics is regarded as a blind alley. Meanwhile, state and international languages garner a wealth of attention and research.

No language develops outside the context of its corresponding ethnos. At the same time, urbanisation and globalisation encourage smaller cultures to merge with the majority, and marginalize themselves. The knowledge and historical and cultural experience stored within these cultures gradually vanish, as well as the culture's/language's potential. Cultural and linguistic marginalisation is thus an interrelated and multifaceted process; with the death of a language, its unique carrier culture vanishes¹.

All these and other factors lead to a dramatic decrease in the number of active speakers of minority languages resulting in further marginalization (and extinction – in extreme cases) of the less equipped languages with the smallest number of speakers.

These issues are salient for nearly every country where two or more languages cohabit. What can we do to stop or at last to hinder the process of language

¹ *In the context of this article, the term “culture” is used in the broadest sense to denote the entirety of salient material, intellectual and emotional features of a given community or social group, comprising the arts and literature, as well as lifestyle, the status of human rights, value systems, education, customs, traditions and philosophy.*

marginalisation, and to enhance the fitness of endangered languages? Who can do it, and whose duty is it?

Let us examine how Russia, one of the most multiethnic, multilingual and multi-religious countries of the world; tackles these issues, and to what extent it solves them.

180 world languages are spoken in Russia belonging to the Indo-European, Altaic, and Ural language families, the Caucasian and Paleo-Siberian language groups. Those are not languages of new immigrants; various Russian population groups have been speaking them for centuries. Over a hundred of these languages belong to **indigenous ethnic entities historically formed within the present-day Russian borders or living there for centuries.**

The Constitution of the Russian Federation declares all languages of Russia to be common cultural assets. Almost all languages use graphic systems, even if some have acquired them somewhat recently.

There are four most widely used languages except Russian with between 1.5 and 5.5 million speakers: Tatar, Chuvash, Bashkir, and Chechen. A further nine languages have between four hundred thousand and a million speakers – Yakut language belongs to this group. A further fifteen are spoken by between fifty thousand to four hundred thousand people.

Intensive cultural dialog, mutual exchange and enrichment that took place on the territory of the Russian Empire and Soviet Union feature modern Russia as well. Respect for other peoples' cultures has been intrinsic to Russian policies allowing to preserve the richest cultural and linguistic diversity we are justifiably proud of.

Unlike many other multilingual countries, Russia offers education (primary and even higher for certain fields of humanities education), television and radio broadcasting, internet resources, books and newspapers in nearly all of its languages. These activities find financial support by the state. Russia is unique in another respect as well: **close to forty of its indigenous languages enjoy official status.**

All languages except Russian are minority languages, and all are marginalized to varying extents. Future perspectives are limited for representatives of indigenous peoples speaking their native language only. Proficiency in Russian is required for building a career or realizing one's potential, especially in the intellectual sphere.

Problems and issues for concern are still numerous. Out of a hundred indigenous languages of Russia nearly thirty are minority languages of the

peoples of the Far North, Siberia and the Far East less than fifty thousand speakers strong. Some of the languages have less than 100 speakers. Despite official efforts at every level of Russian bureaucracy to nurture these languages (especially those with less than two thousand speakers) and their corresponding cultures, the risk of extinction remains high due to globalization, urbanization and active migration processes leading to the rapid assimilation of these peoples.

It is reasonable that Russia's top priority is protecting, preserving and developing the major state language – Russian – as the language of interethnic communication within the country, an instrument of transnational communication and an official language of international organizations. At the same time Russia advocates for linguistic and cultural diversity. While actively supporting this concept at the international level, Russia makes it a point to implement it consistently in home politics and everyday life – despite the tremendous complexity of this cost-demanding problem, especially in the context of numerous burning challenges and systemic problems our country is facing in the course of drastic changes in all spheres of life.

Support for multilingualism is of great importance for modern Russia. Aside from preserving and developing languages as the basis for the cultural heritage of our country, i.e. the heritage of the Russian people and all other peoples living here, it has always been relevant for tackling political, economic, social and cultural problems, in particular those dealing with interethnic communication in polyethnic environments.

In order to maintain and develop in our modern world languages should be in-demand in cyberspace and get representation there. ICTs open possibilities to decelerate languages' extinction, preserve and even develop them. This chance should not be lost.

Three years ago here, in Yakutsk, I talked on the measures taken in Russia to preserve languages and on the ways of organizing these activities on political and practical levels. This communication is included in the proceedings of the first conference. Our Yakut colleagues can describe in details the problems they are facing and the exemplary ethnolinguistic and sociocultural policies implemented in the Republic of Sakha.

On the basis of Russian political and practical experience, I would like to define the roles and tasks of social institutions that can – and should – hold responsibility for languages' preservation and development in cyberspace.

To develop in cyberspace languages should first of all get development in real life. Three more factors are important, however.

First, tools for multilingualism promotion in cyberspace are required.

Second, institutes are needed to a) elaborate and implement these tools creating attractive and useful content in minority languages and b) teach others to develop, create and use their own content. Information literacy is crucial for both representatives of language support institutions and ordinary users.

Third, positive environment should be created to allow institutes and users work on the elaboration of instruments, creation and preservation of content and on providing proper training.

As we analyse the experience (both positive and negative) of the Russian Federation and one of its entities, the Republic of Sakha (Yakutia), we can attempt to draw out some more general conjectures regarding the question of how to guarantee the continued functioning of minority languages in the shadow of a dominant language in a national context.

Potential contributors to the promotion and development of a language are manifold and diverse, and include:

- national authorities;
- local authorities;
- educative systems;
- research establishments (universities and institutes for scientific research);
- memory institutions (libraries, archives, museums);
- artistic establishments (theatres, philharmonic halls, musical and folk groups, art galleries working in close contact with local painters, sculptors and architects);
- film studios;
- cultural centres, principally in remote settlements, which unite the functions of memory institutions and art and educational centres;
- book publishers and traders;
- media outlets, including digital media;
- the ICT industry;
- public organisations and private persons; and
- businesses.

Let us now consider each of the above-listed contributors, and their corresponding targets and lines of action.

National and Local Authorities

National and local official policies and activities are prime. Effective policies include a combination of active, consistent, complementary steps to stimulate and add value to the activities by all major stakeholders. They should be allocated a duty to facilitate the preservation, free expression and development of linguistic, ethno-cultural and religious identity of ethnic communities, their cultural values, traditions, folklore, as well as the expansion of the domains of national languages' use via the practical application of principles of cultural pluralism, bilingualism and multilingualism.

This goal demands the enactment of special laws (and/or update of the already existing ones) creating favorable environment for the preservation and equitable and authentic development of a country's languages. Monitoring and ensuring compliance with these laws is also essential. They can provide a basis for the formation of a broad system of statutory regulation of activities by legal entities and individuals, and for the elaboration of by-laws.

Constitutions of many countries affirm bilingualism and multilingualism, stressing the equality of languages. Education in the state's official languages is guaranteed, and citizens are often free to choose the main language of education.

Federal and regional language laws should stipulate that the acquisition of the state status by certain languages must not encroach on the linguistic rights and expression by all ethnic entities historically inhabiting a particular territory.

Programmes of socio-economic, national and cultural development should be elaborated based on a set of measures to preserve and develop minority languages and cultures, to extend cooperation of all peoples for mutual intellectual and spiritual enrichment. Respect for customs, traditions, values, and institutions reflecting ethnic cultural specificity is a prerequisite for such programmes.

Authorities should contribute to systemic language studies and multilingualism promotion in education, administration, law, cultural education, news media and cyberspace.

The attainment of those goals can be facilitated by:

- establishing a regulatory framework for the development of languages at the national level (the national constitution and federal laws, along with constituent entities' constitutions, statutes and laws);
- forming and implementing cultural and educational strategies, policies and programmes explicitly aiming to promote minority cultures and languages;

-
- targeting federal funding and soft taxation of both governmental and non-governmental programmes for language preservation and development;
 - granting state or official status to the largest minority languages either at the national level, or within regions densely inhabited by speakers of those languages; whenever possible, language equality must be affirmed in law;
 - affirming a given minority language's official status in the records of government and municipal authorities: using the language in governmental work, publishing federal and republican legal acts in it (and guaranteeing their equal legal force), and granting the language equal standing with the principal state language during elections, referendums and industrial, office and administrative activities;
 - creating official document databases in the language;
 - establishing councils on language policy within central and/or regional governments, and determining their rights and duties;
 - guaranteeing social, economic and legal protection of the language in legislative, executive and judicial bodies;
 - providing material incentives for experts to use both national and minority languages in their work;
 - signing (or lobbying for signing) and ratifying international acts promoting multilingualism;
 - promoting ethnic entities' interest in the development of their languages;
 - establishing targeted regional programmes to preserve culture and language;
 - helping and legally assisting the development of the language's body of literature through financial and other support of book publication and media dissemination, particularly that which is oriented to children and youth;
 - forming and implementing strategies and programmes promoting reading in the native language;
 - partnering with ethnic cultural associations outside the administrative territorial boundaries that are historically densely populated with members of the given ethnic;
 - supporting libraries, museums, archives and other cultural agencies in the preservation and development of minority cultures and languages;

-
- establishing ethnic schools to intergenerationally transmit experiences, traditions, culture and ethics;
 - promoting the ethno-cultural component of education and extending it wherever necessary and possible;
 - equipping public schools with minority language and literature classrooms;
 - contracting the governments of other regions densely inhabited by speakers of a particular language to assist in measures to preserve that language, for example by supplying literature to public and school libraries to enable the study of a given language, and participating in the graduate and postgraduate training of teachers for ethnic minorities; and
 - creating graphic systems for non-literate languages.

Securing languages representation and development in cyberspace gains in importance in the context of rapid Internet penetration in all spheres of modern life. Use of ICT has both positive and negative consequences. On the one hand, it may decrease linguistic and cultural diversity, on the other hand, it opens new prospects for preserving and even developing languages and cultures in cyberspace.

Promoting linguistic and cultural diversity in cyberspace is a new field of action, expanding our opportunities for preserving languages and cultures and for extending the sphere of influence of under resourced languages rather promptly and efficiently.

That is exactly why the Declaration of Principles of the World Summit on the Information Society stresses that the Information Society should be founded on and stimulate respect for cultural identity, cultural and linguistic diversity, traditions and religions, and foster dialogue among cultures and civilizations. The creation, dissemination and preservation of content in diverse languages and formats must be accorded high priority in building an inclusive Information Society. The development of local content suited to domestic or regional needs will encourage social and economic development and will stimulate participation of all stakeholders, including people living in rural, remote and marginal areas.

To promote multilingualism in cyberspace, authorities can take both general and goal-oriented measures to create a multilingualism-friendly environment:

- designing and implementing ICT penetration programmes;
- drawing up action plans to promote public use of the Internet, including information literacy programmes for both dominant and minority languages;

-
- providing telecommunication networks to remote areas;
 - elaborating information resource development programmes in minority languages;
 - promoting training in ICTs and information, especially in local languages;
 - promoting the creation of local content, translation and adaptation;
 - promoting the translation of world literary classics into minority languages, and of minority speakers into other languages, and posting these translations online;
 - establishing integrated multilingual information resource networks;
 - introducing electronic documentation and record management in at least two languages; and
 - promoting the research and development of operating systems, search engines and internet browsers, online dictionaries and term reference books, and their adaptation to local demands.

Research Centres (Universities and Institutes for Scientific Research)

Research centres provide the theoretical basis for governmental and non-governmental efforts on multilingualism promotion and make fundamental and relevant applied research. Their duties may comprise:

- studies of ethnic cultures, traditions and quotidian life;
- studies of languages and their history;
- studies of the current linguistic situation and related issues;
- studies of language-promoting policy and practice in other parts of the country/world, display and dissemination of pioneer experience;
- elaboration of proposals on adapting cutting edge experience;
- elaboration and implementation of permanent monitoring tools to measure language use by social groups;
- elaboration and implementation of permanent monitoring tools to qualify and quantify the work of language-promoting institutions;
- proposals to the government for draft regulatory legislation on language protection and promotion;

-
- initiation and organization of theoretical and applied conferences addressing the various aspects of minority language preservation and development;
 - establishment of research, education and information centres of minority languages and cultures, aimed at conducting research and training for relevant experts;
 - popularization of minority languages and cultures;
 - elaboration of national reading promotion programmes, in particular for minority languages, in cooperation with libraries, educational institutions, media outlets, and book publishers/traders;
 - elaboration of best practices guidelines for relevant offices and organisations, charged with the task of languages and cultures support;
 - publication of bilingual dictionaries that include audio recordings of words;
 - establishment of terminology and orthography commissions;
 - creation of text corpora and phonetic databases;
 - linguistic and folklore field studies and expeditions;
 - establishment of centralised archives, including electronic archives, for minority languages;
 - acquisition of private archives of researchers and community activists (including foreign) engaged in minority language support, and entrusting those archives to state memory institutions;
 - establishment of clear standards and guidelines for recording and representing texts, alphabets and graphic systems for non-literate languages – this is of particular importance for oral languages and languages having recently acquired a graphic system;
 - establishment of a unified literary language, if absent;
 - documentation of minority languages;
 - research and development of operating systems, search engines and information scanning systems; and
 - development of fonts in cooperation with relevant experts.

Education

Primary, secondary and higher educational establishments should cooperate with federal and regional executive and legislative bodies, as well as research and cultural institutions, to support and develop minority languages and multilingualism.

Their sphere of activity includes:

- participating in writing the regional/local component of national educational standards;
- training minority language teachers for schools and universities;
- training experts on languages, history and traditional culture of ethnic minorities;
- implementing postgraduate teacher training programmes;
- elaborating basic curricula;
- elaborating academic curricula and learning packages;
- elaborating language teaching and speech improvement methods;
- making recommendations to implement new language teaching technologies;
- establishing university classes in minority languages;
- using minority languages as educational tools in all places of learning, including pre-school institutions, secondary schools and universities;
- teaching minority language as part of core curricula for students who speak it as a second language in all educational establishments in areas where an ethnic minority makes up a considerable section of the population;
- organising specialist language and literature classes;
- organising educational competitions on minority languages and literature;
- organising conferences and events on linguo-cultural and ethno-cultural issues;
- organising off-campus minority language courses, especially on interregional and international levels (including e-learning in higher education institutions);
- organising summer camps conducted in minority languages; and
- organising online conferences in minority languages (on diverse topics).

Cultural Institutions

Cultural institutions and activists are tremendously important in language support, not only those directly connected with preserving written cultures, but also theatres and conservatories, art schools, folklore performers, cultural centres in remote areas, and individual artists and cultural workers.

Memory Institutions: Libraries, Archives and Museums

It is the duty of these institutions to preserve, store, popularise and offer for public use all essential testimonies of a particular people's history; to elaborate all possible ways and forms of accessing its cultural and written heritage, intellectual and artistic products; to contribute to saturating public space – both real and virtual – with them.

Libraries and archives must search, acquire, describe, study, popularise and store all printed matter, sound and video recordings emanating from a language, both in the geographical area that is densely inhabited by its users, and other areas (even foreign countries) where those languages are used. Not only materials in minority languages but all information about them published in other languages is important.

The activities of memory institutions include:

- gathering, preserving and extending comprehensive and thematic collections of all published and unpublished materials in a minority language;
- creating full-text databases of periodicals in the given language;
- constructing an exhaustive bibliography of printed and written resources in the language;
- making available centralized catalogues of publications in the language (especially important for languages that have recently acquired a graphic system);
- including bibliographic descriptions of works reflecting the history and culture of an ethnic minority in electronic national catalogues of all libraries at both national and international levels;
- popularising these works, especially by organising readers' conferences, reader clubs, and meetings with writers, critics, publishers, illustrators, and others;

-
- digitising documents and museum exhibits that reflect an ethnic entity's history and culture, establishing corresponding electronic libraries, museums and archives, and granting public access to them;
 - establishing electronic and other museum expositions in the given language or bilingual exhibitions using that language;
 - creating electronic catalogues in museum systems in the given language; and
 - preparing archives of electronic publications and exhibitions on cultural and linguistic diversity and memorable dates and events.

Together with other cultural, research and educational establishments, libraries, museums and archives can launch multimedia projects pertaining to the founders of ethnic cultures, folklore collectors, writers, artists, composers and performing musicians to be applied in various fields. Texts, photographs, digital copies of paintings and sketches, sound and video recordings can be recorded on discs for broad circulation, and their online versions be posted on the websites of cultural, research and educational institutions with due respect for copyright.

Mass Media

Today mass media tend to become most important and efficient tools for influencing public opinion, their effect being even stronger than that of education. Federal, regional and municipal media outlets can be purveyors of cultural and linguistic diversity and promote spiritual values exchange. The contemporary mass media should focus on:

- preserving and developing periodicals in minority languages and sections in those languages in other periodicals;
- organising television and radio broadcasting in minority languages, especially the release of programmes entirely or partly conducted in those languages and dedicated to the areas of their active use and the original ethnic culture of their speakers;
- organising internet broadcasting in minority languages; and
- establishing information portals.

Book Publishing and Circulation

Book publishers and traders can make a tremendous contribution to the support of minority languages and development of multilingualism: a language without access to the book industry is a language excluded from intellectual

community life. UNESCO says that “books are in fact a means of expression which live through language and in language” and stresses the importance of translation in strengthening multilingualism, and the urgent need to “give languages broader access to publishing, so as to promote the exchange of books and editorial content, and thus the free flow of ideas by word and image”.

Publishers can contribute to the promotion of minority languages through:

- printing research, popular science and fiction books, periodicals and translations in a minority language;
- promoting literary work in a minority language and its emerging authors;
- assuring that libraries of educational institutions include books in minority languages; and
- helping minority language speakers to acquire books, especially in remote areas that are historically densely populated by the given ethnies, and the diaspora outside the traditional settlement areas.

Public Organisations

Non-governmental language promotion activities include:

- establishing weekend schools, clubs and ethno-cultural associations to provide supplemental linguistic and literary education;
- organising competitions, festivals and creative events to promote cultural and linguistic traditions;
- participating in language and culture days in and outside the traditional settlement areas of a given ethnies;
- participating in folk festivals; and
- communicating with and supporting a language’s expatriate population.

Private Initiatives

Individuals and groups of individuals can also participate in language preservation and promotion by:

- establishing and supporting Wikipedia in minority languages;
- establishing and supporting websites, blogs, Twitter and other social networks.

The ICT Industry

Issues of linguistic diversity in global information networks and universal access to information in cyberspace dominate the agenda of the discussions on information society. The ICT industry should therefore become a crucial participant in supporting and enhancing a language's status. The ICT industry can channel its energy into the following areas:

- articulating and promoting technical standards, taking into account ethnic minorities' demands;
- creating complete computer fonts for minority languages;
- participating in the establishment of international UNICODE standards and the implementation of the unified keyboard layout;
- localising existing software and creating free software to support local languages;
- elaborating computer language models and machine translation systems;
- supporting minority languages in e-mail, chat and other messaging utilities;
- uploading electronic study books and dictionaries in minority languages;
- establishing multilingual domains and e-mail addresses;
- creating software for multilingual internet domain names and content;
- establishing localised, minority language retrieval systems;
- creating information and other websites and portals in bilingual versions;
- making information resources available electronically; and
- developing the non-textual sphere of the internet (such as voice over IP, data streaming, and video on demand).

The above measures can be efficient and bring about their desired results only when the entire ethnos – not only its cultural, intellectual and ruling elite – makes major intellectual and emotional efforts, and displays goodwill, desire and interest in the survival and development of its unique culture and linguistic identity.

Yoshiki MIKAMI

*Leader, Language Observatory Project;
Professor, Nagaoka University of Technology;
(Nagaoka, Japan)*

Katsuko T. NAKAHIRA

*Member, Language Observatory Project;
Assistant Professor, Nagaoka University of Technology
(Nagaoka, Japan)*

Measuring Linguistic Diversity on the Web

The Language Observatory

The Language Observatory [1] Project was founded in 2003. The main objective of the project is to observe the real state of language use on the web. When the first workshop of the project was held on 21 February 2004, the Russian Committee of the UNESCO Information for All Programme kindly reported it in Russian. Then we received several responses from various language communities around the world. This really encouraged us.

How It Works

The Language Observatory is designed to measure the use of each language on the World Wide Web. Measurement is done by counting the number of written pages on the Web in each language.

The observatory consists of two major components. The first is a data collection instrument from the Web, a crawler robot developed at the University of Milan. It can collect millions of Web pages per day.

The second component is a language identification instrument. We have developed software to identify language, script and encoding properties of Web pages with high accuracy and maximum coverage. The first version of the identification algorithm LIM (Language Identification Module) was developed in 2002 [2] and implemented in 2004. The most recent updated version is called G2LI. You can use it on the Web.

According to a recent verification examination G2LI is capable of identifying 184 languages in ISO Language Code (ISO 639-1) with an average accuracy of 94%. In addition to a wide coverage of languages, it can identify various types of legacy encodings², which are still extensively used by many non-Latin-script user communities.

² *Legacy encodings are non-standardised, and often proprietary encodings.*

A Hidden Component: The Universal Declaration of Human Rights

Hidden inside the language identification instrument is a set of training texts for the software. Considering that the richness and quality of training texts is the most critical in language identification task we used a set of translated texts from the Universal Declaration of Human Rights (UDHR) provided by the UN Higher Commission for Human Rights (UNHCHR).

Of note is that not all translated UDHR texts are provided with encoding; some are available only as image files. Image files can be read by humans but not directly by computers, necessitating that we transform images into text data. Table 1 illustrates how many transformed texts are given in image format (322 languages were available at the date of the first search, in early 2004). More than two hundred languages use Latin script, with or without diacritics, and only three of them were given in PDF or GIF file format. In contrast to this, among languages using so-called Abugida script³, not a single language was presented in the form of encoded text. This fact might itself point to the existence of a digital language divide, or in this particular case, a “digital script divide”.

Table 1. Number of available UDHR texts from the UNHCHR website by format

	Latin	Cyrillic	Other alphabet	Abjad	Abugida	Hanzi	All others	Total
Encoded	253	10	1	1	0	3	0	268
PDF	2	4	2	3	10	0	4	25
GIF	1	3	0	9	15	0	1	29
Total	256	17	3	13	29	3	5	322

Other alphabets: Greek, Armenian and Georgian; Abjad: Arabic and Hebrew; Abugida: Amharic and all Brahmi origin scripts used in South and Southeast Asia; Hanzi: Chinese, Japanese and Korean; All others: Assyrian, Canadian syllabics, Ojibwa, Cree, Mongolian and Yi.

Around the same time as we launched the Language Observatory Project, Eric Miller launched UDHR-in-Unicode project. The objective of this project was to demonstrate the use of Unicode for a wide variety of languages, using the Universal Declaration of Human Rights (UDHR) as a representative text. Currently, UDHR-in-Unicode is housed on the Unicode Consortium website and the texts are used in the study of natural language processing.

³ *Abugida scripts are syllabic scripts, most of which are generated from Indian Brahmi scripts and currently used in South and Southeast Asian regions. Another important Abugida script is Amharic.*

Sponsors and Collaborators

The Language Observatory project was initiated by the authors in 2003 and received funding from the Japan Science and Technology Agency from 2003 to 2007. At the kick-off event, held at Nagaoka University of Technology on February 21, 2004, UNESCO Communication and Information (CI) sector was represented by Mr Paul Hector.

The project interacted and collaborated with many partners from various parts of the world, and joined with the African Academy of Languages (ACALAN) at WSIS in Tunis in November 2005 at a workshop on African languages. Among the attendees were the President of ACALAN Adama Samassekou, Daniel Pimienta of FUNREDES, and Daniel Prado of Union Latina.

We agreed to organize a joint African web language survey project. In 2006, we held a workshop in Bamako, Mali, in cooperation with ACALAN. Many African researchers interested in language diversity and the digital divide issues attended the workshop.

After a fruitful workshop in Bamako, we planned a set of events to publicize our project. Workshops were held at the UNESCO headquarter in Paris on the International Mother Language Day in 2007 and 2008.

The first complete observation report was published in 2008. It was the first article on language distribution on the Asian Web providing an overview of web pages collected from Asian domains. The authors concluded that there is a serious digital language divide in the region. English was very widely used especially in South Asian countries and in the majority of South East Asian countries (60% of the web pages were in English). In West Asia, English dominance was less outstanding, and in some countries Arabic was the most widely used language. In Central Asia, Russian was the most widely used language except for Turkmenistan where English was used at 90% of the web pages. It is also important to notice that some of the indigenous languages, Turkish, Hebrew, Thai, Indonesian, Vietnamese and Mongolian were the most used languages in their country domains.

Linguistic Diversity on the Web

Lieberson's Diversity Index

Lieberson's Diversity Index (LDI) [4] is a widely used index of linguistic diversity that is defined by the following formula, where P_i represents the share of i -th language speakers in a community:

$$LDI = 1 - \sum P_i^2$$

If anyone in a community speaks the same language, then $P_1 = 1$ and for the speakers of other languages, $P_i = 0$. Thus the LDI of a completely monolingual community is zero. If four languages are spoken by an equal number of people, then $P_1 = P_2 = P_3 = P_4 = 0.25$ and the LDI can be calculated as $LDI = 1 - (0.25)^2 * 4 = 0.75$. Thus a higher LDI means larger linguistic diversity and a lower LDI means lower diversity.

The basic idea of LDI can be explained by the illustration in Figure 1. A square of P_i means the probability that the i -th language speaker meets with a speaker of the same language. And the sum of P_i squares represents the combined probability of any speaker meeting with a speaker of the same language in the community. Then the sum of P_i squares is subtracted from 1, indicating the probability that any speaker will encounter different language speakers in a society. The dark-colored areas of the square in Figure 1 correspond to this probability.

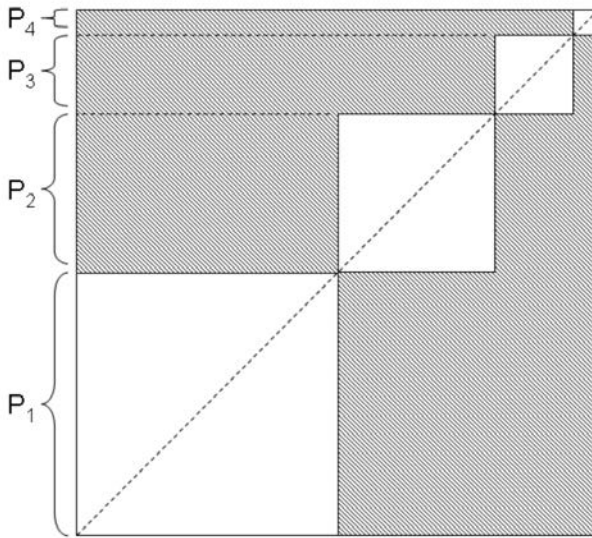


Figure 1. A graphic interpretation of LDI

Ethnologue provides a complete list of LDI data for each country or region, together with population size and the number of indigenous and immigrant languages. Based on this data⁴, Figure 2 was prepared to show how LDI changes across countries and across continents. Each circle represents a country in this chart. The circle's size corresponds to the country's population, and its vertical axis represents the country's LDI. The two large circles on the axis of Asia correspond to India (LDI = 0.94) and China (LDI = 0.51).

⁴ Based on the Web version, an equivalent of the 16th edition of *Ethnologue*.

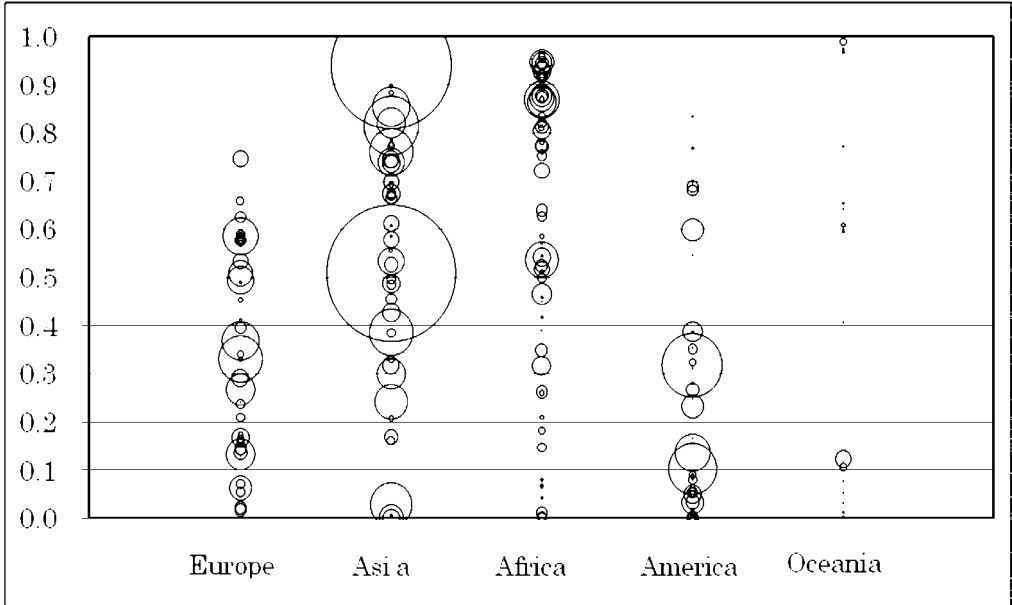


Figure 2. Lieberman's Diversity Index of countries by continent (based on the data from *Ethnologue*)

As the chart illustrates, countries in the African continent have the highest language diversity among the continents, followed by Asia, Europe, America (North and South America included) and Oceania.

Local Language Ratio

In the previous section, we reviewed the overall condition of linguistic diversity of the world based on data by *Ethnologue* that reflects the situation *in the real world*. And what about language diversity *in the cyber world*?

Since being launched, the Language Observatory has focused its attention on two continents, Asia and Africa. The first observation results were reported during a workshop organized at UNESCO headquarter in February 2005; they are fully documented in an article published in 2008 [3]. Recently, the project has completed another round of surveys of Asia, Africa and the Caribbean region based on 2009 data. The following sections will introduce an overview of this most recent study.

Here we propose a two-dimensional chart, which is tentatively named the LL-chart, because the chart has the Local Language Ratio on the horizontal axis and the LDI on the vertical axis.

The purpose of this chart is to solve a problem we encountered when preparing an LDI chart based on data from cyberspace. It often happens that languages used on the Web are completely different from languages spoken in the real world. In many cases, the latter consists of local languages while the former mainly consists of global languages like English, French or Russian. And in those cases, the LDI of languages in cyberspace and that in the real world are not considered to be the same. We have to take into account some measurements about the presence of local languages, as presented in Figure 3.

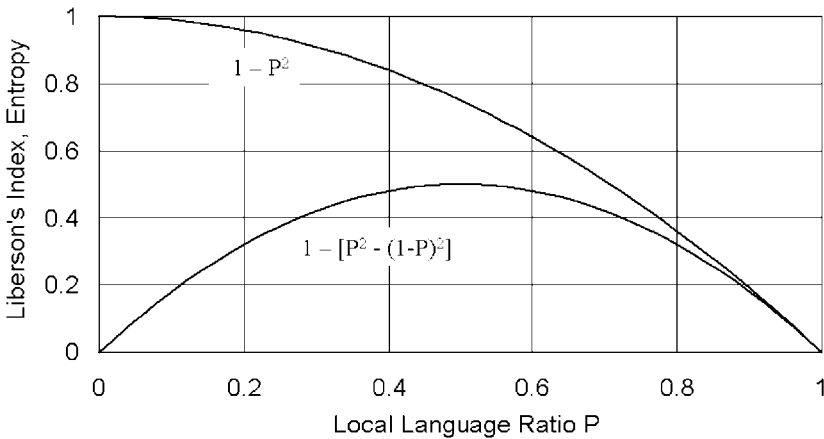


Figure 3. Schematic diagram of the LL-chart

Notice that all countries with a local language ratio P fall within the area between the two curves $1 - [P^2 + (1-P)^2]$, Liebersson's index in the case of two languages, and $1 - P^2$, which gives the maximum value of Liebersson's index⁵. When P becomes larger than 0.5, the LDI becomes smaller and the plotted point will move towards the bottom-right corner. When P is small, there are two possibilities: either the vacancy of local language is filled by a dominant foreign language, in which case the LDI shrinks and the point moves down

⁵ Two curves provide the upper and lower limits. The upper curve indicates the LDI of a two-language community. As the addition of a third-language speaker to this community increases the average probability to encounter different language speakers, this value is the minimum LDI of more than two language communities. The lower curve indicates the LDI of a very special case, where each member, in addition to the local language, speaks another language, or the maximum LDI.

and to the left; or the vacancy of local language is filled with multiple foreign languages, in which case the LDI grows and the point moves up and to the left.

Comparison by Region: Asia, Africa, Europe

Based on data collected in November 2009, the LDI and local language ratio were calculated for all country domains in Asia and Africa. As we do not have data for European countries, we used Google's page count by language. Figures 4, 5 and 6 show the Local Language Ratio – LDI chart for these three regions.

Asian LDIs are plotted in Figure 4. China, Japan and Korea and some Arabic-speaking countries (Iraq, Saudi Arabia, and Jordan) are found in the bottom-right corner, while Vietnam, Thailand and Indonesia, Israel, Turkey, Georgia and Mongolia show a relatively high local language presence.

Of note here is the context of central Asian countries. Their Web spaces are composed of local languages, with major components of English and Russian, although the emphasis changes by country. Kazakhstan, Kyrgyzstan, Tajikistan, and Uzbekistan have a major emphasis on Russian, while only Turkmenistan has an emphasis on English.

On the other hand, web contents in the Indian subcontinent have a nearly negligible local language presence on the Web. More than 70 % of these Web contents are written in English.

Worth mentioning here is the case of Laos. According to *Ethnologue*, the country's LDI is only 0.674. Why then does it have such a high LDI on the Web? The major reason for this is that the ".la" domain is actively marketed to foreigners, including customers connected to Los Angeles. As the domain is sold mainly to foreign industries and peoples, only 8% of web pages of ".la" domain are written in Lao.

LDIs of African domains are plotted in Figure 5. The presence of local languages in African domains is far rarer than in Asian domains. The local language claims the majority only in Sudan and Libya. However, several countries show high Web LDIs.

The LDIs of European and some Anglophone country domains are plotted in Figure 6. Local language presence is above 50% with the exception of Slovenia and Denmark (those countries' web spaces are dominated by English), which results in a lower LDI. At the opposite extreme is the United Kingdom, which joins other Anglophone countries (USA, Australia and New Zealand) in displaying a characteristically low LDI.

Table 3. Language composition of the Asian and African web domains

African Domains			Asian Domains		
Language	Number of pages	Percent	Language	Number of pages	Percent
English	30,327,396	78.40%	Chinese	7,832,521	20.46%
French	2,737,455	7.08%	Japanese	5,287,655	13.82%
Afrikaans	660,510	1.71%	English	4,867,355	12.72%
Arabic	592,746	1.53%	Russian	1,611,339	4.21%
Chinese	391,745	1.01%	Korean	1,100,232	2.87%
Portuguese	348,131	0.90%	Vietnamese	710,048	1.86%
Russian	307,178	0.79%	Thai	544,561	1.42%
Spanish	276,126	0.71%	Indonesian	308,894	0.81%
Japanese	158,992	0.41%	Hebrew	89,076	0.23%
Others	879,605	2.27%	Others	14,055,334	36.72%
Not identified	2,005,311	5.18%	Not identified	1,867,355	4.88%
Total	38,685,195	100.00%	Total	38,274,370	100.00%

Web data obtained from the country-code domains of Asia and Africa in November 2009.

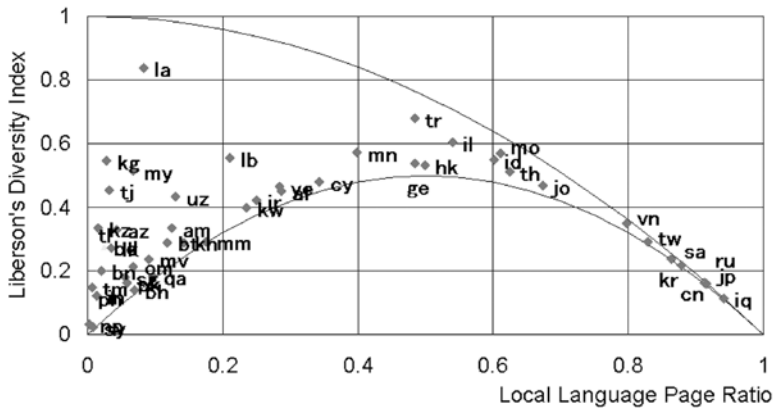


Figure 4. Local Language Ratio and LDI of language composition – Asian domains

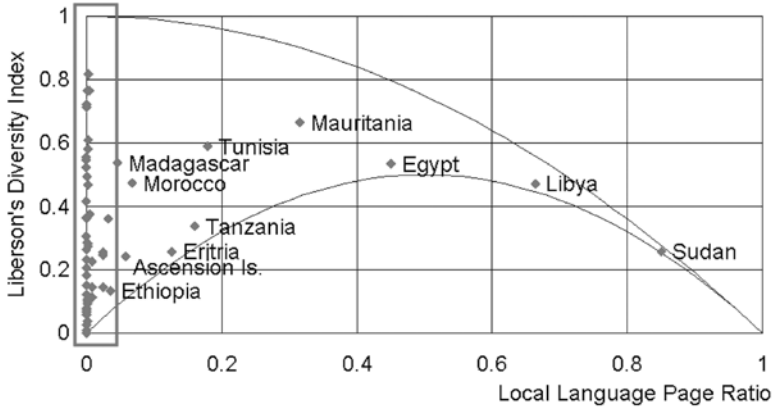


Figure 5. Local Language Ratio and LDI of language composition – African domains

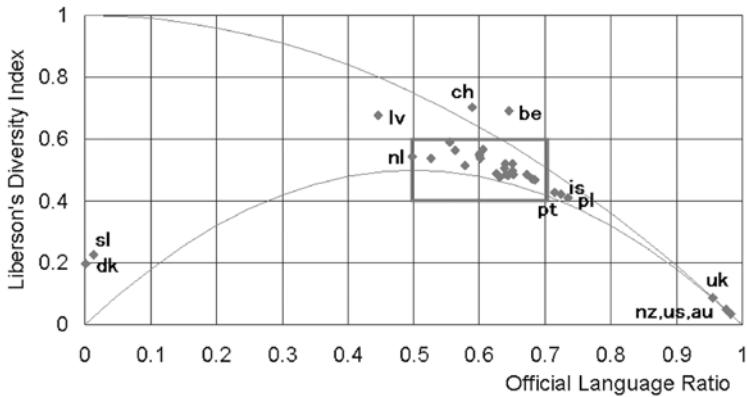


Figure 6. Local Language Ratio and LDI of Language Composition – European Web

Challenges and Future Directions

Though we have managed the project since 2003, the survey of the Web is becoming difficult year by year. The most serious challenge to the surveys comes from the sheer size of the growing Web. Nobody knows exactly how many web pages exist on the entire Web. In 1997, the number was estimated at only 320 million. In 2008, Google announced 1 trillion URLs on the Web, but it has since stopped providing data.

Also we need a deeper analysis of the Web, not just a language-wise counting. The Web and its hyperlink structure can tell a lot about who and how is using the Webs, and what content is written there.

Currently my friend Daniel Pimienta is preparing a new, ambitious project to achieve these goals.

References

1. Yoshiki Mikami, Zavorsky Pavol, Mohd Zaidi abd Rozan, Izumi Suzuki, Masayuki Takahashi, Tomohide Maki, Irwan Nizam, Massimo Santini, Paolo Boldi, and Sebastiano Vigna. The Language Observatory Project. 2005. Proceedings of the 14th International World Wide Web Conference, p. 990.
2. Izumi Suzuki, Yoshiki Mikami, Ario Ohsato. 2002. A Language and Character Set Determination Method Based on N-gram Statistics. ACM Transactions on Asian Language Information Processing, Vol. 1, No.3. pp. 270-279.
3. S. T. Nandasara, Shigeaki Kodama, Chew Yew Choong, Rizza Caminero, Ahmed Tarcan, Hammam Riza, Robin Lee Nagano and Yoshiki Mikami. 2008. An Analysis of Asian Language Web Pages. The International Journal on Advances in ICT for Emerging Regions (ICTer), Vol.1 No.1. pp. 12-23.
4. Stanley Lieberman and Anwar S. Dil. 1981. Language Diversity and Language Contact: Essays. California: Stanford University Press.

Daniel PIMIENTA

*Director, Networks and Development Foundation FUNREDES
(Saint Domingue, Dominican Republic)*

DILINET: Building Indicators for Linguistic Diversity in the Internet

The theme of linguistic diversity, in the broader frame of cultural diversity, is transversal to many society matters (from education to business) and is emerging at the center of many actual debates. Do languages – as assets of humanity - require public policies to be preserved, promoted or supported? Is English the accepted *lingua franca* for international research collaboration and business? To what extent has the business globalization opened inescapable requirements for marketing in other languages?

These questions, that are essential for the development of information societies, gain even more relevance when they are referred to the Internet, a space which has seen the initial dominance of English getting more balanced, both in terms of users and contents, in the last years, as a consequence of the intense spread of the Net in many regions. Is the “digital divide” a simple issue of access or shall the content divide and subsequent linguistic divide also be addressed? Will translation of contents be a workable and acceptable panacea for multilingualism?

The theme of linguistic diversity on the Internet, which has been for a long period dealt by a small group of specialists, is gaining now the attention it deserves among policy makers and many stakeholders. As a side effect from the recent development of internationalized domain names (IDNs) more public awareness has been gained and the theme is becoming a central topic in international agendas as seen in the Internet Governance Forum (IGF) and in the main organizations in the field of the information society, such as UNESCO and the International Telecommunication Union (ITU). Furthermore, the issue is also rising in the agenda of business, and calls for a clear policy framework for the virtual world.

As for indicators, a recent publication by UNESCO [1] reports on the current situation and evaluates future perspectives: the situation it describes is indeed paradoxical and quite alarming. Until the late 90s, this field was marked by a lack of serious indicators; this period was followed by the preliminary work of a handful of pioneers, which provided some indicators, most of them limited to number of users and the split of the World Wide Web per language. However, now that interest in the theme is becoming visible, the existing works aiming

at measuring the linguistic diversity in the Internet are being undermined by the almost infinite size of the web as well as by the evolution of search engines. Accordingly, no reliable indicators have been produced since 2007, when the Language Observatory Project (LOP)⁶ and Funredes/ Union Latina, the two most visible actors in producing indicators, published their most recent results.

In this context, and starting from the fact that it is hardly possible to formulate policies in any field without a clear vision of the situation, only obtained from reliable and frequently produced indicators, **it is urgent to mobilize existing actors and to encourage new ones to engage in an ambitious, serious and collective research effort of building indicators for linguistic diversity in the Digital World.** This effort should both build on existing approaches and explore new methods rather than those involving a static vision of supplied linguistic resources, thus also informing on the demand side (user behavior). DILINET will also introduce the first attempt to measure automatically content characteristics while recognizing languages and use conceptual maps and visual analytics to extract meaning to statistical data on languages in the digital world.

Project Description

In response to this context, the DILINET project aims to develop a set of methods for producing indicators of linguistic diversity on the Internet which will therefore support informed public policies in all fields related to the Information Society, at national and global level.

DILINET will adopt an **exploratory research approach**, taking into account existing measuring methods and adding innovative approaches, including users-based measuring systems. To overcome the limitations created by the size of the Web, the project will develop crawling optimization methods based on non sequential mathematical or statistical approaches and use both distributed and super computing resources while opening new avenues such as recognition techniques for voice or automatic content characterization.

Given the transversal relevance of the theme, the project will devote effort in **engaging all the relevant stakeholders** from the policy field (at regional, national, European and international level) and from the research field (both with public and private background). Specific attention will be paid to **raising awareness** and disseminating project results among users' communities and organizations working in the sphere of linguistic technologies.

⁶ *Language Observatory Project* – <http://www.language-observatory.org/>. See reference [4].

The project stems from the motivation of a group of **key international organizations** such as UNESCO, ITU, Organisation Internationale de la Francophonie (OIF) and Union Latina and will be built on the **professional experience** of some MAAYA (the World Network for Linguistic Diversity) members such as Funredes or the Language Observatory Project.

Results and Impact

DILINET will provide the crosschecked and validated values for a set of indicators of linguistic diversity in cyberspace, as well as trustable and sustained methods and processes to maintain a frequent production allowing perceiving the trends and being able to gauge the results of policy actions. Additionally, DILINET will provide effective awareness raising and training activities for policy and research stakeholders on the matter of linguistic diversity on cyberspace.

DILINET will represent a historic breakthrough for the issue of linguistic diversity in the Digital World opening the floor for a professional approach to this emerging theme and, as a side effect, contributing to the change of paradigm of the vision of the digital divide, switching the perspective from access to content. Indirectly, DILINET will also open new promising avenues for the production of Information Society Impact Indicators and create opportunities for the professional consideration of languages as a key parameter of the digital economy.

References

1. D. Pimienta, D. Prado, Á. Blanco, Twelve years of measuring linguistic diversity in the Internet: balance and perspectives. UNESCO, 2009. Retrieved from: <http://unesdoc.unesco.org/images/0018/001870/187016e.pdf>.
2. ITU, Measuring the Information Society: the ICT Development Index, ISBN 92-61-12831-9, 2009.
3. J. Paolillo, D. Pimienta, D. Prado, et al. Measuring Linguistic Diversity on the Internet. UNESCO, 12/2005. Retrieved from: <http://unesdoc.unesco.org/images/0014/001421/142186e.pdf>.
4. Suzuki I., Mikami Y., and al. "A Language and Character Set Determination Method Based on N-gram Statistics", in ACM Transactions on Asian Language Information Processing, Vol.1, No.3, September 2002, pp.270–279.

Dwayne BAILEY
Director, Translate.org.za
(Pretoria, South Africa)

Multilingual Moments of Truth: Moments that Define a Person's Future Perception of Multilingualism

What is a Multilingual Moment of Truth?

Walk up to a counter at an airport as an irate customer who has just lost her baggage. How the airline employee handles the issue could be the defining moment for you and that airline. You could change your opinion from hate to love depending on your experience. Depending on that moment of truth.

We also have these moments of truth in language, we might not be aware of them, but they do exist. The moment a child decided to forget their mother tongue and speak only a dominant language, that is a moment of truth.

Thinking back on your own life there are these moments, good and bad. Some examples that ***Translate.org.za*** has seen out in the field:

- A teacher telling a passionate Afrikaans linguist that they should choose another career as there is no money in languages;
- Reading a novel that switches on love for your language;
- A praise singer singing the praises of a president in Xhosa;
- Walking into a library to see an empty Zulu shelf, confirming that your language really is only for conversing at home;
- Writing a Xhosa poem and seeing every word underlined in red.

This paper covers a number of activities that ***Translate.org.za***, as well as other members of ANLoc (The African Network for Localisation), have undertaken that try to change these moments of truth into positive moments for multilingualism.

Translate.org.za is a South African non-profit organization focused on removing the barriers that exist in technology that prevent people from working in their mother tongue. The organisation also works to increase the volume of content and content platforms that allow mother tongue speakers to produce local language content.

Immersion

Dipping your toe into the sea is not immersion. Dipping into language is not immersion. Immersion is where you take your whole body and submerge it into the sea. Immersion in language is the ability to do everything in your language.

At *Translate* immersion is our vision, while for most languages it might not be a reality it is a goal for which we strive.

For a language that is fully immersed this would mean: you switch on your computer and it boots in your language, your emails are written in your language and spell checked correctly, your keyboard works for your language. Your cellphone, TV, ATM are all in your language. When you surf the web you can get content: news, wikipedia, facebook in your language. You look for books online and you find ones written in your language.

That is a fully immersed language.

First Aid Kit for Languages

Moving from a dream to reality. We often hear talk of the 6000 languages of the world, often that figure is used to indicate how fast our languages are dying.

Why are we so fascinated by languages dying? Is this a morbid fascination with death and destruction? We really shouldn't care if languages die. Time has told us that languages come and go. What we should be very concerned about are the people behind those languages. If people are dying then we really should be concerned.

The reality is in many cases people are dying. Without language they have poor access to education, health care and employment. Without these, people's lives are impacted often leading to them not reaching their potential, dying of treatable diseases and not being able to sustain their own families.

Instead of bemoaning the dying languages we need to focus on how to give them first aid. First aid of course is medical intervention before real medical care commences. What is critical for a written language is the ability to:

1. Capture content
2. See content
3. Store content

We achieve this with a keyboard, to type in content; fonts, to see the content and a locale to store the information correctly on a computer. With these

things in place it is possible to do all the other language interventions that are needed. You can't translate any application without a locale, you can't enter content into a blog or wikipedia without a keyboard and you can't see any of this content in documents or on the web without a font for your language.

All the exotic language applications such as automatic speech recognition (ASR), text-to-speech (TTS) and machine translation (MT) cannot proceed without these basics.

In the ANLoc project we created almost 100 locales, 12 keyboard layouts and a font to cover all of the latin characters used in African languages. For a large number of languages we can capture, see and store content – we applied first aid for these languages. And it can be applied to any marginalised language across the globe.

Empowered Communities

Marginalised languages often have limited resources, but one resource that they do have is a community of speakers. For marginalised and under resourced languages to prosper is really about empowering these communities because then the community can take our work to new heights and in fact take over our work.

Within the scope of limited resources this often means that there are no funds to pay contributors so there needs to be a body of committed volunteers. But scarce resources also means that good skills are in limited supply. Thus it is important that these scarce skills are used optimally. What that means is that those skills are deployed to important tasks while volunteers assist in other areas. Empowered communities in this case are assisting in optimising resources allocation for the benefit of the language.

The following are examples that we have used to empower communities:

1. Translate@thons – ***Translate.org.za*** adopted this approach to community translation for working on various pieces of software. Google now uses a similar approach to use communities to translate the Google interface.
2. Pootle – a web-based tool developed by ***Translate*** to allow easy community participation in translation tasks. During the ANLoc project seven teams translated Mozilla Firefox into African languages using Pootle as their tool for translation. A number of open source and commercial companies now use Pootle for community translation.

Translate is currently working on the following areas that involve community participation:

1. Terminology – with terminologists being scarce resources it is important that communities assist in gathering, proposing and voting for terminology. **Translate** are currently developing an expert lead terminology process and platform together with the Kamusi project with funding from International Development Research Centre (IDRC).
2. Dictionaries – in South Africa most children do not have access to a printed dictionary in any language. Such a critical language resource is simply not available. **Translate** is helping to develop an online dictionary platform that will ensure that dictionaries can be built by communities and be made available on many different formats.
3. Books scanning – many African language resources are locked away on the shelves of libraries. These public domain works are critical resources for speakers of the language to be able to read and be inspired by works in their language. They are not easily available as they are often in rare collections or out of print. These resources in digital format are also critical for the creation of linguistic resources such as spell checkers. **Translate** has begun a project in South Africa to recover African language books through scanning and digitisation.

While Talking About Books

Children love to read and they don't need encouragement to read. What they do need are exciting books in a medium that they enjoy. For young readers to find exciting books we need lots of books and the medium is most likely the cellphone.

Thanks to the m4lit project for having this quote demonstrating the value of content on mobile phones:

It's great ... for me it really hard to pick up a book to start readin but i don mind readin on my phone (Dotty1)

We hope that by scanning the books found in marginalised languages we can provide a resource: firstly, lead to someone discovering their first exciting book. Secondly, but more importantly, it should create a reading culture and produce new creative works in the language.

Copyright and Public Domain

Where do we get large volumes of books? From the public domain, which contains books whose copyright term has expired and thus have been returned to the people. In our project we are scanning books in order to make them available at low cost. In this way anyone can own a collection of Tswana classic literature.

Our primary objective is of course reading, the secondary being linguistic resources. It is these linguistic resources that can lead to some exciting language tools. The linguistic data present in public domain works allows us to create spell checkers, grammar checkers, text-to-speech engines and machine translation.

Other sources of public domain data that we are interested in and in which *Translate* is developing solutions include:

1. Assisting with the capture and dissemination of Hansard (the verbatim transcripts of parliamentary debates produced in a number of commonwealth nations)
2. Government website translation – in countries where translations of government resources is regularly performed or mandated by law these translations are a valuable linguistic resource for the language.

These sources of open linguistic data are critical for the advancement of marginalised languages. Thus it is important that marginalised languages are active in ensuring that linguistic resources such as these are made available for the advancement of the language. In making these resources available it is critical that their availability be judged not by access to the resources, but by what new resources can be produced from them. Thus licensing that allows academic use but prevents the creation of a commercial spell checker are not in the best interests of the language or the community.

In a similar vein the ongoing extension of copyright is problematic for marginalised languages. By extension we mean the move from the current international norms of 50 years of copyright to terms in excess of 70 years. Each extension means that more works are not available to the language. If we consider that copyright approximately 100 years ago was anything from 14 to 28 years and that now it ranges from 50 to 70+ years we realise that marginalised languages have lost much of their public domain content. For marginal languages it is important to consider what benefit copyright extension has to the overall health of the language.

Of interest to these languages is the championing of terms such as those found in the Egyptian copyright law that give authors 3 years in which to exercise their right to translate. If they fail to translate their work into Arabic within that time then anyone may translate their work into Arabic. This simple clause could dramatically stimulate the production of content in local languages.

Choosing Closer Languages

Linguistically speaking English is very far from either of the languages, Xhosa and Zulu. Xhosa and Zulu are both part of the Nguni language group in South Africa and are therefore linguistically closely related.

Many efforts to use machine translation focus on the long distance translation of English to Zulu, French to Zulu or similar. While we won't discount the value of these it is worth considering that by machine translating from Zulu to Xhosa we grow a 10 million strong community into an 18 million strong community.

There is an unexploited strength in these close communities that goes beyond machine translation. Zulu is a Bantu language and so is Swahili. Swahili has an estimated 140 million speakers. It is much easier to translate content from Swahili to Zulu. And by translation in this case we don't only mean machine translation, we mean human translation as well as the fact that the closer alignment of the cultures makes translation of cultural metaphors so much easier.

For marginalised and minority language we really should be examining how we can exploit the closer relationships to grow the language speaking community, grow the financial viability of languages and grow the limited resources by the pooling of resources.

Journey

Our efforts are part of a journey. Travellers share their meager possessions, but we in the language community are not very good at sharing our resources.

When it comes to sharing, the creation of open linguistic data seems to be most critical. Resource rich languages can afford not to share because ultimately there are enough resources available. But resource poor languages have to share. But even in the act of sharing we have witnessed resource poor languages creating open resources that could not be combined or could not be used in open source applications. These situations prevented open resource sharing.

Thus it is important that we create a common ground on how resources should be shared. As we said earlier this should not be defined by the act of sharing but by the outputs that can be created from the resources.

As an example, in South Africa, we estimate there are four Zulu morphological analysers. This is clearly a waste of scarce resources in a language with limited access to funds to advance the language. The reason there are four is that there is no framework for sharing these resources. But ultimately it is because there is no clear objective on how these resources and tools should impact on the lives of real people. When it comes to tools for processing linguistic data, the same logic applies. We really do need to learn how to share so that we can focus on the work that impacts language speakers.

This Is a Wave and This Is You

If you look at the wave that is marginalised languages, those 6000 all crashing down on us at once, then we all want to run away. But if we focus on riding the wave then it could actually be quite fun.

We at *Translate* have realised that it is in some ways about changing our thinking. We want to build solutions that meet the needs of language speakers. But we've realised that sometimes we need to address other people languages needs, and in so doing we create linguistic resources that we can employ to address the needs of our marginalised languages.

For resource poor languages using modern technology to give oral literates the resources and tools to be active e-literate participants is the type of thinking, or wave riding, that is required.

William W. McLendon, Jr.

*Associate Director,
International Information and Technology Services,
SIL International
(Dallas, USA)*

New Developments and Trends to Facilitate Multilingualism in Cyberspace

Introduction

SIL International is a faith-based, non-profit organization committed to building the capacity of language communities worldwide for sustainable language development. We define language development as the series of ongoing planned actions that a language community takes to ensure that its language continues to serve its changing social, cultural, political, economic, and spiritual needs and goals.

From its beginnings in 1934, SIL has had the privilege of working with over 2,590 language communities representing more than 1.7 billion people in nearly 100 countries. SIL's staff see their work as an outgrowth of their Christian commitment, valuing service, academic excellence, sharing of knowledge and partnership as we serve language communities in linguistics, literacy, translation and other language-centered development activities.

In addition to language development activities undertaken with individual language communities, SIL takes an active role in advocacy for minority languages at the local, national and international level. Please see our web site at <http://www.sil.org/> for more information.

The increase in global collaboration on the Internet in recent years has opened significant opportunities for minority language communities. We are witnessing an increased visibility of and support for the needs of these communities from the governmental, commercial and non-profit sectors. Collaborative efforts to facilitate the use of all the world's languages in cyberspace are growing.

This report highlights several significant efforts and SIL's participation in them. It reflects the work of many SIL colleagues and partners worldwide⁷.

⁷ Some information in this report is repeated from the author's 2008 report at the I International Conference on Linguistic and Cultural Diversity in Cyberspace (Yakutsk, Russian Federation, 2008).

Organizing Information About the World's Languages

In order to facilitate the use of all languages in cyberspace equally, members of the language communities themselves and other interested parties need to be able to collect, organize and share information in and about all languages accurately and consistently. Multiple organizations, including the International Organization for Standardization (ISO), the Unicode Consortium and others, develop and support standards that facilitate this work.

SIL International supports these efforts to enhance data sharing based on these standards. The following paragraphs will provide information on some recent developments.

Introducing ScriptSource

In June 2011, SIL International released a powerful web-based resource called ScriptSource. ScriptSource is a dynamic, collaborative reference to the writing systems of the world, with detailed information on scripts, characters and languages.

A writing system for an individual language is composed of several elements including the script and individual characters. A language may be supported by several writing systems based on historic, political, educational or other reasons. Scripts, characters and language names are all currently codified by international standards, however the complex relationships between these elements are often not well understood or well documented.

ScriptSource seeks to bring together authoritative information from standards and other recognized resources with contributions from the global community to document the writing systems of the world. We hope this resource will help bring visibility to the richness of languages and writing systems, as well as to help researchers, designers, linguists and software developers with the complex task of supporting the world's languages in different information and communication technologies.

While almost 7000 living languages are spoken in the world today, only around 130 scripts are used to write them—if they have a written form. An individual script, such as Latin, Arabic or Cyrillic, is thus used to write many different languages. The set of characters used to write a specific language within a script is based on the linguistic characteristics of the language. These, and other elements, go into defining a writing system for a language.

For minority languages that use a majority language script, documenting the elements of the writing system makes it possible to develop computing solutions that support the language. As an example, while the use of Cyrillic

script for major languages is well documented, its use for some minority languages is not consistently documented or well understood, and that use continues to evolve and change.

SIL is investigating the possibility of an extended Cyrillic documentation project that would bring together information from a wide variety of sources to catalog and describe the use of Cyrillic script by all languages that use it. We believe that ScriptSource can be a resource to facilitate that global discussion, and would like to see all interested parties included in the process. Please contact us if you feel this would be a useful project or if you would like to be involved.

Comprehensive font development work is already underway. We acknowledge and appreciate the work of the font foundry, ParaType, in developing and releasing the PTSans and PTSerif fonts. These free/libre fonts are designed to support minority-language use of Cyrillic and Latin scripts in the Russian Federation. This development represents a strong partnership between government, non-government and commercial organizations to benefit the minority language communities. The fonts are licensed for the widest possible sharing and use. SIL also supports a wide range of extended Cyrillic characters with the Gentium, Charis, Doulos, and now Andika free/libre fonts. We intend to improve and broaden that support as more information becomes available.

Beyond font development is a need to document which specific characters are used in a given language and to develop the standard locale information and other components for a complete writing system reference. Commercial and open source software developers use this information to support a specific language in computing applications. We hope to simplify the complex process of documenting the information needed for an individual language to be used in cyberspace and telecommunications. Once the information is clearly organized, we hope the process for submitting it to the appropriate standards-setting bodies will be clearer and more accessible.

To investigate and contribute to ScriptSource, please visit the website at <http://scriptsource.org/>. For further information on the ParaType fonts, please see their website at <http://www.paratype.com/public/>. SIL's fonts are available from <http://scripts.sil.org>. Finally, to learn more about the Cyrillic documentation project concept, please contact Victor Gaultney of SIL at <Victor_Gaultney@sil.org>.

Update on the Open Language Archives Community

The Open Language Archives Community, or OLAC, is an “international partnership of institutions and individuals who are creating a worldwide

virtual library of language resources”. OLAC facilitates sharing of information on languages by “developing consensus on best current practice for digital archiving of language resources”, and by “developing a network of interoperating repositories and services for housing and accessing such resources”. Currently 45 organizations are participating in OLAC and SIL is an active contributor.

In 2010, OLAC achieved a significant milestone in making available the OLAC Language Resource Catalog, a web-based search engine that accesses information from the combined catalog of over 100,000 items of the 45 participating language archives. The search engine supports faceted search, enabling a user to pinpoint resources based on various categories, or facets, of information. This powerful tool illustrates the benefit of shared standards for cataloging language archives. OLAC continues to invite other institutions to link their language archives into this global resource.

Further information is available from the OLAC website: <http://www.language-archives.org/>. The new search engine is available at: <http://search.language-archives.org/>, and is linked from the OLAC home page.

Update on ISO 639-3 Standard Language Codes

The ISO 639 family of standards was expanded in February 2007 with the formal adoption of ISO 639-3. This standard seeks to provide a comprehensive list of human languages, including living, extinct, ancient, and constructed languages, whether major or minor, written or unwritten. It provides a unique three-letter code for each language along with limited meta-data about the language.

As the Registration Authority (RA) for Part 3 of the standard, SIL processes requests for changes to the language codes. We receive and review requests for adding new language codes and for changing existing ones according to criteria defined in the standard. All update requests undergo a period of public review before being acted upon in a yearly review cycle.

In 2010, thirty-seven requests were considered, recommending fifty explicit changes in the code set. After a public review and comment process, 32 were fully approved and 5 were rejected:

- 4 new language codes were created: 3 for living languages and 1 for an extinct language.
- 19 language names were either changed, or additional name forms added.
- 1 language had another language variety merged into it.
- 8 language codes were retired: 2 language codes were merged and 4 language codes were split.

The ISO 639-3 standard is used in many applications that benefit minority language use in cyberspace. One such use is by the Wikimedia Foundation. New language editions of the Wikimedia projects are created based on the language having an ISO 639-1, 639-2, or 639-3 code. SIL is working with the Wikimedia Foundation to ensure that individual language communities can benefit from this global resource.

Further information, including details on the change request process and documentation of requests, is available at the official ISO 639-3 website at <http://www.sil.org/iso639-3/>. For more information on Wikimedia sites, please visit http://meta.wikimedia.org/wiki/Meta:Language_proposal_policy.

Endangered Languages and Documentary Linguistics

A grave concern for many in the world is the large number of minority languages that are endangered or at risk of extinction. SIL shares this concern and is partnering with many organizations to determine ways to document and share information about these languages before they become extinct, with the prospect that these languages can be preserved. We believe that every language has inherent value, and that speakers of minority languages and other interested parties should have the tools and techniques available to protect and enhance their cultural and linguistic heritage. In many cases around the world, we have witnessed renewed language vitality as members of a language community are equipped with the capacity to use their language in new and different areas of life.

SIL continues to develop, and make freely available, various software applications that facilitate language documentation and sharing. One example is the FieldWorks suite of language tools. The FieldWorks Language Explorer (FLEx) component provides a comprehensive application that supports gathering texts in a language, analyzing the grammatical structure through interlinear text analysis, and developing a comprehensive lexicon. FLEx supports complex scripts and writing systems, and the user interface is localizable into different languages. Currently, localized versions are available for the following languages at different levels of completion: Chinese, English, French, Hindi, Indonesian, Korean, Persian, Portuguese, Russian (81% complete), Spanish and Turkish.

In addition, we are now working on solutions to publish dictionaries from FLEx on the web for further collaboration and knowledge sharing. With the advances in Unicode and web typography, complex scripts can now be utilized on the web, thus enhancing the language situation for minority languages.

We have also developed software that helps non-linguists build a dictionary in their own language. WeSay has various ways to help indigenous speakers to think of words in their language and enter some basic data about them. The program is customizable and task-oriented, giving the advisor the ability to turn on or off tasks as needed and as the user receives training for those tasks. WeSay uses a standard XML format, so data can be exchanged with linguist-oriented tools like FieldWorks. Community-level collaborative dictionary development can be a rallying point for language communities. It's something they can get involved in and it opens their eyes to a brighter future for their language.

For further information on the FieldWorks suite of software applications, and to download the latest version, please visit the website at <http://fieldworks.sil.org/>. For an example of a multi-script lexicon with more than 6,600 entries created using FLEx, see the Nuosu Yi–Chinese–English glossary at <http://www.yihanyingcihui.net/?lang=en>. To learn more about WeSay and download a copy, visit <http://www.wesay.org/>.

Mother Language-Based Multilingual Education

Governments around the world are increasingly aware of the difficulties that children face if the language of instruction at school is not the language the children speak at home. As more education systems seek to use the mother tongue of the children in the early years, demand is increasing for local language documentation and educational resources such as schoolbooks and multilingual dictionaries. SIL is developing software to help create early stage reading materials.

Conclusion

SIL's long history in supporting language development activities with minority language communities around the world has given us the unique opportunity to develop and contribute technical expertise from the local to the international level. We are grateful for the privilege and are at the same time aware of the responsibility to share what we have learned.

The growth of the Internet and the increase in global opportunities for collaboration, information sharing, and standardization on behalf of all language communities provide significant benefits for enhancing multilingualism in cyberspace. SIL continues to seek to be a valuable partner to support the particular needs of minority language communities. We are grateful for the significant progress being made.

Daniel PRADO

*Executive Secretary, MAAYA Network;
Director, Terminology and Language Industries, Latin Union
(Paris, France)*

Language in the Virtual World and in the Real World: Parallels to Take into Account in Language Policy

The Case of Romance Languages

It is often said that using a language for professional, administrative, educational, legal and other purposes helps it stay alive, because speakers who are forced to switch language according to context tend gradually to use the language that allows them the widest variety of expression. In our knowledge society, a language loses value in the eyes of its speakers if they cannot find knowledge or access to the rest of the world through it. As we said here, in Yakutsk, in 2008, with communication playing a growing role in the balance of power between two competing languages, in the information age this phenomenon favours the languages that are the best equipped or the most “prestigious” to the detriment of the others.

We know that the day is not far off when all, or at least the great majority of humanity will have access to cyberspace. In this context, if a language is absent from cyberspace, its speakers might, eventually, turn to other languages.

There is a high risk of the disappearance of more than nine out of ten languages which are not represented in cyberspace, because their speakers will have to use other languages for information, education, making purchases, administrative procedures, offering services, connecting to the rest of the world and so on. Furthermore, of the minority of languages that do have access, that is, between 300 and 500 according to different estimates, very few are really well equipped and have a relevant presence on the Internet.

Despite some clear progress in multilingualism in recent years, only a handful of the world’s languages have a noteworthy presence on the Web. English is still the language most in use on the Internet, but, as all serious studies show,⁸ its relative presence is falling. Corbeil told us back in 2000 that “very soon the presence of English should fall to about 40% when sites are created in different countries as they connect to the network”⁹, although we do not have scientific confirmation of this data.

⁸ *The full study Langues et cultures sur la Toile is available online: <http://dtiil.unilat.org/LI/2007/index_fr.htm>.*

⁹ *Corbeil Jean-Claude, «I comme informatique, industries de la langue et Internet », in B. Cerquigliny, Tu parles!?Le français dans tous ses états, Paris, Flammarion, 2000, p. 129.*

We should recall that most of the world's languages are represented in an essentially symbolic way with a few pages dedicated to them, and only a minority of languages are genuinely present.

Given that Facebook, Google and Wikipedia are strong trends on the Internet, it is by no means trivial to note that the famous social network has menus in less than 1% of the world's languages, that the almost global search engine only provides at present language recognition for about 50 languages, and that only 5% of languages are represented in the famous encyclopaedia, which nevertheless seems to be the virtual location that is the most open to languages. In fact, there is at present no system of language recognition for any language of American or Oceanian origin on Google¹⁰.

We said it in 2008, but it is worth repeating, that the phenomenon of the disappearance of languages, caused by several factors in the near past (colonization, genocide, epidemics, war, displacement of populations, prohibition on the use of the language, etc.), has grown as globalization has gathered pace, with its technological, political, and socio-economic consequences, in particular migration and urbanization.

As a general rule, languages that used to play an important role in the past have experienced a significant decline in sectors linked to knowledge, to the benefit of English. All languages of European origin, apart from some minority or sidelined languages that have been able to make a come-back in recent years, have been affected: German, French, Spanish, Russian, Italian, Portuguese, and so on.

What is the Situation of Romance Languages?

Among these European languages, the Neo-Latin languages are also affected. Despite the fact that the main Romance languages played a major role for nearly 1,000 years, in particular French, Spanish, Portuguese and Italian (without forgetting the historic, and literary and lexical mark made by Occitan, Catalan, Norman, Venetian and many other Romance languages), they are today diminished, and the action of languages as vectors of knowledge and international negotiation is reduced.

Of course, we understand that concerns about the ground lost by the Romance languages might seem misplaced when 99% of the world's languages have an uncertain future. However, the Romance languages are losing ground in international organizations, scientific and technical expression, international governance, higher education and international negotiations. They are

¹⁰ *It should be noted that we refer to "recognition" when the engine can search in a language and find results. Google might offer interfaces in a given language –in 120 to date – but that does not mean that the engine recognizes the language.*

certainly gaining in terms of demography and education as second languages (particular Spanish, Portuguese and French), but their use is above all related to tourism, culture, migration and owing to new populations becoming literate, and less and less to the sectors reserved for knowledge and negotiation.

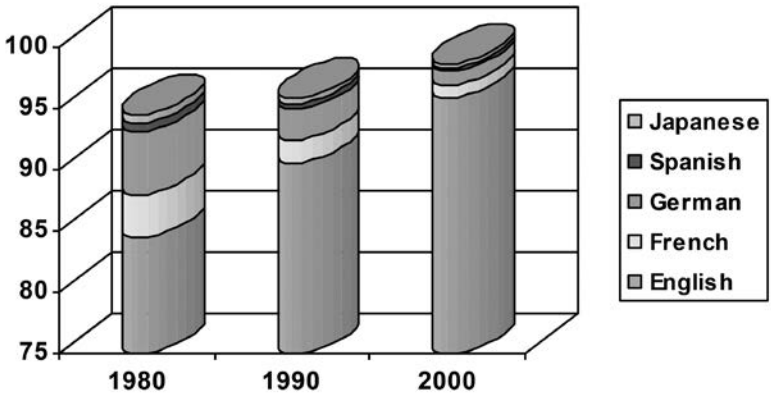


Fig.6 – Evolution of the language of publication of science books in the Science Citation Index¹¹

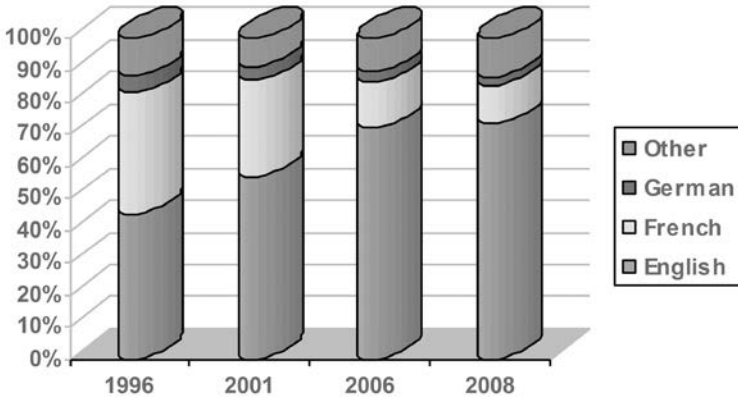


Fig.7 – Original language of European Commission documents (source: European Commission, Translation Directorate, 2009)

¹¹ Bordon María, Gómez Isabel, "Towards a single language in science? A Spanish view" in *Serials*, vol. 17, No.2, July 2004, pp. 189-195.

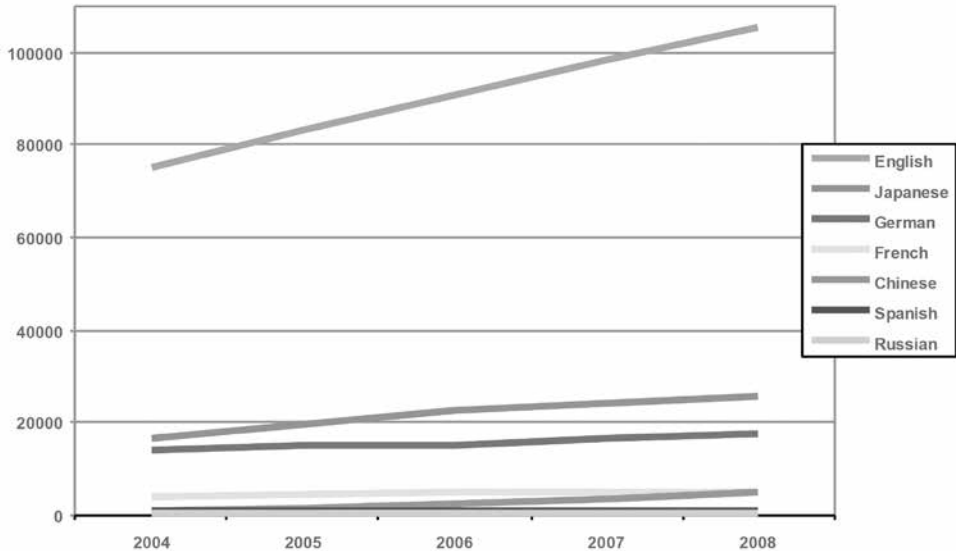


Fig.8 – Language of publication of WIPO patents between 2004 and 2008¹²

The outlook is not disastrous, as for a few years now speakers of French, Italian, Spanish and Portuguese have been working hard in many international bodies either to give their language specific status or to ensure that it is used as provided for in the rules and regulations. However all these languages are in decline in science and higher education, with our own research institutes preferring to publish their discoveries directly in English without providing a translation into the national language, and our higher education institutions proposing more and more courses in English only. The risk of “domain loss”, well known in Nordic countries, leading to the disappearance of whole segments of the language and meaning that engineers will no longer be able to talk among themselves other than in English, is on the way to becoming a reality.

Lack of Indicators

We lack an accurate vision of the situation of languages that could help us propose actions for a readjustment in favour of linguistic diversity. The statistics are incomplete, indicators are skewed, studies are biased and above all studies to measure languages are in their infancy.

¹² *PCT Annual Review: The international patent system in 2008*, WIPO, 2009, p. 20 of French version, [online]. <http://www.wipo.int/export/sites/www/pct/en/activity/pct_2008.pdf> (page consulted on 12 April 2010).

This leads us to certain questions: what makes a language attractive and enables it to act as a common language, a language of negotiation, education? In a word, what gives it prestige in relation to others? It is certainly not merely a matter of demographics, as otherwise Hindi would be as well regarded as Spanish or English, and Bengali as much as Portuguese or French and far more than German. Neither is it official status in a large number of States, as otherwise Portuguese would be more valued than Chinese, Russian or German. And it is definitely not the supposed qualities of a language to better express a situation more easily or more directly – a weak argument that is nevertheless heavily relied on by activists in favour of a single language of communication.

No, it is a series of characteristics in which demographic weight is important, as is official status, and whether the language is easy to learn. But it also includes other parameters such as history (we should not forget that the official languages of the United Nations are the languages of the winners of the Second World War), politics, economics, tourism, science and technology, standard of living, literacy, cultural industries, migration and so on and so forth, and today, without a shadow of a doubt, presence on the Internet.

Of course, these parameters are not sufficient without political determination behind them. Churchill paved the way followed by the numerous and powerful political and economic interests of the English-speaking world when he said that the widespread use of English would be “a gain to us far more durable and fruitful than the annexation of great provinces”. As a result, the English language is the main source of income for the United Kingdom, and Grin has reminded us that it thereby saves between 10 and 17 billion euros a year in translation costs owing the predominance of English in the drafting of European Union documents.

Political will has also enabled societies such as those in Quebec and Catalonia to recover a professional, institutional and educational use for their languages, and has even enabled the restoration to life of dead languages like Hebrew.

The Real World and the Virtual World

Access policies for all languages are needed, but especially what is needed is for their speakers to feel ownership. In order to formulate a policy though, reliable indicators are required as well as substantial written or audiovisual material. From the outset we have lacked any kind of indicator on languages and we still lack such indicators on the presence of languages in cyberspace at present, but when we do have some, for written languages in particular, we see that the virtual world seems to reflect the dynamic of languages in the real world.

In 2008, we compared the first 30 languages to have a language recognition system in Google, and we noted that they were, roughly speaking, the most productive languages in terms of traditional literature. Recent public and private initiatives to digitize library holdings might only serve to reaffirm the status quo of linguistic diversity on the Web.

Should we then conclude that the Web can only be added to when publishing comes first? Probably not really, because of the 50-odd languages with a language recognition system under Google, four fifths of them are the most productive of literature and translations and the remaining dozen might have far lower productivity but do have a larger quantity of articles on Wikipedia and are well represented on Facebook.

As far as we know, there is no global study giving us an oversight of the place of languages at the world level on social networks. We have however noted, through an increase in the number of specialized studies and an accumulation of various statistics, that written production through these means is far higher than the production of web pages, even if it is often ephemeral¹³. The studies carried out by Semiocast¹⁴ in 2010 for Twitter, for example, showed that Malay and Portuguese were used far more than Spanish, German, Russian and Italian, for instance, with a greater presence on the traditional Web and far more robust policies on the translation and digitization of works. The research has not been repeated since, but the languages spoken in Indonesia might be far more present today as it is the country with the third highest Generation 140¹⁵ in the world.

Do social networks represent a second chance for languages? Probably, because cyberspace actually opens the door to forms of expression of no interest to traditional publishing circuits. After all, science publishing in languages other than English has found a place, albeit a modest one, thanks to the ease and low cost of publishing on the Web, and traditional publishers do not want to run the risk of publishing articles that would be of concern to a very small number of readers.

The Internet has undoubtedly enabled minorities absent from traditional publishing to express themselves, but we should not think that this is enough. There is still an inversely proportional relation between Internet access and global linguistic diversity as we showed in 2008; the language divide corresponds only too closely at the moment to the digital divide.

That is why it is important to stress aspects relating to infrastructure as much as the ownership of technologies and content (text and multimedia) production.

¹³ *The observatory site Portalingua* <<http://www.portalingua.info>>, created by the Latin Union aims to respond to this problem by compiling and placing in parallel studies and statistics on the presence of languages in the various spheres of the knowledge society.

¹⁴ See: <http://semiocast.com/static/downloads/Semiocast_Half_of_messages_on_Twitter_are_not_in_English_20100224_fr.pdf> (page consulted on 9 July 2011).

¹⁵ A name for the Twitter-using generation, as the system only admits messages of 140 characters or less.

Promoting the use of a language must take place at every level: educational, administrative, scientific and technical, even for leisure, and in the regional or national bodies concerned. It must take place basically at the level of the access of languages to technology.

Machine Translation

Of all the language technologies, those whose evolution has been followed most closely are those related to machine translation. As language is perceived to be what sets the human species apart, the thought that a machine might replace us – or even go further than us – in the major linguistic and cognitive exercise that is translation cannot fail to awaken deep-seated fears. And yet, after many years of failure, we are apparently not so far off the goal ... at least, for a few pairs of privileged languages and still in the field of specialized translation which, we must recall, concerns between eight and nine of every ten pages translated in the world.

If new public programmes do emerge with the aim of democratizing the use of machine translation and favouring language pairs that have not been studied in any depth until now, it is also at the root of competition between businesses. Thus, the United States administration sees it more as a way of making businesses more competitive¹⁶, on the grounds that on average 52 % of consumers would not buy a product not described in their own language, according to a study carried out in 2006 in eight developed countries.

However, although everything seems to indicate that machine translation will be de facto integrated into all our applications, and although the quality seems to be quite satisfactory for some language pairs and the outlook broader in terms of the languages concerned, when it comes to most of the pairs treated, quality is lacking, and in any case concerns barely 60 of the world's languages. The geo-linguistic imbalance is clearly visible in machine translation. Although it is reaching maturity, although it is effective, although it is profitable, it only concerns very few languages, mostly used in North America and Europe, and to a lesser extent China and Japan, and the rest of the world is disregarded.

As the most effective systems derive their performance from the corpora already existing (thanks to “translation memories” and systems that process statistics) we can see the challenges that persist for languages with small written corpora.

To return to the Romance languages, although they are suitably equipped¹⁷, developments are still needed to give full satisfaction both to machine

¹⁶ See the final paragraph of the document presenting the White House's Innovation Strategy: <http://www.whitehouse.gov/assets/documents/SEPT_20__Innovation_Whitepaper_FINAL.pdf>.

¹⁷ Namely Catalan, Spanish, French, Italian, Portuguese and Romanian. Galician is finding it hard to find satisfaction despite various initiatives and the other Romance languages, for different reasons, are far from being considered suitably equipped to be used in contexts of specialization.

translators and to human translators working with these languages. For instance, in addition to overdue spelling reform for Portuguese, and a lack of public resources to automatize Italian, Portuguese and Romanian¹⁸ other than those provided by the European Commission, there is a blatant lack of terminology policies for all these languages, other than French and Catalan. This latter situation is moreover the reason civil society initiatives have been launched, notably that of the Pan-Latin Terminology Network (Realiter)¹⁹, which brings together the main actors in the terminology of seven Romance languages, and which despite some remarkable work, is, clearly, far from being able to meet the needs of all these languages.

The Three Linguistic Spaces²⁰ are preparing an interoperability project for Spanish, French and Portuguese terminological data banks, which might give strong impetus to the terminological vitality of the three languages, but the road ahead is long. Let us hope that a similar project on language technologies for all the Romance languages can follow on from it.

We are striving today to give effect to Action Line C8 of the World Summit on the Information Society Action Plan “Cultural diversity and identity, linguistic diversity and local content”, to recommendations of international meetings. Documents that came out of the MAAYA seminars in Bamako, Havana and Barcelona, and those that came out of the 2008 Conference in Yakutsk should be applied for equitable representativeness of languages and cultures in cyberspace²¹. The World Network for Linguistic Diversity MAAYA is well on the way to this goal, with constant support from the Latin Union. Today subject to the policy of preservation, promotion and modernization should be not only the Romance languages descended from Latin, but also all those with which they share spaces and concerns for the future.

The Latin Union is convinced that supporting MAAYA ideas is essential, in particular the idea of holding a summit on linguistic diversity and multilingualism, because we are sure that the political will exists today more than ever before to renew most of the world’s languages.

¹⁸ *Despite the sustained activity of two institutes of the Romanian Academy, the linguistics institute and the artificial intelligence institute.*

¹⁹ *See <<http://www.realiter.net>>.*

²⁰ *Group consisting of five intergovernmental organizations wishing to act in favour of French, Spanish and Portuguese languages and cultures: Community of Portuguese-Speaking Countries (CPLP), International Organisation of La Francophonie (OIF), Organization of Ibero-American States (OEI), Ibero-American General Secretariat (SEGIB) and the Latin Union.*

²¹ *Read the full text of the Action Line in the UNESCO list of actions for WSIS: <http://portal.unesco.org/ci/en/ev.php-URL_ID=15927&URL_DO=DO_TOPIC&URL_SECTION=201.html>.*

Sozinho F. MATSINHE

*Executive Secretary, African Academy of Languages
(Bamako, Mali)*

African Languages in Cyberspace: Challenges and Prospects

*One thumb alone cannot crush a louse
African proverb*

Introduction

The advent of Cyberspace marked the dawn of a new era in the development of human languages. More and more human languages are gaining space and research is deepening. More cooperation amongst scholars across the globe is strengthening than ever before and more networks amongst them are also established. In short, Cyberspace has lent a fresh impetus to the development of human languages. However, many factors, including the lack of proper language policies; regulations, and business inspired interests have effectively excluded African Languages from Cyberspace. The Mozambican daily newspaper, *Noticias*, in its issue of 2 October 2012, while citing a report on the world stage of the use of broad band and digital inclusion launched in New York recently, points out that the vast majority of Africans have no access to the internet. Similarly, Prado (2012:34), in his paper entitled Language Presence in the Real World and Cyberspace observes that²²:

Barely 5% of the world's languages have a presence in cyberspace, and among these few, there are still considerable differences. Only a tiny handful of privileged languages offer a genuine production of content.

This paper takes a holistic approach to the language issue in cyberspace. Specifically, on the one hand, it calls for fair language and regulatory policies that take into account the African Linguistic Mosaic and, on the other, it argues that ACALAN, as the sole language agency of the African Union entrusted with the task of fostering the development of African languages in collaboration with the member states, should play a pivotal role in the process of allocating proper space to African languages in cyberspace.

²² Prado, Daniel (2012). *Language Presence in the Real World and Cyberspace*. In *NET.LANG: Towards the Multilingual Cyberspace*. MAAYA NETWORK: C&F Éditions: 35-51

The African Linguistic Mosaic



The lighter area of the diagram represents mainly the urban areas dominated by the former colonial languages, i.e. English, French, Portuguese and Spanish. These languages, though spoken by small minority elite, were retained as official languages, when African countries achieved their independencies in the early sixties. As a result, they are associated with power, access to socio-economic privileges, including education, justice and well-remunerated jobs. In other words, they dominate the socio-economic mainstream. As a result, these former colonial languages are regarded as passports for upward social mobility and, as such, given preference at the expenses of African languages, which occupy the darker area of the diagram, representing the rural areas where the vast majority of Africans live and communicate solely in these languages. As Negash²³ points out, while discussing globalization and the role African languages can play in the development of Africa,

African governments and the elite still continue to channel away their resources and energies into learning 'imperial' languages that are used by a tiny minority of the population.

The preference for former colonial languages has not only resulted in maintaining the status quo, but has also led to the exclusion and marginalisation of the vast majority of Africans keeping them on the periphery of the socio-

²³ Negash, Ghirmai (2005). *Globalization and the Role of African Languages for Development*. Paper presented at the conference "Language Communities or Cultural Empires", February 9-11, 2005, University of California at Berkeley.

economic mainstream. In this regard, while discussing language, dominance and control in Africa, Wolff²⁴ expresses a similar view, when he states that:

Post-independence language policies in Africa, however, were largely in favour of maintaining or installing the colonial foreign language as official language of the newly-independent country. This worked much to the benefit of the former colonial masters: all official dealings with the new governments could be conducted in the language of colonialism. Likewise, the new governments saw no other way of smoothly taking over control and power from the colonial master than insisting on official monolingualism in their favour. Any change of language policies in favour of national languages to complement, if not replace, the foreign colonial languages as official languages on the national level in the long run, would put them at a disadvantage in terms of communication and control. The ex-colonial foreign language, therefore, comes in handy for “mass exclusion” from control and access to power and resources.

The language policies described here have also created a fertile ground for the thriving of negative attitudes towards African languages. As a result, there have been a widespread believe that African languages are linguistically ill-equipped and cannot be used in techno-scientific domains, including the cyberspace. On the false ill-preparedness of African languages, Negash has this to say²⁵:

Some say that African languages may be effective for daily interaction but not for coping with the demand of modern, high technology, science, the arts, literature, cinema, the internet, international communication of diplomacy and trade, etc., all aspects of a too complex contemporary world.

Unwittingly, Africans buy into this belief and regard linguistic diversity that is the hallmark of the African linguistic mosaic shown in the diagram presented above as rather being a liability than an asset, and therefore that the exclusive use of former colonial languages should prevail over African languages. This is why the former colonial languages are regarded as passports for upward mobility and, as such, they are given preference at the expenses of African languages. Tied to the beliefs, there is a plethora of factors militating against the development, promotion and use of African languages in all domains of society, including cyberspace.

²⁴ Wolff, Ekkehard (2011). *Background and history – language politics and planning in Africa*. In Ouane, Adama and Christine Glanz (2011) (Eds.) *Optimizing Learning, Education and Publishing in Africa: The Language Factor: A Review and Analysis of Theory and Practice in Mother Tongue and Bilingual Education in Sub-Saharan Africa*, Tunis: UIL/ADEA: 51-100.

²⁵ Negash (2005:11).

Factors Militating Against the Presence of African Languages in Cyberspace

There are many factors militating against the presence of African languages in cyberspace. Major of them include:

- The lack of political will to put in place effective language policies in Africa;
- The lack of proper regulations favouring African languages;
- The lack of human and financial resources;
- The lack of effective training programmes that are informed by the African linguistic mosaic referred to above;
- The lack of incentives, user-friendly as well as practical programmes offered in the institutions of higher learning;
- The work to foster the presence of African languages in cyberspace is generally inspired by business interests, including competition that neither leaves space for experience sharing and cross-fertilization of ideas nor properly takes into account the African linguistic mosaic; and
- The content of the African languages present in cyberspace tends to be confined to non-standardised translation engines and programmes.

African decision makers have not yet gone beyond making ambiguous statements on language policies to which very often they pay lip service. This may explain why most constitutions of African countries contain varied clauses on the status of African languages. The constitution of the newly independent South Sudan is the very epitome of what is stated here. In its Part 1, article 6, clauses 1 and 2 on languages²⁶ it states that:

(1) All indigenous languages of South Sudan are national languages and shall be respected, developed and promoted.

(2) English shall be the official working language in the Republic of South Sudan, as well as the language of instruction at all levels of education.

As is well known there is no better way to foster the development of a language than using it as a medium of instruction. The exclusion of African languages from the education system makes it difficult to pass regulations in favour of developing African languages and to accord them a proper place in cyberspace. Returning to the training aspect, apart from lacking lustre, most programmes on computational linguistics institutions of higher learning offer across Africa do not

²⁶ *The Transitional Constitution of the Republic of South Sudan 2011, Government of South Sudan, Page 13.*

take into account the African linguistic mosaic and they are mostly concerned with theoretical issues with little impact on the development of African languages. Furthermore, work directly linked to African languages in cyberspace is generally informed by business imperatives and, as such, is broadly limited to developing spell checkers and online dictionaries while taking into account the number of speakers of the targeted language. Last but not least, is the lack of resources. The lack of funds to support research and training is another challenge facing the presence of African languages in cyberspace. This is exacerbated by the absence of clear language policies and regulations, as described above.

Language is a cross-cutting issue and, as such, there is no single solution to the problems facing the presence of African languages in cyberspace. We therefore require a paradigm shift that will not only take into account the African linguistic mosaic, but will also allow a holistic approach, which takes into account various initiatives and resolutions taken at various forums.

Paradigm Shift

As the African proverb cited above reminds us “one thumb alone cannot crush a louse”. Researchers on matters related to African languages and cyberspace tend to work on silos. It is therefore necessary to broaden the research agenda on Human Language Technology, while strengthening macro and micro cooperation across various disciplines and stakeholders as well as creating synergies. The African Academy of Languages (ACALAN) can play a vital role, particularly in lobbying for the support of African decision makers, urging them to put in place effective language policies that favour the development of African languages. It can also facilitate the mobilization of resources at national and international levels. As we all know ACALAN is the official language agency of the African Union whose statutes were approved during the Summit of the Heads of State and Governments that took place in Khartoum in 2006. Its mandate is to work in collaboration with the member states of the African Union towards the development, promotion of African languages so that they are used in all domains of the society in partnership with the languages inherited from Africa’s colonial past; i.e. English, French, Portuguese and Spanish. The partnership component is vital as the minority elite referred to earlier always see the efforts to develop African languages as an attempt to replace former colonial languages with African languages. Far from it, all that is required is a linguistic equity in the same way Africa has been calling for gender equity. However, for all that to materialize, the quest for creating space for African languages in cyberspace should be part and parcel of the quest for developing African languages in particular and for poverty eradication in Africa in general.

Taking into account the work on the harmonization of the orthographic systems of the Cinyanja/Chichewa, Fulfulde, Hausa, Mandenkan, Kiswahili, Setswana, Vehicular Cross-Border languages, ACALAN organized a workshop on African languages and cyberspace in Niamey from 14 to 15 in December 2011. The workshop brought together researchers working on African languages and cyberspace from Botswana, Djibouti, Kenya, and Nigeria. It took stock of current work on African languages and cyberspace with special reference to the vehicular cross-border languages mentioned here. One of the recommendations of the workshop was that ACALAN should create space on its website where researchers on African languages and cyberspace could post information on current work. This would allow ACALAN to consolidate the information and clearly define priority areas on which to focus its activities. As of the outcomes of the workshop, ACALAN has commissioned the development of spellcheckers to some of the researchers who participated in the workshop. Apart from contributing towards according equitable space to African languages in cyberspace, the spellchecker will allow to disseminate the harmonized orthographic systems of the vehicular cross-border languages mentioned above.

As mentioned above, it is also necessary to take into account the various decisions, plans of actions and resolutions which directly and indirectly have bearing on the efforts to develop, promote and use African languages in all spheres of society, in particular:

- The Language Plan of Action for Africa;
- The Second Decade of Education for Africa;
- The Charter for African Cultural Renaissance;
- The Khartoum decision on the linkage between education and culture.

The use of African languages has been one of the major preoccupations of the Organization of African Unity (now African Union) since its creation in 1963 in Addis Ababa, as indicated in Article XXIX of its founding charter stating that “The working languages of the organization and all its institutions shall be, if possible, African languages”. To render operational the resolutions on languages, the Heads of State and Governments adopted resolution AH/DEC8 establishing the OAU Inter-Bureau (OAU/BIL) during the 1966 summit held in Addis Ababa. This institution started its functions in 1972 in Kampala under the directorship of the late Kahombo Mateene. In 1985, the OAU/BIL convened a meeting of prominent linguists from Africa and beyond to draft the Language Plan of Action for Africa. It was subsequently adopted by the heads of state and governments through resolution C11/1123 (XLVI)

during the Addis Ababa summit in 1987. It is to be noted here, that OAU/BIL is in fact the precursor of ACALAN's work and approach to the development of African languages, particularly the focus on vehicular cross-border languages as a pragmatic strategy for Africa's development and integration.

The Language Plan of Action for Africa defines priorities and a programme of action the African Union Member States should undertake in their efforts to develop, promote and use African languages in all domains of society. Its main objectives are as follows:²⁷

- To encourage each and every Member State to have a clearly defined language policy;
- To ensure that all languages within the boundaries of Member States are recognised and accepted as a source of mutual enrichment;
- To liberate the African peoples from undue reliance on the utilisation of non-indigenous languages as the dominant, official languages of the state in favour of the gradual take-over of appropriate and carefully selected indigenous African languages in this domain;
- To ensure that African languages, by appropriate legal provision and practical promotion, assume their rightful role as the means of official communication in the public affairs of each Member State, in replacement of European languages, which have hitherto played this role;
- To encourage the increased use of African languages as vehicles of instruction at all educational levels;
- To ensure that all the sectors of the political and socio-economic systems of each Member State is mobilised in such a manner that they play their due part in ensuring that the African language(s) prescribed as official language(s) assume their intended role in the shortest time possible;
- To foster and promote national, regional and continental linguistic unity in Africa, in the context of the multilingualism prevailing in most African countries.

The Language Plan of Action for Africa addresses some of the challenges described above, which need to be addressed so that African languages attain a significant position in cyberspace. First and foremost, the need for the member states to define clear language policies. There is also a list of

²⁷ *The Language Plan of Action for Africa, Council of Ministers, Forth –Sixth Ordinary Session, 20-25 July 1987, Res. 1123 (XLVI), Organization of the African Unity, General Secretariat, Addis Ababa, Pages 2-3.*

priorities in the Language Plan of Action for Africa, including modernisation that has bearing on the research whose output will pave the way for African languages to enter cyberspace.

The draft plan of action for the Second Decade of Education for Africa (2006 to 2015)²⁸ produced by the Department of Human Resources, Science and Technology of the African Union in June 2006 specifically states that:

Language is another area of concern in African education systems. The preponderant use of 'foreign' languages as media for instruction disadvantages learners and erects barriers between school and community. It has been shown that learning outcomes are higher for children who learn in their vernacular in the early years of school.

As correctly stated here the use of African languages as medium of instruction can enhance learning and improve the learner's performance. This can put African languages to the test and allow them to consolidate and systematize terminology to express various scientific concepts. By doing so, addressing one of the main concerns mentioned above. For this to be possible, the Second Decade of Education for Africa calls for the revamping of the education systems that would allow the effective use of African languages as a medium of instruction. The efforts to place African languages in cyberspace should be part and parcel of that process.

The Charter for African Cultural Renaissance was adopted in January 2006 during the Summit of the Heads of State and Government that took place in Khartoum. In order to popularize the charter and urge the African Union Member States to ratify it, the African Union Commission launched the Campaign for African Cultural Renaissance (2010-2012). The launch of the campaign coincided with the celebration of the centenary of birth of Kwame Nkrumah, one of the founding fathers of the Organization of African Unity. In part IV, Article 18 of the charter for African Cultural Renaissance²⁹

African States recognize the need to develop African languages in order to ensure their cultural advancement, and acceleration of their economic and social development. To this end, they should endeavour to formulate and implement appropriate national language policies.

As was the case with the resolutions referred to above, the Charter for African Cultural Renaissance also underlines the role African languages play in propelling Africa to economic and social development and the need to adopt

²⁸ *Second Decade of Education for Africa (2006-2015), Draft Plan of Action for the Second Decade of Education for Africa, June 2006, Department of Human Resources, Science and Technology, African Union Commission, Page 11.*

²⁹ *Charter for African Cultural Renaissance, African Union Commission, January 2006, page 9.*

and implement policies conducive to developing these languages. Furthermore, Article 19 of the charter urge African member states to effect reforms to integrate African languages in their education systems. In order for the objectives of the charter related to African languages to be attained, African languages should be accorded a proper place in cyberspace.

Another decision taken during the Khartoum summit in 2006 relates to the linkage between education and culture. According to this decision, when African Union Member States undertake the reforms in their curriculum, within the ambit of the Second Decade of Education for Africa, they should ensure that the content of the curriculum is informed and inspired by African culture. ACALAN has been assigned the task of monitoring the process and regularly report to the African Union Commission.

Conclusion

A holistic approach to the efforts to put African languages in cyberspace is therefore required. This approach will not only take into account the various decisions and resolutions pertaining to the development, promotion and use of African languages, but will also broaden the research agenda on Human Language Technology, as it would be difficult for African languages to gain an equitable place in cyberspace without strengthening applied research in that domain. In other words, in order to address the various challenges facing the efforts to accord equitable space to African languages in cyberspace a collective effort is required. Such effort should go beyond the concerns of linguists, language practitioners, teachers of African languages and other stakeholders whose work involve African languages regularly. As is well known, language is not everything; but it is in everything. This implies that according equitable space to African languages in cyberspace should be part and parcel of the search for viable strategies to bring about sustainable development that would change the lives of the vast majority of Africans for the better. Once again, we can remind ourselves of the African proverb cited above “one thumb alone cannot crush a louse.”

Gilvan Müller de OLIVEIRA

*Director, International Portuguese Language Institute
(Praia, Cape Verde)*

Margarita CORREIA

*Researcher, Institute for Theoretical and Computacional Linguistics
(Lisbon, Portugal)*

Portuguese and the CPLP Languages in Cyberspace

Introduction

The CPLP, an international organization founded in 2000, brings together all Portuguese-speaking countries, covering a territory of 10.7 million km² in America, Africa, Europe and Asia, and a population of about 241 million. The member countries are: Angola, Brazil, Cape Verde, Guinea Bissau, Mozambique, Portugal, Sao Tome and Principe and East Timor.

The degree of proficiency in Portuguese language in different countries varies from almost 100% in Portugal to less than 10% in Guinea Bissau and East Timor, as other 339 languages are also spoken in the CPLP, of which 215 languages in Brazil, which accounts for 5% of the number of languages in the world, set in roughly 6500.

Although the CPLP is the attempt of construction of an international parity and democratic block, the expansion of Portuguese language was due to the construction of a colonial empire in the same way as European commercial colonialism in the sixteenth and the seventeenth centuries. The Portuguese colonial empire was the last of this tradition to disappear, and this only occurred with the independence of African countries and East Timor in 1975. The process of independence was achieved through armed struggle, called the Colonial War (1961-1975), a conflict that lasted in some countries, in new forms, till the late 1990's or even the early 2000's.

The colonial situation left two main by-products when it comes to languages.

On the one hand there was the impossibility of building the modern concept of citizenship and the consequent lack of interest in schooling of the population called "native", with a low participation in the Portuguese language community, low penetration of the Portuguese language, and very low level of literacy - the monarchical Brazil becomes Republic in 1889 with 98% of illiterates, a figure similar to that of Mozambique at the time of independence in 1975.

On the other hand, colonial language policies or those already independent States, as in the case of Brazil, led to the exclusion of the other languages of virtually every prestigious areas of circulation, which meant that, at best, its use continued in oral language, out of institutions without building standards or instruments related to writing.

The Portuguese in Cyberspace

The era of digitization of the languages found the world of Portuguese languages unprepared for the challenges of the Millennium Goals and the conformation of the Information Society in relation to the *corpus* of the languages, but also with regard to *Internet access* and to the necessary schooling of populations to participate in virtual communities.

Nevertheless, the Portuguese is the fifth most used language on the Internet, with 87 million users, a figure growing rapidly, following the rapid growth of school population in the last 15 years and the improvement of logistics for the provision of access, as the supply of electricity.

The Portuguese on the Internet is currently treated as two languages: *Portuguese* and *Brazilian Portuguese*, which results from the fact that we are facing a pluricentric language (rather bicentric), presenting two very different standards (Portuguese and Brazilian), particularly with regard to spelling: Portugal followed for decades the established spelling Orthographic Agreement of 1945, while Brazil was ruled by the Orthographic Form of 1943.

This treatment of the Portuguese on the Internet as if it were two different languages weakens its position and international exposure, so it would be desirable that this situation should be reversed.

The Orthographic Agreement of the Portuguese language in 1990, which was ratified by six of the CPLP countries, which is already in force in Brazil and Portugal (although in this country the transition phase between the two spellings is still on going until 2015) opens the prospect that Portuguese might come to be treated in a unified way on the Internet, as in this context, the spelling issues are very important to the weight they have in the construction of computational tools of production and recovery of contents.

One of the aspects relevant to evaluate the weight of an international language and its vitality is its ability to respond to the challenges posed by science and technology in terms of scientific output in that language.

As such, the presence of scientific texts in Portuguese is one of the parameters to consider when trying to portray the presence of this language on the Internet.

In this particular aspect the Portuguese position on the Internet is clearly different from the two centers which so far have determined the bicentric character of this language: Portugal and Brazil. When comparing the scientific productions in Portuguese language available on the Internet, there is a clear imbalance between the two standards: the presence of scientific texts in Brazilian Portuguese on the internet is many times superior to the presence of Portuguese texts.

To explain this, some data have to be taken into account, which go beyond the geographical and demographic differences between the two countries: Portugal has a total area of 92,389 km² and about 10 million inhabitants, while Brazil has 8,514,876 km² and approximately 191 million inhabitants.

Portugal is a European country, which is part of the European Union since 1986. As such, since then the parameters of evaluation and funding of science and technology in this country have been marked by certain standards by the European Commission, who value scientific publication in English, at the expense of publication in Portuguese.

Furthermore, in Portugal, given its size, geographic location, emigration (which has always been a constant) and bet on tourism since the 60's, the teaching of foreign languages not only has been encouraged but also has had very positive results: the literate population can express in one or two foreign languages, often acquired outside the formal school system. Currently English is the language preferred by young students and with more support at the level of government structures. It is to be noted that the implementation of compulsory teaching of English in schools from the third grade occurred in 2005 (Order 14753 / 2005 by the Minister of Education), to the detriment of other languages so far studied in the education system, French and German.

All these factors contribute to the fact that a lot of scientific literature in Portugal, especially in the areas of the hard sciences and technologies, is produced primarily in English: this case is all the more visible when even in areas in which Portugal was once a pioneer, as the nautical and ship construction, the Portuguese language has ceased to be practically used and has been replaced by English in a professional context.

In turn, the Brazilian government has not adopted policies of scientific literature in the English language in the same way, valuing, also, the scientific literature in Portuguese at the level of science funding agencies. On the other hand, Brazil, an emerging country, has structural problems at the level of basic education that are reflected in teaching and learning of foreign languages, very deficient when compared with the case of Portugal.

If we add to this, economic power and the number of universities and research centers in Brazil, as well as the impact of areas in which Brazil is currently a leading producer of science and technology (note for example the case of biofuels), it is easily understood that the Brazilian scientific production in the Portuguese language is truly thriving.

All these data have a direct impact on the amount of scientific texts (theses, reports, scientific articles) available on the Internet in Portuguese, and it is mostly ensured by Brazil – and in some subjects almost exclusively.

Added to this framework, policies such as the universal availability of master's dissertations and doctoral theses of all Brazilian universities on the Internet, existing since 2000, which enhance the circulation of knowledge in Portuguese. Furthermore, we call attention to the SciELO Network, which indexes the scientific literature in Portuguese and Spanish, creating a broad scientific area in two very close languages, and which together are spoken by 580 million people in 30 countries.

This last argument shows the opportunity for Portuguese-Spanish bilingualism, already adopted as communication policy of the MERCOSUR (Southern Common Market) in South America, and is seen increasingly as a strategy to promote the use of these two major languages as an alternative to the exclusive use of English.

The Other Languages of the CPLP

Within CPLP there are between 300 and 340 spoken languages, according to the way of counting, 215 languages being spoken in Brazil, including indigenous languages, languages of immigration, sign languages, Creole and Afro-Brazilian languages.

The twenty-first century presents a more purposeful framework for the presence and promotion of minority languages in public compared with previous centuries, when this set of languages was ignored by the public power, or at various times, was largely suppressed by the colonial Portuguese or Brazilian power .

The process of affirmation of linguistic diversity is very recent throughout the Portuguese speaking states, but practically all are moving to create new laws and practices in this field. Thus, we find cases of:

- **Officialization of minority languages** in Timor, Portugal and Brazil;
- **Actions of heritage language valorization** in Angola, Brazil, Cape Verde, Portugal, Timor and Sao Tome and Principe;

-
- **Bilingual education programmes involving minority languages**, although experimentally, in Brazil, Portugal, Timor and Mozambique;
 - **Corpus Development actions** in Cape Verde, East Timor, Angola, Sao Tome and Principe, Mozambique and Brazil;
 - **Inclusion Actions of minority languages in Internet instruments** in Cape Verde, Brazil and East Timor.

The very initial degree of preparation of corpus of the CPLP languages (scripturalization, standardization and regulation), as well as the incipient literacy of speakers in their own mother tongue, given the exclusion of these languages in most education systems, have been an impediment to further their presence in cyberspace. However, it is expected that soon changes will be experienced in this field.

We can cite the example of Nheengatu in the Brazilian Amazon: a language that until the early twentieth century was spoken in much of the 4 million square kilometers of this territory, and now is spoken in an area of about 35,000 km² by no more than 7 000 people. Nevertheless, it was favored by the legislation emanated from the 1988 Federal Constitution, so it was possible to initiate in 1997 the Intercultural Bilingual Schools Programme and a teacher training programme. It was made official at the municipal level in São Gabriel da Cachoeira, along with the Tucano and Baniwa, by ordinary law of the City Council in 2002, in a process unprecedented in Brazil, and the law was legislated in 2006. It has been a year since the first Indigenous Language Degree, offered by the Federal University of Amazonas (UFAM) was established in the country; this programme selected 40 students of Nheengatu. Currently, the students are busy in the Cucuí community, Alto Rio Negro, on the border between Brazil and Venezuela, with the support of IILP, creating a Wikipedia in their language, virtual encyclopedia that will receive, in the form of entries, the knowledge researched and produced by them in the course. This path taken by Nheengatu can be followed soon by many other languages.

I would like to remind you that the official minority languages at the municipal level in Brazil has already reached the figure of seven languages made official (Nheengatu, Tucano, Baniwa, Pomerano, Talian, Hunsrickisch and Guarani) in eleven cities (São Gabriel da Cachoeira (AM), Santa Maria Jetibá (ES), Laranja da Terra (ES), Pancas (ES), Vila Pavão (ES), Domingos Martins (ES), Antonio Carlos (SC), Serafina Correia (RS), São Lourenço do Sul (RS), Paranhos (MS) and Tacuru (MS)).

In other words, this is a time of preparation of logistic conditions for the access of speakers and languages from CPLP into Cyberspace, and if the current trend continues we will have visibility of our major languages on the Internet very soon.

For this reason the International Institute for Portuguese Language will hold The Maputo Colloquium on the Linguistic Diversity of the CPLP, from 12th to 14th September of the current year, which will assemble for the first time, in Mozambique, programme managers in linguistic diversity of the eight member countries. On that occasion there will be an exchange of experiences on the modus operandi of the institutions responsible for language rights, bi or multilingual education, the promotion of Portuguese in complex sociolinguistic contexts and other aspects related to the field.

Also, and in a complementary way, the International Institute for Portuguese Language will hold at the end of January 2012, in the Brazilian state of Ceará, The Fortaleza Colloquium on the Portuguese Language in the Digital World and the Internet, in order to, also, establish contact between internet managers of the eight countries to think collectively about the future of our language in this medium. Both colloquia will have the opportunity to benefit from the knowledge conveyed here in this extraordinary conference in Yakutsk.

The Maputo and the Fortaleza Charter will contain experts and managers advice to the II International Conference on the Future of Portuguese in the world system, which will take place in Lisbon, Portugal, in 2012, and shall prepare the Lisbon Action Plan for the Promotion, Diffusion and Projection of the Portuguese Language (2012-2014); this plan needs to strongly consider the global movement for language rights, the ecology of knowledge and the building of the future by all our citizens, speakers of many languages in which solidarity in diversity should be built up within the Community of Portuguese Speaking Countries.

Javier LÓPEZ SÁNCHEZ

*Director General, National Institute of Indigenous Languages
(Mexico City, Mexico)*

Revitalization, Strengthening and Development of National Indigenous Languages of Mexico in Cyberspace

The 2nd International Conference *Linguistic and Cultural Diversity in Cyberspace* gives a unique opportunity to share information on some projects aimed at making a better use of the opportunities that cyberspace is offering to the global society in the 21st century. These are particularly important for us, who lead and manage public institutions responsible for the revitalization, strengthening and development of national languages, working for the recognition and diffusion of cultural and linguistic diversity in our countries and regions, as well as the elimination of social practices of exclusion and discrimination.

Although these issues have been ignored, I will discuss the importance of carrying on investment in infrastructure, equipment and human capital necessary to ensure the access of indigenous communities in cyberspace, promoting their incorporation with linguistic and cultural relevance to the intensive use of the Information and Communication Technologies (ICT).

I propose the implementation of inter-institutional agreements, for carrying out joint actions, gathering global resources and using ICT in the revitalization, strengthening and development of indigenous languages.

Cultural and Linguistic Diversity

Nowadays, biodiversity and cultural diversity are facts that globalization and the use of ICT have made us more evident. Thus, linguistic diversity is as important for the cultural world development as biodiversity is for the sustainability of the planet.

The mother tongue is an essential mechanism for our species to pass on knowledge and ways of seeing the world from generation to generation. According to the latest report from Ethnologue (2009) there are 6,909 languages worldwide. Official research data show that at least half of them are reducing the number of speakers and are at risk of disappearing (UNESCO, 2003); and 90% of these languages have no presence on the Internet (UNESCO; 2009).

In the American continent, there are over a thousand languages which are approximately 15% of the world's total. The National Institute of Indigenous Languages (INALI), established in Mexico in 2003, is a decentralized agency of

the federal government, dedicated to promote the strengthening, preservation, and development of 68 linguistic groupings and 364 variants, all organized into 11 linguistic families, according to the Catalog of the National Indigenous Languages (INALI; 2008).

According to the latest XIII Censo General de Población y Vivienda (2010) nearly 7 million representatives of 62 indigenous peoples residing in the country speak national indigenous languages. However, migratory movements have prompted the dispersal of speakers of indigenous languages (SIL) on the length and width of the country, without considering the significant presence in the United States of America.

It should be noted that the SIL live mainly in the States of Oaxaca, Chiapas, Veracruz and Yucatán, a million of them are still monolingual.

Náhuatl, Maya, Mixtec, Zapotec, Tzeltal and Tsotzil are linguistic groupings which “concentrate” 54% of SIL, although it should be mentioned that only these 6 groups add in total 184 linguistic variants, many of them as close as it can be Spanish with other Romance languages such as French, Italian or Portuguese.

Diversity and Linguistic Rights in Mexico

In an international level, Mexico has signed the *Universal Declaration of Human Rights* (1948), the *Convention 169* of the UN International Labour Organization ILO (1989), and the *Universal Declaration of the Rights of Indigenous Peoples* (2007) among other instruments that recognize rights to indigenous peoples.

At the same time, at federal, state and municipal levels institutional changes are still important, such as the reforms to Article 4 in 1992, and Article 2 in 2001, by which the Mexican State has recognized at a constitutional level that: “The nation has a pluricultural composition originally based on their indigenous peoples...”.

2003 witnessed the enactment of the Federal Law to Prevent and Eliminate Discrimination; the law of the National Commission for the Development of Indigenous Peoples (CDI) and the General Law of Linguistic Rights of Indigenous Peoples; as well as the creation of new agencies, responsible for ensuring their rights: CDI, INALI and National Council against Discrimination (CONAPRED).

It is up to the INALI to promote knowledge, recognition, evaluation and appreciation of the national indigenous languages by the approach to multilingualism and interculturalism. This institution has the following strategic lines of action:

1. To promote public policies for indigenous languages, encouraging the participation of all social agents (indigenous peoples, public and private institutions, universities, research centers, experts, etc.), usage of different means of communication and inter-institutional

linkages and coordination of efforts with the federation, states and municipalities.

2. To encourage the use of the national languages in governmental practice and daily life, and
3. To ensure linguistic planning at national level with focus on the formalization of the national indigenous languages (cataloguing) and the standardization of writing, grammars and dictionaries and specialized lexicons.

However, it should be noted that while the indigenous population is not able to seize the institutional framework that has been described, it is harder to reach more ambitious development goals. The only way out is organizing intensive diffusion campaigns to show that it is possible to break the cycle of poverty and discrimination, historically associated with the use and preservation of indigenous languages.

In addition, jointly defining objectives and working out clear strategies international agencies, developers, and users of cyberspace would be able to:

1. Establish leadership to promote awareness, respect and the strengthening of the global, regional and local cultural and linguistic diversity.
2. Reinforce support programmes to install infrastructure and equipment, as well as develop focused applications to revitalize and fortify linguistic diversity.

Cyberspace as a Window of Opportunity for Diversity

Today the options that cyberspace and public media provide are essential to make visible and to spread knowledge about the social and regional realities of minority groups, traditionally excluded and isolated.

The expansion of coverage and the accelerated growth of cyberspace and its applications, offer new opportunities to developers and users for revitalizing, strengthening and developing the cultural and linguistic diversity at all levels.

At the same time, cyberspace is an “ideal place” where respect, freedom and democracy prevail, representing a real and huge “window of opportunity”, allowing “the others” to be “visible and audible”, to communicate among themselves and to share their ideas and creations with others.

According to a recent United Nations report by Frank La Rue (UN Human Rights Council, 2011) Internet is a medium where the right to freedom of expression can be exercised; and that access should be included by the Member States as a human right to develop effective policies to achieve universal access.

While this is a fast access medium even from remote places, members of indigenous peoples and their communities immediately get engaged in online communications in a very active and creative way due to the similarity of such activities to traditional forms of participation in community work, where information and results are freely shared.

Since 2005 the INALI produces and promotes among the indigenous peoples and their communities multimedia, which are mostly available for free downloading at the official website: <http://www.inali.gob.mx>.

There are music and audio CD's with testimonies and letters of speakers in their indigenous languages; as well as books of poetry, stories and riddles, dictionaries, vocabularies, and alphabets.

Also, there have been animated productions from the presentations of the General Law of Linguistic Rights of Indigenous Peoples in order to spread linguistic diversity among the children, radio programmes and the ongoing campaign of "Los Guardavoces"; as well as DVD's with video in indigenous languages subtitled in Spanish. In the middle term INALI will look for opportunities to produce some TV shows.

On the other hand, the cooperative model to develop free software has been used for generating some applications such as: the Evaluation and Analysis System for the Program of Revitalization, Strengthening and Development of Indigenous Languages 2008-2012; the National Register of Interpreters and Translators of Indigenous Languages "PANITLI"; and the Catalogue of the Indigenous Languages of Mexico. Self-naming and reference material are also available for a free of charge consulting, providing opportunities for international cooperation.

Efforts in the diffusion and teaching of indigenous languages are also taken by other Mexican institutions such as the Autonomous University of Querétaro "YAAK".

Challenges and Final Ideas

In conclusion, once again I want to insist that a joint definition of objectives and clear strategies would make it possible for those who are active in cyberspace and the ICT:

1. To establish a leadership to promote and strengthen cultural and linguistic diversity; and,
2. To reinforce investments on infrastructure for ensuring indigenous peoples and their communities' access to cyberspace and ICTs as a priority human right.

The following proposals for action represent new possibilities for international cooperation in the construction of space enabling intercultural and multilingual dialog with due respect and tolerance, and the dismantling of stereotypical social representations which generate discrimination, racism and social exclusion:

- Speeding up the installation of essential infrastructure and equipment so that more indigenous people have access to cyberspace.
- Promotion of spaces which foster respectful intercultural and multilingual dialogues.
- Providing technological support to the standardization of the writing of indigenous languages, particularly with regard to the handling of special characters.
- Developing thematic agendas and inter-institutional agreements of international cooperation.
- Promoting the opening of new markets, from approaches to intercultural and multilingual communication.
- Recognizing leadership and granting awards for projects aimed at the dissemination and reassessment of linguistic and cultural diversity (“giving visibility and audibility”).

In the 21st century equality depends on our ability to recognize that we are different, and in this regard, citing Delors, to bet for a real ethics of alterity we must learn to live among different people (Delors, 2001). And one of the biggest challenges for us, users and developers of cyberspace, is to be able to recognize and to assume social responsibilities in multicultural and multilingual societies.

References

1. Delors, Jacques. *La educación encierra un tesoro*. UNESCO, 2001.
2. Instituto Nacional de Estadística y Geografía (INEGI). *Resultados Definitivos. XIII Censo General de Población y Vivienda, 2010*.
3. La Rue, Frank. United Nations General Assembly. Human Rights Council. *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*. 17th session, May 2011, UNESCO.
4. Lewis, M. Paul (ed). *Ethnologue: Languages of the World*. 16th ed. USA, 2009.

J. Prabhakara RAO

*Professor in Linguistics & Coordinator,
Centre for Study of Foreign Languages,
University of Hyderabad
(Hyderabad, India)*

Multilingualism in India and Indian Languages in Cyberspace

Recent years have witnessed a number of significant changes in language management. For instance, largely as a result of globalization and scientific and technological advancements, there has been a considerable focus on frequency and intensity of use of languages in cyberspace. However, cyberspace is vastly available in developed and in some of the fast developing economics like India, China, Brazil, South Africa, etc. The reasons are obvious – qualified human resources, free capital movements, transcontinental trade resulting in consumption identical products, etc. Besides, the distribution of population and big improvement in the level of communication access (telephony, mobile phones, Internet, etc.) during last five years have given boost in these countries for designing tools and implementing them vigorously for effective use of tremendous huge size of cyberspace.

The paper is divided into two parts. The first part deals with the background for multilingualism and its development. The second part concentrates on various governmental initiatives to promote multilingualism using information technology.

1. Genesis and Development of Multilingualism in India

Multilingualism represents a historical phenomenon in the Indian sub-continent. It began with the migration of Dravidians and then contact of Aryans with Dravidians. It is important to note that ‘Dravidians’ and ‘Aryan’ are not racial terms. As Krishnamurti (2003:3) observes, “still there is no archeological or linguistic evidence to show actually when the people who spoke the Dravidian languages entered India. But we know that they were already in northwest India by the time Rigvedic Aryans entered India by the fifteenth Century BC E.11”. Scholars still debate on this issue and “a truly convincing hypothesis has not even been formulated yet” (Zvelebil, 1990:123). However, this situation had provided basis for the birth of multilingualism in the Indian subcontinent. Multilingualism flourished with the spread of both the tribes across India and their contact with local austroloid tribes (a hypothesis yet to be established). In later centuries, Sanskrit which was the language of rituals became archaic and new forms of Sanskrit such as Pali, Prakrits came into

existence (Deshpande 1979). When Buddhism came into existence, it played an important role in consolidating multilingualism. It encouraged to write all its scriptures in Prakrit and Pali. Sanskrit gradually remained as a language of rituals when Buddhism began spreading over central and southern Dravidian territories. The Buddhist scripts and Official Orders were written in both Pali and regional language. This explains further development and maintenance of multilingualism in the Indian subcontinent. Cultural fusion between Aryan and Dravidian tribes which has taken place almost since beginning of the contact was further intensified. This resulted in the formation of the unique Indian culture in which these two cultures occupy major part. The extensive linguistic borrowing among languages and cultural amalgamation of different tribes in India have continued in subsequent centuries. For instance, Emeneau (1956) highlighted that many features shared between Dravidian and Aryan at linguistic level allow formulating the concept of 'India as a linguistic area'. My aim of looking briefly into the linguistic and cultural history of India is to draw attention on the fact that 1) multilinguality and multiculturalism are being unconsciously maintained in the Indian society; 2) despite amalgamation of different cultures, each linguistic community in India preserved its specific cultural characteristics; 3) India represents a linguistic area, where local languages exist along with the national languages.

The unconscious existence and maintenance of multilingualism has resulted in coexistence of diverse languages in Indian society. They can genetically be classified into four groups; 1) Indo-Aryan; 2) Dravidian; 3) Munda; 4) Tibeto-Burman. Due to their co-existence over thousands of years in one geographic area, these language groups share common areal features, while preserving their distinctiveness and identity. It should also be noted that because of 'peaceful' coexistence for long period, Aryan group of languages even altered their entire grammatical system under the influence of Dravidian and became similar to that of Dravidian (Prabhakara Rao, 2000). To put it in typological linguistic terms, after getting in contact with Dravidian languages which are agglutinative in type, Aryan languages which were inflectional in type slowly converted into agglutinative type. The typological balance between the major group of languages has contributed to maintenance of multilingualism in Indian society. As it was mentioned, despite the fact that there is an amalgamation of cultures among different groups in India, cultural diversity is well maintained. Hence, scholars sometimes speak about 'pan-Indian language' and 'pan-Indian culture'. Therefore, India represents an illustration for the dialectical principle of unity in diversity and diversity in unity, which has to be thoroughly studied.

After the independence in 1947, Hindi was declared as the official language of Union Government. A Three Language Formula (local, national and

international (English)) has been introduced to give equal rights to all languages in the country. However, English is also used for government notifications and communications between central (federal) and state (local) governments and also between local governments. It is often said that there are 1650 dialects spoken by different communities across the country. The government of India has recognized 22 languages as constitutionally approved languages. And there are demands for inclusion of some languages from time to time. The state governments are free to use their local languages in all domains of life (i.e. education, official, etc.). In addition to regional language(s) at state level, there exists good number of tribal languages. State governments are engaged in promoting and preserving tribal languages and cultures by providing scripts, in most of the cases, the script of regional languages, and preparing language text books at primary level. They are also making efforts to prepare dictionaries from tribal to regional languages. Andhra Pradesh state government could be named as an example in this regard. A commendable step has been taken by the Government of India to establish a Central University for tribal languages in the state of Madhya Pradesh. Some state governments are also seriously thinking to follow the same suit.

Central Institute of Indian Languages (CIIL) was established in 1969 with the objective “to assist in and coordinate the development of Indian languages, to bring about the essential unity of Indian languages through scientific study and inter-linguistic research and to promote the mutual enrichment of the languages and thus contribute towards emotional integration of the people of this country”. Simultaneously regional language centres were also set up “to take steps for implementing fully in all states the three language formula” which was expected to protect multilingual character of Indian life. Language institutes have been established in states for effective implementation of regional languages. Translation departments are existing in the state administration to translate central government documents, from Hindi and English to regional language as well as to translate the state government documents (especially communication with central government) into Hindi and English.

Apart from departments of regional languages in central and state universities and other academic institutions, regional language institutions and universities have come up in many states which concentrate on research, teaching and publication of materials in respective languages. Central institute of Hindi, Agra, Maulana Azad National Urdu University (MANU), Hyderabad, National Council for promotion of Urdu (NCPU) under MHRD, National Commission for Scientific and Technical Terminology, etc. are working under the Government of India. There are universities, institutions and departments at both central and state level to carry out research in Sanskrit language and to investigate in its rich scientific and cultural heritage.

All the above academic, language and language planning institutions established by central and state governments show the governments' intentions to protect the multilingual and multicultural fabric of Indian society and also to strive for cultural integration of the country with affecting individual cultures. However, it should be mentioned that these institutions and bodies do not always function effectively due to lack of vision, authority and funds.

Another major step towards protection and promotion of multilinguality and multiculturalism in India is the establishment of the National Translation Mission (NTM) on the recommendations of National Knowledge Commission (NKC). The main objectives of the Mission are: 1) to encourage translations among Indian languages; 2) to do translations from foreign languages into Indian and vice-versa; 3) to promote research in Machine translation; 4) to prepare translation manuals and so on. It is well-recognized that translation is one of the major means of transferring knowledge from one society to another. Translational activity enables to enrich languages not only lexically but grammatically as well. Translation among Indian languages (which is one of the objectives of NTM) undoubtedly supports the sustainability of multilingual and multicultural character of India by bringing people of different cultures together and promotes cultural integration of Indian society.

All these initiatives and measures of governments definitely ensure the development of multilingualism in India which has been existing historically since centuries. However, governments should focus on strengthening the already existing institutions and come up with the national policy on language studies and promotion of translation studies in the country. Various documents and recommendations of UNESCO shall be taken as basis while designing such policy. It is not out of place to mention that the University of Hyderabad along with the Mahasarakham University recently submitted a document on 'Preserving and Promoting Asian Languages and Cultures' to the ASEAN Secretariat, Jakarta for its consideration.

2. Initiatives by Central and State Governments to Promote Multilingualism Using Information Technology

Since the promulgation of such significant documents as the Universal Declaration on Cultural Diversity (2001), Recommendations concerning the Promotion and Use of Multilingualism and Universal access to Cyberspace (2003), Convention for the Safeguarding of the Intangible Cultural Heritage (2003), Convention on the Protection and Promotion of the Diversity of Cultural Expressions (2005) and a number of subsequent initiatives UNESCO attracted the attention of its

member states by emphasizing the need to safeguard plurality of languages and cultures as the intangible Heritage of Humanity.

Cyberspace provides a unique opportunity to procure information about everything. India which possess a large quantity of English-speaking and technical human resources, can utilize and implement information technologies in all fields of life. Today India is one of the biggest software services providers and software developers in the world.

UNESCO Recommendations concerning the Promotion of Multilingualism clearly enunciates that "... linguistic diversity in the global information networks and universal access to information in cyberspace are at the core of contemporary debates and can be a determining factor in the development of knowledge-based society". It also underlines that "basic education and literacy are prerequisites for universal access to cyberspace".

Recently India has enacted the Right to Information Act (2005) and the Right of Children to Free and Compulsory Education Act (2009). Now it is contemplating with the idea to encompass preprimary education also in these acts. Government of India is spending huge amount of money through Rajeev Sarwa Siksha Abhiyan ("Education for All" programme) on primary education and literacy. To give free access to information on all Acts, Bills, Decisions of various committees, Judgments, etc, the government made it mandatory to make them available on the Net. The information that is not available on the Net shall be provided to citizens with 5 days of applying for such information.

India is not only a multilingual and multicultural, but also a multiscrypt country. The 22 official languages are written in 10 different scripts. Hence, it is a real challenge for specialists to design tools for information processing in local languages at low cost to bring 'Digital Unity' and to make 'knowledge available for all'.

To build knowledge societies, it is essential to store, to transfer and to transmit that knowledge in a multilingual form and make it easily and freely accessible to people. This enables to build inclusive knowledge society with rapid economic growth. It seems government of India is totally convinced with this fact and initiated accordingly large number of measures to implement it (Vikas Om, 2001; Report by India to UNESCO, 2007).

Department of Information Technology of the Government of India initiated a major program called Technology Development of Indian Languages (TDIL) with the objective to develop tools for Machine Translation in Indian languages, for language studies and research, e-governance, e-learning, etc. for a group of major Indian languages. This job was entrusted to thirteen resource

centres for Indian Language Technology Solutions (RC_ILTS) which cover 10 languages. We are proud that Centre for Applied Linguistics and Translation Studies, University of Hyderabad, is one of the major centres in India where tools for machine translation for Indian languages (IL-ILMT project) are being developed. Today software packages are available for machine translation from one Indian language to another (for 14 languages), spell checks, on-line bidirectional and multidirectional electronic dictionaries, Hypergrammars, morphological analyzers, text-to-text, speech-to-text, etc. Both Central and state governments are actively implementing ICTs, as a result such important domains like health, education, tourism, public services, etc. are made available on-line. The Central government has also established (and some are at planning stage) centres for Indian language Technology and Resources (ILTR) which “would coordinate with state IT department and language departments for providing the inputs feedback on technological issues such as localization, etc.”; Community Information Centres (CICs) – to bridge digital gap between urban and non-urban areas; common service centres (CSCs) – to provide internet access to rural citizens. A significant contribution to the emergence of networking in the country was made by the Education and Research Network (ERNET). National Informatics Centre (NIC) was also established under the Department of Information offering services such as corporate communications, Internet/Intranet emerging voice communication, multimedia, broadcast, distance learning, etc. It should be mentioned that all these services are not always in local languages. However, the government committed to make them available in major Indian languages as early as possible. In addition to governments’ efforts some MNCs like Microsoft, HCL technologies, Google services etc. and National companies like TATA consultancy services, etc. are seriously engaged in developing a variety of software tools for Indian languages.

However, there is still a lot of work to be done. For example, translation is yet to get industry status (Prabhakara Rao: 2001). Language studies should be encouraged as it is crucial for language industry and language engineering. Much remains to be done in localization of software. Both Central and State Governments should more seriously be committed for implementation of regional languages in all domains.

References

1. Cultural and Linguistic Diversity in the Information Society. Paris: UNESCO, 2003 (I-2004/WS/3).
2. Deshpande, Madhav M. Sociolinguistic Attitudes in India. A Historical Reconstruction. Ann Arbor: Karoma Publishers Inc., 1979.

-
3. "Preserving and Promoting Asian Languages and Cultures", document submitted by the University of Hyderabad, India, and Mahasarakham University, Thailand, to ASEAN Secretariat, Jakarta.
 4. Emeneau, M. B. India as a Linguistic Area. *Language* 32:3-16. 1956.
 5. Krishnamurti Bh. *The Dravidian Languages*. Cambridge University Press, 2003.
 6. Prabhakara Rao, J. Towards Commercialization and Industrialization of Translation in India (in Press). *Festschrift for Prof. J.P. Dimri*, EFLU, Hyderabad, 2011.
 7. Higher Education and Industry: Some Thoughts. A paper presented at India-EU meeting, Warsaw, Poland, 2011.
 8. Multiculturalism and Multilingualism in India and in Europe: A Comparative Study (in Press). In: 'Centre for Contemporary India Research and Studies', University of Warsaw, Poland, 2011.
 9. Recommendations concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace.
 10. Report by India to the UNESCO General Conference on the Implementation of the Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace. Government of India, MCIT, Dept. of Information Technology, New Delhi.
 11. Right of Children to Free and Compulsory Education Act of 26th August, 2009, Government of India.
 12. Right to Information Act, 2005, Government of India.
 13. Systemic Linguistics: Some Remarks. *Indian Linguistics*. Vol.60: 1-4, Pune, 2011.
 14. Vikas Om. *Language Technology Development in India*, 2001.
 15. Zvelebil, K.V. *Dravidian Linguistics. An Introduction*. Podicherry: Podicherry Institute of Linguistics and Culture, 1990.

Natalia GENDINA
*Director, Research Institute for Information
Technologies in Social Sphere,
Kemerovo State University of Culture and Arts;
Member, Russian Committee,
UNESCO Information for All Programme
(Kemerovo, Russian Federation)*

The Importance of Information Literacy in the Development of Multilingualism in Cyberspace

Preservation of linguistic diversity is one of the global problems and challenges of cultural ecology. At the end of the twentieth century humanity faced a complex of socio-natural acute contradictions that affect the world in general as well as particular regions and countries. Under generally accepted classification developed in the early 1980s three main groups of global problems are distinguished:

- problems associated with basic human social communities (prevention of global nuclear catastrophe, closing the gap in the levels of socio-economic development between developed and developing countries, etc.);
- issues concerning the relationship between man and environment (environmental, energy, raw materials and food, space exploration, etc.);
- problems requiring special attention to the relationship between man and society (profiting from scientific and technological progress, elimination of dangerous diseases, health care improvement, eradication of illiteracy, etc.).

There are other classifications of global problems, but any of them is arbitrary, since all problems are closely related, have no clear boundaries and overlap each other.

One of the global problems is the rapid loss of linguistic diversity of mankind. Hundreds of languages are endangered: languages with a small number of speakers that have no writing and other signs of high social status, the so-called “small” or “minority” languages. This process can be compared to a decrease in the Earth’s natural diversity. Environmentalists around the world precisely estimate the loss of biodiversity as a catastrophe. However, socio-cultural consequences of language extinction and decreasing linguistic diversity is hardly less dangerous than those of the decline of biodiversity.

The writings of the eminent German linguist W. von Humboldt and his 20th century followers German linguist L. Weisgerber and U.S. ethnolinguists E. Sapir

and B. Whorf carry the theory of linguistic relativity. According to this theory, people speaking different languages, have different visions of the world, so each language reflects its speakers' logic of thinking. Turning to another language, we borrow another way of thinking, another way of understanding the world.

Thus, language loss and linguistic diversity reduction result in the loss of valuable information, because each language is a unique way of describing the world, expressing a people's identity and transmitting social and cultural traditions. It is the most important means of transferring unique knowledge, accumulated by a given nation.

Like any other global challenge, the problem of linguistic diversity preservation is characterized by a number of criteria:

- manifestations of magnitude that go beyond the limits of a single state or group of countries;
- topicality;
- complexity: all problems are intertwined with each other;
- universal character of the problem, understandable and relevant to all countries and peoples;
- requiring solution by the entire international community, all countries and ethnic groups.

Preservation of linguistic diversity can justifiably be classified as an essential problem of cultural ecology.

Search for Methods and Systematic Solutions to Safeguard Linguistic Diversity as One of Mankind's Global Problems

Surge and exacerbation of global problems of mankind requires for developing a complex understanding and choosing best solution methods. UNESCO has long been playing a leading role in coordinating international efforts to preserve linguistic diversity. In recent years the preparation of a number of important events (8) and documents (5; 6; 11; 13) has been initiated.

In the Russian Federation the idea of preserving linguistic diversity is being actively promoted by the Russian Committee of the UNESCO Information for All Programme. Thanks to its efforts information on the ways of solving the problem of the linguistic diversity preservation in the modern world is collected, analysed and interpreted in Russia. The analysis of papers published by the Russian Committee (9; 10; 14) reveals the following set of measures, as well as forms and methods used in Russia to preserve multilingualism in cyberspace.

*Activities of social institutions, organizations and public institutions in Russia
to preserve multilingualism in cyberspace*

Social institution, organization, public institution	Types of Activities
National authorities	Improvement of language legislation of the Russian Federation regulating the functioning of the state and native languages, the development of national education and culture. Elaboration of language public policies, creation of special programmes of language development.
Local authorities	Elaboration of specialized programmes of language development, indigenous cultural heritage maintenance and promotion. Establishment of legal information access centres. Provision of television and Internet broadcasting and online media in the languages of indigenous peoples.
Educational institutions (schools, colleges, universities)	Provision of education in national (native) languages. Creation of textbooks, including media didactic materials, on national languages. Development of multimedia teaching materials in national languages.
Scientific institutions and universities, research institutes	Research on multilingualism preservation and development. Documentation of endangered languages.
Libraries, archives, museums	Formation and preservation of museum collections, documents, books in indigenous languages, images, filmed documents, video recordings, local content. Creation of digital collections of objects of cultural and natural heritage. Creation of digital resources in indigenous languages: <ul style="list-style-type: none"> – portals – sites – databases (thematic, bibliographic, full texts) – electronic directories of local history.
Arts institutions (theaters, philharmonics, music and folklore groups, art galleries, studios, cultural centers)	Promotion of indigenous peoples' works of art, folk art, including performing in indigenous languages.
Publishers (books and media)	Publication of literature in national languages.
Bookstores	Dissemination of literature in national languages.

NGOs and individuals	Advocacy for national associations and tribal communities of the indigenous peoples. Creation of digital resources in indigenous languages: <ul style="list-style-type: none"> - portals - sites - thematic databases.
----------------------	--

Analysis of these data suggests that a system of measures, forms and methods of work of various institutions, from government to private individuals, is being developed in Russia to preserve linguistic diversity. However, this system does not explicitly mark one of the tools without which the preservation of multilingualism in cyberspace and digital environment is not feasible. Such an important and essential tool, in our view, is information literacy.

Information Literacy as One of the Tools for Safeguarding Linguistic Diversity

The term “information literacy” was adopted by the international community to refer to a wide range of competences and skills related to the ability of individuals to use information and communication technology (ICT), in order to confidently navigate the huge flows of information, be able to locate, evaluate and effectively use this information to solve various problems of the modern world. UNESCO and IFLA were the leading international organizations to initiate the promotion of the idea of information literacy. As a result the concept of information literacy has been formed. The following steps have been taken by IFLA and UNESCO to promote information literacy in the world:

- creation of the Information Literacy Section of IFLA (68th Session and the General Conference of IFLA in Glasgow, 2002);
- creation of the IFLA-UNESCO strategic alliance for the implementation of decisions of the World Summit on the Information Society, including problem solving on information literacy (UNESCO Open Forum on the 72nd IFLA World Library and Information Congress in Seoul, 2006);
- publication in 2006 of “Guidelines on Information Literacy Education for Lifelong Learning”, edited by J. Lau, Chair of the Information Literacy Section of IFLA (2);
- publication in 2008 of the book “Information Literacy Indicators”, prepared within the framework of the UNESCO Information for All Programme (4);

-
- publication in 2008 under the auspices of IFLA of the book “Information Literacy: International Perspectives”, translated and published also in Russian on the initiative of the Russian Committee of the UNESCO Information for All Programme (3);
 - holding an international expert group meeting on the development of indicators of media and information literacy by UNESCO, UNESCO Communication and Information Sector, and UNESCO Institute for Statistics on 4-6 November, 2010 in Bangkok (Thailand).

A special place in advancing information literacy has the UNESCO Information for All Programme. The essence of the Programme’s concept is the notion of a world where each individual has access to information important to him and possess necessary skills and ability to use information obtained to build a better society. Information for All Programme raises the awareness of the importance of information literacy for all and supports projects that promote the development of information literacy skills.

The Strategic Plan of the UNESCO Information for All Programme for 2008–2013 places information literacy among the Programme’s top priorities along with information for development, information preservation, information ethics, information accessibility (12). Interpretation of the “information literacy” concept in this document is based on the special declaration adopted on November 9, 2005 at the Alexandria Colloquium on Information Literacy and Lifelong Learning and known as “Beacons of the Information Society: The Alexandria Proclamation on Information Literacy and Lifelong Learning.” As noted in the Alexandria Proclamation, “Information Literacy lies at the core of lifelong learning. It empowers people in all walks of life to seek, evaluate, use and create information effectively to achieve their personal, social, occupational and educational goals. It is a basic human right in a digital world and promotes social inclusion of all nations” (1).

According to the Alexandria Proclamation, information literacy:

- comprises the competencies to recognize information needs and to locate, evaluate, apply and create information within cultural and social contexts;
- is crucial to the competitive advantage of individuals, enterprises (especially small and medium enterprises), regions and nations;
- provides the key to effective access, use and creation of content to support economic development, education, health and human services, and all other aspects of contemporary societies, and thereby

provides the vital foundation for fulfilling the goals of the Millennium Declaration and the World Summit on the Information Society; and

- extends beyond current technologies to encompass learning, critical thinking and interpretative skills across professional boundaries and empowers individuals and communities (1).

Based on this interpretation, we emphasize the role and importance of information literacy as an essential means of facilitating the task of linguistic diversity preservation in cyberspace. If you re-examine the above table showing the main action lines for social institutions, organizations and establishments to preserve multilingualism in cyberspace, almost all of them require for citizens' proficiency in information literacy.

Functions of Information Literacy as a Means of Linguistic Diversity Preservation

The most important functions of information literacy as a means of preserving multilingualism in cyberspace are connected with accessing information and communication.

The “key” function of information literacy is its being a kind of a key that opens the door to information storages. We emphasize that this can be both traditional (libraries, archives, museums), and electronic information storages. Mastering information literacy allows individuals to get access to socially important information contained in the electronic environment, including the Internet. Without information literacy one cannot be provided with public access to modern digital resources: online newspapers and magazines, databases, Web sites and portals containing a wealth of legal, linguistic, educational and scientific information that reflects the rich traditions and culture of the peoples of the world, including small and indigenous nations.

Another important function of information literacy is giving people an opportunity to communicate in digital environment. It expands opportunities for communication and interaction in cyberspace for native speakers and people studying a certain language, facilitates integrating the efforts of all those interested in the preservation and promotion of multilingualism, regardless of their location and distance from each other, through the use of ICT. The major benefits of ICT in this regard are:

- openness – ability to access necessary information resources and communicate with all those interested in the preservation and promotion of multilingualism;

-
- interactivity – active interaction of all stakeholders and usage of network information resources with feedback provided;
 - efficiency – high-speed information exchange, ability to regularly update and promptly amend information;
 - convenience – usability of the digital information environment and the possibility of access for remote users at any time convenient to them.

In our view, the role of information literacy in the preservation of linguistic diversity in cyberspace is not limited to the above-described two functions. Information literacy surely performs the adaptive function as well ensuring individuals' adaptation to the new challenges of a rapidly changing information society. Moreover, information literacy is also an important factor of development, as it is aimed at enriching one's mental capacity and inner world.

Information literacy is the foundation of any cognitive process, including education, and scientific research. It is a tool for tackling practical vital tasks requiring for the use of appropriate information and relevant knowledge and skills.

The protective or preventive role of information literacy should be emphasized, allowing individual to protect himself from the negative effects of computerization and ICT development. Mastery of information literacy skills gives people a tool for protection from risks and challenges of the information society, connected with the huge volume of information, often unreliable and contradictory, from ICT penetration into all spheres of life and danger of manipulation of human consciousness.

In this regard, developing critical thinking is essential for current training programmes on information literacy. Critical thinking allows to select, analyze and interpret information, draw one's own conclusions and form own point of view on various social, cultural, political, and other aspects of life instead of blindly trusting other's opinion.

In conclusion, we want to emphasize once again the complexity of the problem of multilingualism preservation in cyberspace in the context of globalization. Its solution lies outside the scope of simple and unambiguous decisions, and requires for the integration of efforts by national and local governments, education, science, memory institutions (libraries, archives, museums), art institutions, both traditional and electronic media, public organizations and private individuals. It involves large-scale, long term and, most importantly, systemic activities including raising the level of information literacy of citizens.

Information literacy development for preserving multilingualism, in turn, requires for state support and provision of the following conditions:

-
- 1) organizing citizens' training in information literacy through educational institutions and libraries of all kinds and types;
 - 2) organizing professional training for specialists to teach information literacy to different categories of learners, including native speakers of various languages, representatives of small and indigenous peoples. This problem might be solved by using the potential of teachers and librarians (information specialists) mastering the technologies of information literacy development with due account for ethnicity, age and type of activities of students;
 - 3) creating and using distributed information and learning environment, including specialized information resources (primarily documents in different languages, including small and indigenous peoples' languages, information publications and guides to electronic resources reflecting the culture and traditions of various nations, traditional and electronic library catalogs), computer equipment, means of access to remote domestic and world information resources. "Distributed" information and learning environment presupposes that its constituent components are concentrated in educational, library and information institutions;
 - 4) creating educational literature on information literacy, profiled in accordance with the problem of preservation of multilingualism in cyberspace. Educational materials, facilitating the formation of information literacy must be presented in print and digital form, and should focus on both educators and learners.

References

1. Beacons of the Information Society: The Alexandria Proclamation on Information Literacy and Lifelong Learning [Electronic resources]. – URL: <http://www.ifla.org/en/publications/beacons-of-the-information-society-the-alexandria-proclamation-on-information-literacy->
2. Guidelines on Information Literacy for Lifelong Learning. – Access mode: <http://www.ifla.org/VII/s42/pub/IL-Guidelines2006.pdf>.
3. Information Literacy: International Perspectives / Edited by Jesus Lau – Munich: K.G. Saur, 2008. – 160 p.
4. UNESCO. Information for All Programme (IFAP). Towards Information Literacy Indicators. Conceptual framework paper prepared by Ralph Catts and Jesus Lau. Edited by the Information Society Division, Communication and Information Sector, UNESCO: Paris, 2008. – 44 p.

-
5. Всеобщая декларация ЮНЕСКО о культурном разнообразии // Культурное и языковое разнообразие в информационном обществе. – СПб., 2004. – С. 65 – 75.
 6. «Ленская резолюция». Итоговый документ [Международная конференция «Языковое и культурное разнообразие в киберпространстве» (Якутск, Российская Федерация, 2-4 июля 2008г.)][Электронный ресурс]//Российский комитет Программы ЮНЕСКО «Информация для всех». – Режим доступа: http://www.ifapcom.ru/files/Documents/lena_resolution_rus.doc
 7. Лихачев, Д. С. Экология культуры / Д. С. Лихачев // Москва. – 1979. – № 7 – С. 173. – 179.
 8. Международная конференция «Языковое и культурное разнообразие в киберпространстве» // Вестник Библиотечной Ассамблеи Евразии. - 2008. - N 4. - С. 10-14.
 9. Многоязычие в России: региональные аспекты. – М.: Межрегиональный центр библиотечного сотрудничества, 2008. – 136 с.
 10. Представление языков народов России и стран СНГ в российском сегменте Интернета. Семинар Российского комитета Программы ЮНЕСКО «Информация для всех» и Межрегионального центра библиотечного сотрудничества в рамках Международной конференции «EVA – 2007 – Москва»: сборник докладов. – М. : Межрегиональный центр библиотечного сотрудничества, 2008. – 104 с.
 11. Рекомендации ЮНЕСКО по многоязычию и доступу к киберпространству // Культурное и языковое разнообразие в информационном обществе. – СПб, 2004. – С. 77 – 88.
 12. Стратегический план Программы ЮНЕСКО «Информация для всех» на 2008 – 2013 гг. / Комиссия Российской Федерации по делам ЮНЕСКО; Российский комитет Программы «Информация для всех». – М., 2009. – 47 с.
 13. Хартия о сохранении цифрового наследия // Культурное и языковое разнообразие в информационном обществе. – СПб., 2004. – С. 89 – 96.
 14. Языковое разнообразие в киберпространстве: российский и зарубежный опыт. Сборник аналитических материалов / сост. Е.И. Кузьмин, Е.В. Плыс. – М.: Межрегиональный центр библиотечного сотрудничества, 2007. – 218 с.

Valery S. YEFIMOV

*Head of the Centre for Strategic Research and Development,
Siberian Federal University,
(Krasnoyarsk, Russia)*

Preserving Linguistic and Cultural Diversity in the Republic of Sakha (Yakutia): Foresight Project

Preserving linguistic and cultural diversity is a need no one calls into question these days, with its importance consolidated in a whole number of international documents and intergovernmental agreements. Yet, small indigenous languages are still facing the risk of extinction, and that risk may be exacerbated by the ongoing process of information globalization. It is not unlikely that the expansion of information and communications technology and the emergence of global cyberspace will lead to a narrow group of major languages taking over the world, with smaller ones pushed to the sidelines.

How Fast May Cultures and Languages Disappear?

The extinction of minor languages was an issue prominent on the agenda of the first international conference “Linguistic and Cultural Diversity,” held in Yakutsk in 2008. By that time, Russia had seen three of its 105 languages disappear while in the United States, with a total 238 languages spoken there originally, 76 had become extinct and another 67 had found themselves on the brink of extinction³⁰. A language may go extinct within a relatively short period of time – three to four generations.³¹

In 2010, we conducted a survey on the status of small indigenous languages spoken in Russia’s Krasnoyarsk Region: Evenki, Selkup, Keto and Chulym. The survey showed that over the past century, the share of native speakers in respective generations had decreased, on average, from 100% in the generation born between 1910 and 1930 (great-grandparents) to 8% in the generation born in 1990-2010 (grandchildren).

³⁰ V. Montviloff. *Internet and Language Diversity: Is It Possible? In: Linguistic and Cultural Diversity in Cyberspace. Proceedings of the International Conference (Yakutsk, Russian Federation, 2-4 July, 2008).* – Moscow: ILCC, 2011.

³¹ Мурадова А. Р. Как исчезают языки и как их возрождают // *Языковое разнообразие в киберпространстве: российский и зарубежный опыт. Сборник аналитических материалов / Сост.: Е. И. Кузьмин, Е. В. Плыс.* – М.: МЦБС, 2008. – С.70–75.

The proportion of native speakers routinely using their mother tongue in daily life had shrunk over the period to 3% (grandchildren's generation), up from 65% (great-grandparents). The disappearance of minor languages is a disastrous process, which in some cases may happen over just two or three generations.

A survey undertaken that same year on the Yakut language, though, showed its high degree of sustainability, with a fluent command demonstrated by 100% in great-grandparents' generation and by 86% in grandchildren's.

The decrease in native speaker numbers is more pronounced with urban dwellers. In rural areas, 100% of the ethnic Yakut born between 1910 and 1930 and 79% of those born in 1990-2010 demonstrate fluency in their native language. In cities, meanwhile, the figures are 93% and 61%, respectively.

In ethnic Yakut inhabitants, aged 15-19 and 20-29, the willingness to preserve their cultural identity is much harder to come by than in older community members, aged between 30 and 65.

To gauge the pace of processes related to native language and culture transmission, we have split the ethnic Yakut population into two categories depending on whether the original identity is neglected or preserved. In the former category, people attach little importance to their traditional culture; they are reluctant to cultivate ethnic traits in themselves, and have no willingness to use their mother tongue, nor teach it to their children (only 48% of the respondents intend to teach it, as compared with 95% in the latter category).

The survey's findings show a deformation in the mechanisms of value, language and culture reproduction in the ethnic Yakut during their industrial and post-industrial transition.

Prospects and Risks of Yakutia's Economic and Socio-Cultural Development in Global Context

The limited amount of proven natural reserves and the industrial boom of major Asian economies (such as China and India) are likely to prompt major world powers' political, economic and socio-cultural expansion into the Arctic and circumpolar areas in the next two decades or so. Countries like the United States, Russia, Canada, and Norway are expected to step up their industrial activity in those territories. This will dramatically increase the migrant inflow in the sub-Arctic, along with bringing in different lifestyles, value systems, and socio-cultural standards. Faced with cultural and economic occupation in their ancestral lands, the region's indigenous communities may find themselves on the brink of cultural extinction within two or three generations.

In the next two decades, the Republic of Sakha will be the scene of ambitious socio-economic and socio-cultural transformations, which may drastically change the living conditions of the local indigenous communities. In keeping with the federal government's Strategy for Socio-Economic Development of Russia's Far East and the Baikal Region through 2025, the gross regional product is expected to grow 8.5-fold on the year 2005. That growth could be provided primarily by large mining and transportation projects -- the driving force behind the republic's future industrial advancement.

The potential threat to the reproduction and existence of sub-Arctic indigenous communities is determined by three "waves" in Yakutia's cultural and economic development in the years to come:

- 1) New industrialization, that is, operations in the republic's territory of large Russian, foreign and transnational corporations (predominantly companies involved in mining);
- 2) Innovative progress, that is, high-tech projects to be launched by the government and the business community, and the introduction of new production lines and services relevant to the post-industrial stage;
- 3) Acculturation on the part of major world players, such as the United States and the Eurozone countries, who have a strong influence on global media and the Internet. They will try to impose their culture and value systems on the indigenous population, making extensive use of postmodernist humanitarian techniques, aimed at destroying traditional cultures and building a globalist, consumption-driven society.

In these conditions, the very vitality of the indigenous communities' traditional economic patterns and their mechanisms for intergenerational transfer of cultural heritage and value systems will be put to test.

The related problems include:

- dilution of indigenous communities' livelihoods against the backdrop of large-scale development of local natural resources by large corporations;
- exacerbation of social and environmental problems arising from the narrower spread of traditional occupations and the impossibility of the native population's full-fledged integration into the emerging industrial and postindustrial realities;
- young people's loss of ethnic identity and breakaway from their native culture (including language, communication and conduct patterns, as well as value systems, under the pressure of mass culture and consumerism);

-
- destruction of the genetic fund as a result of migrations and birth rate decline amidst growing urbanization and the emergence of new medical and socio-medical problems.

Life shows indigenous communities' high sensitivity to aggressive industrial and postindustrial development and the likelihood of their consequent marginalization. Their role in humanity evolution risks being brought down to that of conserved and protected "relics," eventually.

Having said that, the prospective expansion of mining operations in Yakutia and the advancement of high-tech sectors, with the North-Eastern Federal University among the major R&D contributors, will provide indigenous communities with vast opportunities for breakthrough. There is a possibility of creating an economic model that would enhance the financial and economic foundation for their sustenance and development. A broader scale and variety of products, services and socio-cultural activities would create a window of opportunity for expanding the competency range of Yakutia's indigenous inhabitants, thereby enriching its social and human resources.

Measures to Preserve Cultural and Linguistic Diversity: Opportunities and Constraints

International documents³² adopted on the issue in the past few decades recognize indigenous communities' political, economic, and cultural rights to the preservation of their social, cultural, religious and spiritual values and economic practices, as well as the rights of property and ownership of their ancestral lands. All this may be instrumental in helping preserve such communities' cultures and languages.

Cyberspace and ICT have created an additional communications dimension and a new form of existence for languages and cultures. The international community and Russian authorities should do more to overcome the digital gap and to expand the use of cyberspace in efforts to preserve indigenous cultures and languages.

However, such efforts alone will not be enough to counter the processes of Western culture, values, activities and lifestyles conquering the world. The domination in the information landscape of major languages, such as English, Chinese and Spanish, ensures the spread of world powers' value systems globally, depreciating the significance of minor languages to their native

³² *International Labor Organization Convention No. 169 on Indigenous and Tribal Peoples, Geneva, 1989; UN Declaration on the Rights of Indigenous Peoples (adopted by Resolution 61/295 of the General Assembly, September 13, 2007); Concept for Sustainable Development of Small Indigenous Peoples of Russia's North, Siberia and Far East. Moscow, 2009.*

speakers. The presence of indigenous cultures and languages on the Web cannot guarantee their wider use by speaker communities in real life. There is also a risk of turning living indigenous cultures into a mouthballed heritage for museum conservation and display. The preservation of cultures and languages is no easy task, and there are still no established managerial practices to implement it.

The problems of culture and language preservation management arise from a need for long-term commitment (50-100 years), the difficulty of foreseeing economic and socio-cultural scenarios, and the complexity of the object to be preserved.

Foresight Project as a Tool for Preserving Linguistic and Cultural Diversity in Yakutia

The Foresight Yakutia project is being run by two leading higher education establishments: the Siberian Federal University and the North-Eastern Federal University. Launched in 2010, it follows up on the republican government's policy toward preserving cultural and linguistic diversity.

This project is a complex and multidisciplinary one. Its research and planning groups are comprised of methodology specialists, economics, sociologists, demographers, medical doctors, cultural anthropologists, specialists in culture studies, ethnographers, teachers, historians, and philosophers. The research and development work involves more than 50 scientists and scholars, many of whom hold high academic degrees.

As a new tool for “working with the future,” Foresight includes the following dimensions:

- foreseeing the future (identifying basic trends in the development of large socio-economic systems, countries, regions, corporations, etc.);
- managing the future (comparing forecasts by key players and coordinating their strategic goals);
- promoting the future (drawing roadmaps that would show possible routes, bifurcation points, and windows of opportunity).

Foresight is to be rerun every five years for verifying the findings, carrying out a critical analysis of the practical results, and updating the tasks set.

A modern centre will be set up on the North-Eastern Federal University grounds to perform the following tasks: conducting research and monitoring sociocultural processes in circumpolar areas; elaborating modern socio-humanitarian techniques; and implementing pilot projects in socio-

humanitarian practices for the preservation and development of northern ethnicities.

The project has no analogues in Russia's cultural policy practices.

Activities:

- arranging comprehensive study and systemic planning of a long-term future for the indigenous communities of Yakutia; pushing the foresight boundaries to 40-50 years through the engagement of a broad range of competent Russian and foreign experts;
- forming a public consensus, creating a broad-based public coalition for translating the most preferable of scenarios into reality, organizing and supporting the processes of long-term preservation and reproduction of Yakutia's cultures and languages;
- elaborating principles of socio-cultural policy, building strategies and programmes for the preservation and promotion of Yakutia's languages and cultures within the format of strategic partnership between the government, the business community, and the general public.

Aims and goals:

1. analyzing the problems of reproduction and development of socio-anthropological systems amid global transformations; elaborating a systemic model for ethnicities' socio-cultural reproduction processes;
2. studying exterior economic, political and cultural processes that define the possibilities and prospects for the existence and development of Yakutia's indigenous communities in the long term;
3. analyzing the demographic, economic, social and cultural performance indicators for the Republic of Sakha (Yakutia); the traditional economy's potential; the potential of basic economic sectors; communities' demographic and socio-cultural reproduction processes and institutions supporting that reproduction; cultural processes; socialization patterns (in adults and children alike); communities' traditional occupations, culture, language, value systems, mentality and forms of communication;
4. exploring sensitivities of Yakutia's indigenous communities, including their socio-psychological, demographic, cultural, anthropological and other characteristics, as well as specific features in socio-cultural, demographic and other processes that are seen as sources of risks and threats for the preservation and further development of indigenous cultures and languages;

-
5. analyzing preservation and development practices for northern indigenous communities in and outside Russia, including their economies, languages, culture and cultural standards, social life, and mentality; identifying problems in these practices and working out efficient models and formats for getting them solved;
 6. forming a coordinated vision of the future of Yakutia's indigenous communities, including aspects such as economic and cultural positioning, traditional economy, demographic development, socio-cultural and socio-anthropological development);
 7. identifying critical techniques essential to advancing into the future, including humanitarian technology, tools for socio-cultural engineering; new educational, cultural, healthcare, and socio-political practices;
 8. determining a 'scenario field' and describing major scenarios for the development of Yakutia's ethnicities, to cover the entire range from best-case through worst-case; selecting the 'basic scenario,' based on key players' consensus and vision for the future. Making an economic assessment of the scenarios, including the amount of investments required, the possibility of recouping the costs, etc.
 9. developing project proposals in major areas of activity: healthcare (type of treatment, the desired profile of medicine, medications, and diet; defining the notion 'healthy lifestyle' in reference to the indigenous communities of the North); education (socialization for integration into industrial and postindustrial formats of activity; transmission of traditional culture; specificity of childhood as an institution and the logic behind growing up into adulthood; result-oriented education); cultural practices (transmission of mentality specifics, systems of values, methods of identification, patterns of early-age and mature-age socialization, etc.);
 10. Drawing a roadmap that would allow to coordinate efforts by the authorities, members of the business community and the public, and welfare organizations involved in the implementation of the basic scenario;
 11. Developing guidelines for a policy of sociocultural development in the Republic of Sakha (Yakutia) through the year 2025.

SECTION 1. INSTRUMENTS FOR LANGUAGE PRESERVATION AND PROMOTION IN CYBERSPACE

Marcel DIKI-KIDIRI

*Senior Researcher, French National Centre for Scientific Research
(Poitiers, France)*

Creating Content for Less-Diffused Languages on a Web Site

A web site is a prominent and very important tool for promoting any given language in cyberspace. Therefore, it is worth using it to support less-diffused languages. Let us start with a brief recall of what is a web site.

1. Different Kinds of Web Sites

Basically a web site is a page of a document written in HTML or XML language in order to be published on the Web. This page may include different elements such as texts, pictures, movies and sounds. In the early age of the Web, only the webmaster could modify his website by adding to it or withdrawing from it any element of information. Interaction with the visitors of the site was limited to specific devices such as a guestbook, a forum or newsgroup. Since then, the Web technology has undergone a very deep evolution leading to what is known as “web 2.0”, a thoroughly interactive kind of web site that allows a fast development of social networks.

In this new model of web sites, every visitor may sign up and then is given a personal page with her/his profile. Users can access a large set of tools to act like a webmaster, post any document they want, and interact with an endless number of people individually, selectively or collectively. So, it becomes possible to organize all kinds of collaborative activities directly on the web site in real time. Now, the most important question is not “how to do?” but “what to do?” Indeed, the structure of your web site depends on what you want to do with it! That is where the content issue comes in.

2. Different Kinds of Content

There is no limitation to the diversity of the content that can be included in a web site. Yet, as we are interested in web sites designed to support any less-diffused language, let us draw the outlines of some of the content which may be expected on this kind of web sites. On such a site information should be found on the language

itself, on the language's situation, e.g. on where and when it is spoken and on the people who speak it, the way they live and their culture. One of the first decisions to make is whether the site should be written in the less-diffused language or in some more widely diffused one. The first option gives a good visibility to the supported language as it makes it become a working language on the web site. But only those who speak it can read and appreciate the web site's content. The second option reveals the less-diffused language to much more people in the world, but let it remain an object one can talk about instead of becoming a mean of communication. The best approach is likely to be a bilingual web site making use of the less-diffused language as well as a more widely spread one.

2.1. Linguistic Data

Information about the less-diffused language encompasses phonetic and phonology data, morphology and syntax data, orthography and writing systems, all kinds of literary texts (proverbs, tales, poetry, short stories, novels, songs, mail and CV models, advertisements, and so on.) Each of these items can be developed on several pages and where it is suitable video files can be included and lessons can be designed for learners of the language. We'll come back on this last point further down.

One of the most important language data which might deserve a whole separate website is the language lexicon. The technical approach to online dictionaries much depends on how they are intended to be used. It may be a word reference with a search field and a browser that lays out the result for each looked-up word. It may be presented as a full page of a dictionary with several entries in an alphabetic list. Multilingual dictionaries are usually based on semantic links that emphasize both similarities and differences between languages.

The growth of a language entails the enrichment of its vocabulary. As a less-diffused language becomes a working tool in cyberspace, it will necessarily need a technical terminology to talk about the Web site itself and activities carried online. A new way of speaking will come out and it is a good practice to talk about it and discuss it on a wiki space. The need of a series of specialized technical vocabulary will soon come up leading to a growing terminological activity in a variety of knowledge fields, which deserves a dedicated web site.

2.2. The Language's Situation

Linguists love classifying languages. So, one of the basic kinds of information looked for about a less-diffused language is its classification amongst its cognates. In which language kinship and typology does it fit? In which country or countries is it spoken? What can be learned about its history? What is its

social status? In a multilingual situation, which are the other languages spoken in the same area? All these questions can be answered in series of articles that fill several pages of a web site dedicated to the less-diffused language.

As mentioned above, in order to increase the number of its speakers, a whole e-learning system can be built featuring sets of lessons for beginners, and other sets for advanced learners. Beginners' lessons include phonetic exercises, and common sentences used in everyday life or typical social conversations, while advanced lessons introduce more to the people's culture and literature aiming at a better mastery of the language. E-learning requires an important online interactivity between a local staff and a growing number of faraway students, each of whom should have a profile, a personal page and follow-up. Because everybody has different reasons to take lessons in any given language, it is a good practice to design the architecture of the whole course as a tree that allows different ways to progress through the lessons. Thus, each student can choose the better way for her/him, eventually with the advice of their teacher or mentor.

2.3. The Language's Speakers

How many people speak the less-diffused language, either as native speakers or as second language speakers? Do all the speakers share the same cultural area, way of life and custom? A good description of these features provides content of a high interest. This includes all aspects of the society's life such as social organization, administration, justice, labour, education, religion, philosophy, arts, food, architecture, transport, leisure and entertainment, and so forth.

It is very common throughout the world that traditional communities develop a close relationship with specific animals they depend on for their life. For instance, a seal for the Inuits, a horse for the Mongolians, a cow for the Fulani, the Tutsi or the Texan cowboys, a sleigh dog for the Greenlanders, a reindeer for the Sami of northern Europe, a camel for the Sahara Tuaregs, a llama for the Indians in the Andes, a yak for the highlanders of Himalaya and neighboring regions, and so on. This special relationship between a specific animal and a human being community generates a very rich diversity in culture and civilization that all deserve being known to the whole mankind as a global human heritage. Who else can reveal them better than representatives of those traditional communities whose languages are usually less-diffused?

Finally, a social network web site would give the opportunity to the most committed supporters of the less-diffused language to enhance the language's life on cyberspace by using it to chat and exchange all kind of messages and private data. By so doing, they automatically strengthen the vitality of the language that will thus become more and more used and diffused.

2.4. News and Newspapers

Most of the time, people who are really committed to promote a language organize themselves into a legal cultural association. In this case, it is a good practice to give news of the association's activities on a web site in the supported language. Beside of this, a real online newspaper can be created on the same web site to inform and comment on current events in the supported language. News, comments, points of view, debates are the ingredients of these web site pages. But some items like calendar, horoscope, crosswords and similar games, meteorology information, sports, lottery and different polls can be added to make the page even more attractive.

2.5. Games and Leisure

Since Olympic and global competitions are regularly organized, some sports like soccer and rugby have become popular all over the world. During these competitions that take a huge place in the news, comments are usually given only in widespread languages, never in a less-diffused one. So, it is a good idea to find a way of talking about these games in every native language, and this can be done on a web site. On the other hand, very few is said, if any, about games, sports and leisure practiced only in local communities. Since they are part of the local culture, it may be easier to talk about them in the less-diffused language. Here again, a multilingual presentation can help disclosing them to the rest of the world.

3. Final Remarks

As I said it right from the beginning, there is no limitation to the topics which can be addressed on a web site. I simply outlined here some of the most obvious ones that I would expect from a supportive web site dedicated to the promotion of a less-diffused language in cyberspace. The multilingual dimension is also useful to allow sharing knowledge worldwide. In the modern globalised world, being informed is, of course, very important for taking decisions and acting, but it is even more important to share knowledge and learn from each other in order to build a better world locally.

Virach SORNLETLAMVANICH

*Principle Researcher,
National Electronics and Computer Technology Center
(Bangkok, Thailand)*

Linguistic and Cultural Knowledge Co-Creation on Social Network

Introduction

Social network plays an important role in establishing an online collaborative society. Being seen in current occurrence, it provides a great impact in social, political, educational, and many other movements. Such a way of the change in the way of people collaboration can be promoted in developing humanity resources to enrich our linguistic and cultural knowledge.

We developed a platform for Asian WordNet (AWN) (Sornlertlamvanich et al. 2009) co-creation. It is prepared by connecting the existing bi-lingual dictionaries to the core Princeton WordNet (Fellbaum, 1998) based on the degree of English equivalent list (Charoenporn et al., 2008). WordNet is one of the most semantically rich English lexical banks and widely used as a resource in many aspects of research and development. Word knowledge can be logically represented by a set of synonyms called synset. Currently, there are 13 Asian languages semantically connected via Princeton WordNet, and the WordNet Management System (WNMS) is prepared for cross language access and additional term co-creation (Sornlertlamvanich et al., 2010).

Under the collaboration between Ministry of Culture and Ministry of Science and Technology by National Electronics and Computer Technology Center (NECTEC), a collection of cultural knowledge has been embodied since 2010. We did not start the work from scratch. Actually the work has been done some years ago by forming a set of servers individually operated by each province. Each province has to take care of their own contents about their responsible area. The initiative has been carried out for the purpose of creating a reference site of the local cultural knowledge. The distributed system's aim is to decentralize the management and to maintain the uniqueness of each specific area. However, there is a trade-off between the independent design and cost of maintenance that covers the service operation, interoperability and integrity. There are currently 77 provinces in Thailand, and each province is allocated an office for provincial cultural center. With the approach of the above-mentioned distributed system, it is too costly to maintain the service and the standard for data interchange.

The newly designed platform-based approach for digital cultural communication has been introduced. It is to build a co-creative relationship between the cultural institution and the community by using new media to produce audience-focused cultural interactive experience (Russo and Watkins 2005). First, we collected the existing provincial cultural knowledge and convert them to conform to a standardized set of metadata. This is to prepare the cultural knowledge for an open data schema and interoperability. The metadata is defined to follow the Dublin Core Metadata Element Set with some additional elements to fulfill the requirements for the information during the recording process. Second, we assign representatives from each province and train them to be a core cultural content development team for community co-creation. The contributed content needs to be approved by the core team before public visibility. Third, the cultural knowledge will be put on service to the audience of such scholars, who may be interested in the cultural practice, or business developers who may benefit from attaching the cultural knowledge to their products, or tourists who may seek for cultural tourism.

This cultural media assets will be linked and annotated by a governed conceptual scheme such as Asian WordNet (Sornlertlamvanich et al. 2009). The semantic annotated and linked data will be serviced as a fine-grained cultural knowledge for higher-level applications. The new media for recording the cultural knowledge is in the form of narrative, photo, video, animation, image incorporated with GPS data for visualization on the map.

Cultural Knowledge Co-Creation

The existing cultural data has been collected and cleaned up to conform to the designated standard metadata. The absent data are supposed to be revised and augmented by the experts from the Ministry of Culture. A few tens of thousands of records have been collected but most of them are captured in a coarse-grained image. Narratives and images are revised by a group of trained experts to create a seed of standardized annotated cultural knowledge base. Some new records have been added together with animation, video, panoramic photograph, etc. New technique in capturing the cultural image is aggressively introduced to create value added and gain more interest from the audience.

The standardized annotated cultural knowledge base is presented through a set of viewing utilities to the audience. Filter according to the location and province is prepared for customizing to page for each province. This is to allow the unique presentation of each province. The administration of each province will be responsible for its content correctness and coverage. Actually, the attractive presentation and narrative are required to attract the audience.

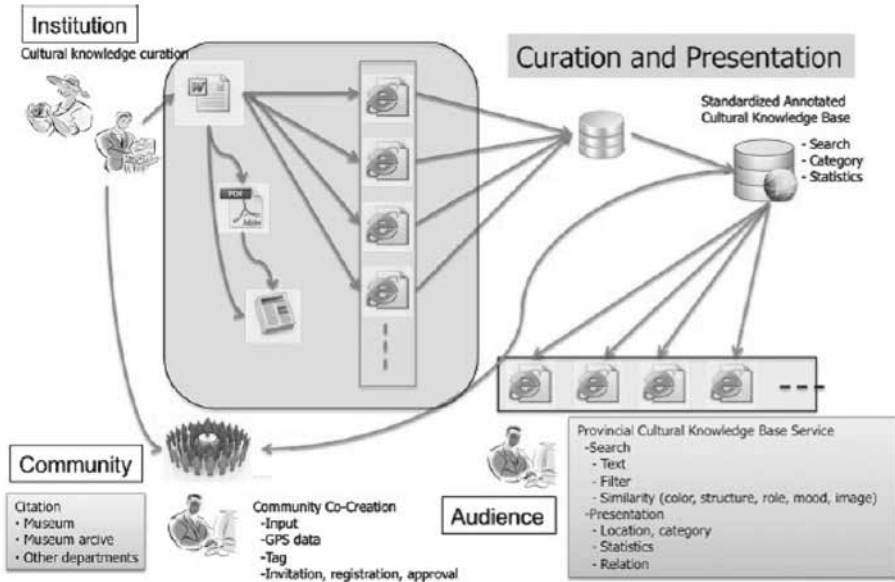


Figure 1. Community Co-Creation Cultural Knowledge Base

Social networking system is introduced to invite the participation from the communities. The institution representatives are actively encouraged to create their own community. The results from the community co-creation will keep the content maintained and clean up to compete with each other.

It is significant that the provided framework can encourage the data accumulation and fulfill the needs from the audience. Community co-creation will feedback the actual requirement that can improve the quality of the content. Institution plays an important role in mediating between community and the audience. As a result, the multiple types of content are generated on a designated standard. The annotated metadata can be used as a guideline for higher level of data manipulation such as semantic annotation, cross language and link analysis.

References

1. Charoenporn, T., Sornlertlamvanich, V., Mokarat, C., Isahara, H., Riza, H., and Jaimai, P. (2008). Evaluation of Synset Assignment to Bi-lingual Dictionary. Proceedings of the Fourth Global WordNet Conference (GWC2008), Szeged, Hungary.

-
2. Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, Mass.
 3. Russo, A. and Watkins, J. (2005). Digital Cultural Communication: Audience and Remediation. *Theorizing Digital Cultural Heritage eds. F. Cameron and S. Kenderdine*, MIT Press, Cambridge, Mass.
 4. Sornlertlamvanich, V., Charoenporn, T., Robkop, K., Mokarat, C. and Isahara, H. (2009). Review on Development of Asian WordNet. JAPIO 2009 Year Book, Japan Patent Information Organization, Tokyo, Japan.
 5. Sornlertlamvanich, V., Charoenporn, T., Isahara, H. (2010). Language Resource Management System for Asian WordNet Collaboration and Its Web Service Application. Proceedings of The Seventh International Conference on Language Resources and Evaluation (LREC), Mediterranean Conference Center (MCC).

Yunseok RHEE

Professor, Hankuk University of Foreign Studies

(Seoul, South Korea)

A Web-Based Participation Model to Preserve Linguistic and Cultural Diversity

1. Introduction

One of the main challenges in preserving linguistic and cultural diversity is to encourage small indigenous nations to use their own languages at all possible circumstances. In reality, however, such possibilities are getting rare even in cyberspace. Fortunately, recent advances in information and communication technologies give a new potential to help minor languages be more widely used and sustained.

By far the Web has been mostly used to share static information. With the advent of Web 2.0, however, its role is significantly extended to facilitate users' participation and collaboration beyond information sharing. Wikipedia, an Internet-based encyclopedia, might be a representative collaboration work enabled by such new features. Collective collaborations enable effectively filing up a huge volume of linguistic and cultural information, and sharing and reviewing the information-base among open communities. Web-based participations, in addition, support easy and efficient composition in information or service which might result in more valuable and newer outputs. In the end, they could open a totally new horizon differentiated from typical information sharing model.

In this report, we present a web-based participation model to support linguistic and cultural diversity, and discuss alternatives to design the model in some crucial aspects. The participation model includes key features to deal with linguistic and cultural assets: collection, review, and publication. It is designed to ultimately promote users' participation, and to embrace a diversity of linguistic and cultural information in terms of information characteristics and sources. Especially, such design consideration is required in the circumstance where diverse and heterogeneous devices including mobile systems are getting popular, and with parallel, the representation of linguistic and cultural data would be also diverse and complicated.

For the purpose, our model is designed to allow a self-defined data representation in which a user can provide his/her own data. It has no doubt that such data representations should be effectively handled by computers as

well as humans. Furthermore, the participation model plays as an open service platform which enables easy creation and launch of another new applications and services. Based on an open dictionary service, for instance, a phrase translation service or another dictionary with more specific interests can be built by others than original authors. In summary, such a participation model is expected to provide a ground to wider sharing and use their linguistic skills of the small ethnic groups, and further efficiently build up new challenging works as well as glue to naturally linked individual's thoughts.

2. Barriers to Minor Languages in Cyberspace

As we know, earlier web services had focused on how to publish and share their information which is mostly static and unchanged over time. But, since the advent of web 2.0, it has become a true communication space far beyond information sharing. Such new web features support collective collaboration, active social networking, and component integration. Collective collaboration is achieved by participatory creation and sharing.

Active social networking enables users to dynamically interact to each other, based on personal profiles, their social links, and additional services. All the social media including blogs, twitters are enabled by this feature. Component integration combines data, presentation, and functionalities from multiple sources to create a new service or application. For example, Google opens their own map service as public APIs which other developers can freely use.

In such circumstances, we are facing two barriers in supporting minority languages in cyberspace. First, sometimes their computing environments would be less developed. For instance, we need operating system supports with character codes, fonts, and input methods. It is also true for language supports in some major applications or services from Microsoft and Google. Also, the gap in leading technologies is getting larger.

Secondly, the most important problem I think is the less participation in using a minority language in cyberspace. Especially, the cyberspace is not familiar with agro-typed rural residents and senior citizens who have much experience in their languages. In addition, younger generation prefers using official languages which are strongly recommended and advantageous in their community.

Of the barriers just mentioned, the eager participation into the cyber community is the biggest challenge we have to address. In essence, it is not a problem linked to technology only, rather it is strongly related to social and cultural surroundings.

3. Approaches for Web-Based Language Sustainability

In terms of technology, we have to support a communication model which satisfies the following requirements. Firstly, the model should encourage language users within a minority community. We expect that it will reinforce the solidarity among the indigenous people, and make a positive cycle to use their languages. But, we have to be very careful as a closed society is subject to threaten creativity or diversity in communication, and in the end, it may lose motivation to participate.

So, the communication model should be able to bridge their local community and global world. Such mutual exchange could facilitate their linguistic, cultural, and creative activities. The Korean pop entertainment is a good example which has made huge achievements by such exchange.

As shown in Figure 1, we may have two separate systems for local and global communities. Also, the local systems may be proprietary and incompatible with the global systems. In that case, those could work for local community, but cannot interact with the global systems.

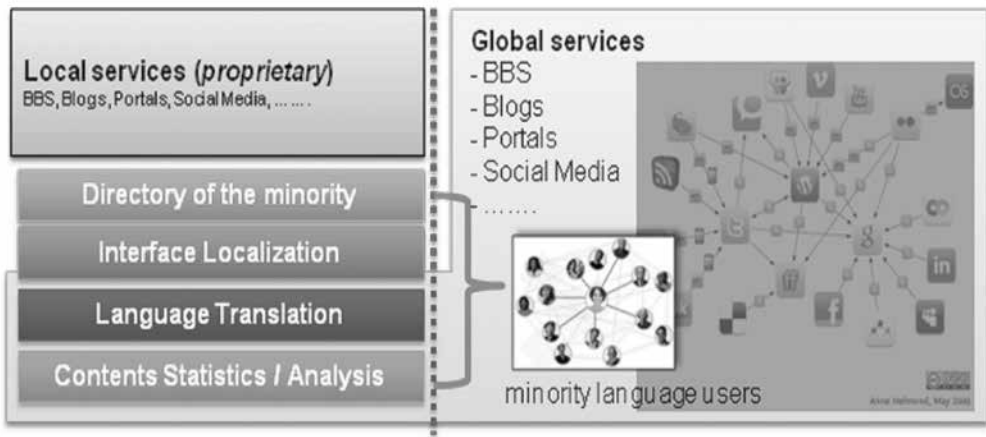


Fig. 1. Web-based Minority Language Service Components

By any means, we have to develop how to exploit global systems for the both. As an approach to support such a model, we need to design a gateway service between local and global communities on the Web. It consists of 4 components, such as directory of the minority users, interface localization, language translation, and contents analysis. We emphasize that the main purpose of this approach is to efficiently unite minority people while helping them play in the global space.

Looking closer, the first one, directory service helps the indigenous people find out their colleagues or relevant contents with ease. It needs to maintain a directory for each web activity such as blogging and social media. Eventually, it acts as a contact point for someone to get into their local community.

Two screen shots of Figure 2 show a simple case which is developed by Dr. Kevin Scannell at St. Louise University in America. The site is named *indigenoustwitter.com*. The left one shows a list of indigenous languages he found from the Twitter. And, the right one shows a list of active users in a specific language. So, you can easily find who stars in your language Twitter, and get into your community.

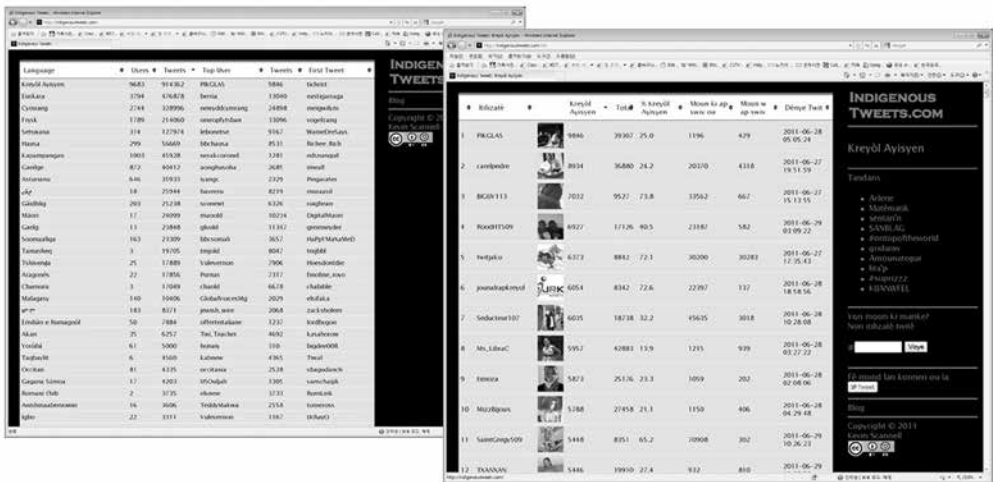


Fig. 2. Web-based Minority Language Support Example: *indigenoustwitter.com*.

Next, we need to localize some major applications such as Google, YouTube, and Twitter. Mostly, menus, popup messages, and on-line help are major targets to localize. For instance, Google is running the *Google Technology User Group (GTUG)* which officially supports language problem of each country. In another way, we can personally contribute to the localization through the Google In Your Language service. For example, we can put into the system a set of localized messages. The language translation is strongly required to interact with global community. It should automatically translate user's texts into a common language such as Russian or English.

In case of Google, they provide a general translation service for about 50 languages. They also provide open translation APIs which can be employed by another application. In the above mentioned example users' tweets are translated in Twitter by the Google APIs (it is also a typical example of component

integration mentioned prior). For a language they do not support, you should develop a translation module which can be plugged into a Web browser. It is not easy, but it is worth it. Even though translation accuracy is not quite good, it is strongly required to link the indigenous community into global community.

Lastly, we need to collect and analyze statistics on the web content. It can be conducted by crawling web data, and after analyzing, it can be used to rebuild a directory of minority users. It can be used to build a corpus and linguistic resources for minority languages including multimedia. Crubadan is an example module designed for this purpose by Dr. Scannell. It acts as a web crawler which gets a small set of seed texts, and forwards a query including these seeds to Google. Then, Google returns a list of documents potentially written in the same language. It is also performed by Google open APIs for web search. You may refer to this URL, <http://borel.slu.edu/crubadan/index.html>.

4. Conclusion

We need to provide a communication model which can support solidarity within local community as well as mutual exchange with global community. At present, web-based participation is definitely a feasible option option, and gateway service with following features may be considerable. I would like to finish with this saying by Nancy Hornberger from the University Of Pennsylvania, "Language revitalization is not about bringing a language back, but bringing it forward." I believe that our efforts and trials for minority languages will strengthen their future.

Tunde ADEGBOLA

*Director, African Languages Technology Initiative
(Lagos, Nigeria)*

Multilingualism, Multimedia and Orature in the Information Age: Using Multimedia to Include Oral Cultures in Cyberspace

One of the advantages of digital technology is that it facilitates access to information in various forms and modes. This could be in the form of sounds such as speech or music and in the visual forms such as written texts as well as static images and motion pictures. The presentation of information in these various forms is what we now refer to collectively as multimedia. The relevance of digital technology to modern use of multimedia stems from the fact that the digital approach to the representation of natural reality offers uniformity in the storage, processing and retrieval of information on the same hardware by the use of appropriate software. This leads to economy of scale and many other benefits.

The development of writing is a great mile stone in human civilization. Writing as a means of documenting human thought offers efficiency, portability and permanence. Writing is efficient in the sense that large volumes of information can be stored within relatively small volumes of media space, hence offering great advantages in the storage space of written ideas. It is portable in the sense that written texts can be moved to locations other than the one occupied by the producer of the idea and it can be at more than one location at the same time. Finally, written texts are said to be permanent in the sense that spoken words, fizzling into thin air after they have been uttered, when written, remain available for consultation for as long as the media on which they are written survive.

Writing has enhanced the production and management of information and knowledge, providing means for productive storage, retrieval, transfer and dissemination. From the humble beginnings of the documentation of human experiences in cave drawings, totem poles and other semiotic endeavors to the precursors of the present information revolution facilitated by Gutenberg's printing technology, mankind has benefited immensely from the improved capacity for the storage, retrieval, transfer and dissemination of information and knowledge.

Unfortunately however, not all the languages spoken in the world today have been reduced to writing. Furthermore, writing present a steep learning curve particularly for adult learners who did not get the opportunity to learn to read and write as children. Hence there are many segments of our so-called global information society in which a significant number of people still do not have the

capacity to read and write. Worse still, there are whole communities that speak languages that are still unwritten languages. People in these sorts of situations are disadvantaged as they are pushed to the fringes of the information society. Their inability to read and write excludes them from many vital aspects of the life of the wider communities they live in. Consequently therefore, such people are systematically excluded from active participation in the development processes of both their local and global communities and their communities are the worst for it since these people cannot fully contribute their own required quota to development processes.

To alleviate this exclusionary condition, multilateral agencies such as UNESCO as well as national governments and various NGOs put a lot of efforts into improving literacy levels in various communities around the world. They organize programs in developing writing systems for languages that remain yet unwritten and adult literacy programs for adults who live within literate cultures but did not manage to acquire literacy skill as children. Despite the best of these efforts however, there are still a lot of illiterate people in our world of the information age. It may be quite depressing but we must admit that it would appear that the death of illiterates is still one of the primary ways by which we are able to improve literacy in the world today.

Despite the efficiency, portability and permanence of written text as discussed above, it must be noted that speech still remains the preferred mode of human communication. It is the most natural way by which humans communicate. This is easily demonstrated by the lengths to which we as humans still go in organizing conferences, seminars and workshops to sit together and consider issues of paramount importance based on the use of speech. These we still do despite the near ubiquity of written text. Given this reality, we need to pause and ask a pertinent question: if sound recording had been developed before writing, would writing still have developed in the same direction? Would it have still grown to acquire the importance it now has in our global society of today? Yet, our present bias for writing seems to manifest as a distraction in the ways we treat other media and modes of communication. The historical exploits of writing and the vibrant cultures and industries that have grown around it tend to ascribe an overrated importance to writing. For this reason, cyberspace has so far been developed based on the importance we attached to writing. This direction of growth of cyberspace is not inevitable and there are viable alternatives to the heavy demand of reading and writing that cyberspace as we now know it presents.

Cultures using languages that yet do not have writing systems and many more that are shaped strongly by oral forms of information and knowledge sharing would not benefit maximally from cyberspace if we do not provide alternatives

to the heavy dependence on reading and writing as we have now grown cyberspace. In such cultures, teaching and learning are still based primarily on memorization and recitation. People share knowledge by telling stories, proverbs, riddles, etc. In many of such cultures, the collective memory of society resides with Griots who memorize the history and tell it when required. Hence, the history is accessed primarily by performance. In some other cultures the collective memory is held within society at large based on an elaborate system of cognomens in which people are named and described according to the lives and times of their forbearers.

These cognomens are usually stories of both valiance and villainy. Among the Zulu of southern Africa it is called *Izithakazelo* and among the Yoruba of West Africa it is called *Oriki*. In the Yoruba culture for example, children get some parts of their *oriki* recited to them at least once a day. Parents and other elders in the homestead, particularly mothers would go into a session of recitation of a child's *oriki* in response to a simple good morning greeting from the child. The *oriki* is also freely recited, both to children and adults during ceremonies and sometimes, just in acknowledgement of important achievements or in a bid to encourage prosperity. This way, various portions of the collective memory of the whole community is held in the brains of individual members of the community and is regularly rehearsed in performance for the purposes of retrieval whenever needed.

Despite the known weaknesses of the human brain as a store of information, the sacred texts of the *Ifa* divination system of the Yoruba still remains largely in oral form. Even though some portions have been written, the bulk of it still resides mainly in the brains of *Ifa* scholars as an oral scripture. These sacred texts contain the knowledge of Yoruba philosophy, medicine and many other relevant sciences in elaborate poetry. The knowledge contained in these poems is organized in an equally elaborate system of information look-up based on binary mathematics and probability theory.

How then do we accommodate such traditions that are based mainly on orature in cyberspace? How do we use multimedia to assist such cultures with documenting their histories and their knowledge of their environment in media that are more appropriate than the human brain? If we are to develop cyberspace as a truly multilingual knowledge space, is it pertinent to ensure that cultures that still learn by memorization and recitation also have access to the information superhighway? Apart from giving such cultures access to cyberspace, it is also necessary to offer modern technology as a means of documenting and mobilizing the knowledge that is otherwise held in the brains of mere mortals.

To accommodate orature in cyberspace we need to consciously and deliberately reassess the role of multimedia as a means of information and knowledge sharing. Multimedia allows for the documentation of human thought, ideas and knowledge without literacy. It does not necessarily require the development of a system of orthography and therefore does not present the steep learning curve that literacy presents to adult illiterates. Modern digital technology has facilitated multimedia in unprecedented ways. By virtue of this development, it is now possible to document information in different forms and in various modes, be it in the form of sounds in speech and music or in the form of images in writing, still images and motion picture. Information in various forms and modes can now be stored and retrieved uniformly in various media on the same hardware. This diversified access to information through multimedia should be better exploited to include cultures that are still based primarily on orality in cyberspace. So far, multimedia has been used primarily as enhancement to written texts in cyberspace. Even though such use is welcomed, multimedia needs to be seen as valid and productive in its own right and should therefore not be used merely to enhance written texts.

Literacy is defined as the ability to read and write. It must be noted, however, that the value of literacy is not really in the process of reading and writing but in the results we get from reading and writing; the sharing of information and knowledge. Hence, literacy is valuable only because it provides an efficient means for the documentation and reproduction of human thoughts, ideas and knowledge. If it is possible to achieve the same results without reading and writing, the value of the process of reading and writing becomes diminished and the result of reading and writing takes the full value. Multimedia offers the capacity to document human thought, ideas and knowledge beyond reading and writing and therefore offers the possibility of redefining literacy, changing it from 'the ability to read and write' to 'the capacity to engage literature'. Replacing the notion of ability with that of capacity moves the definition of literacy away from excluding those that are not able to including those that can be capacitated.

The fight against illiteracy is a major preoccupation the world over, but as was observed earlier, death of illiterates still seems to be one of the most effective tools used in this fight. This is not to say that illiterates are deliberately killed in the fight, but the intrinsic link between illiteracy and ignorance weakens the chances of survival of an illiterate person, manifesting as a vicious cycle in which illiteracy breeds poverty and poverty weakens a person's capacity to become literate.

Modern digital information communication technology can be used to break this vicious cycle by introducing multimedia as a means of productive

information sharing without the ability to read and write. Such intervention is bound to lend a hand to the fight against illiteracy by producing e-literates out of illiterates. We may therefore be able to redefine literacy, changing its definition from 'the ability to read and write' to 'the capacity to engage literature'.

The use of multimedia in the fight against illiteracy should be approached at two main levels. The first level involves the basic application of multimedia for documentation. This entails the use of audio recordings to document spoken information as well as the use of static images and motion picture to provide visual complements to such audio recordings. At a more advanced level we can use speech technologies such as Automatic Speech recognition (ASR) and Text-to-speech (TTS) Synthesis to turn speech into written text and written text in to speech respectively. This way we can derive all the advantages of written texts even without the ability to read and write.

The use of hand-held devices such as Mobile Phones and Tablets is widening access to multimedia-enabled devices. Unfortunately however, the applications built for these devices still reflect a clear bias for written texts. As the popularity of these hand-held devices continues to grow, there is a need to reassess this bias for written texts in order to empower that large population of people who still live in oral societies to benefit from the cyberspace. To grow a truly multi-lingual cyberspace, we need to carefully reevaluate the importance of writing within the context of new developments in digital technologies and thereby make deliberate efforts to elevate the status of multimedia. We should not continue to overrate the importance of writing while there are other information media and modes that offer advantage we cannot afford to overlook.

Nicholas THIEBERGER

*Senior Research Fellow, University of Melbourne
(Melbourne, Australia)*

Multilingualism in Cyberspace – Longevity for Documentation of Small Languages³³

1. Introduction

Supporting small languages can take many forms. A key to long term access to information about most of the world's languages is in the curation of existing records and the proper creation of records now. The network of language archives that exist in the world have been developing standards and have also been training practitioners (linguists, speakers or language workers) in good methods for language documentation. Websites can deliver accessible information, but the risk is that unique records will be placed only into websites and will not survive in the longer term. Archival forms of the records should be properly described using standard metadata terms and be created and stored at the highest possible quality, for later delivery in compressed formats suitable to web delivery.

In this paper I outline the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC) (<http://paradisec.org.au>) as an example of a new kind of archive that is emerging, one that is not only a repository of curated material, but one that is involved in training, adopting standard formats for primary records and creating workflows that will result in multiple outputs from linguistic fieldwork. We have also developed a method for presenting interlinear text and media online (<http://www.eopas.org>) in order to encourage the creation of language records in reusable formats and to work towards a language museum in which samples of language in performance can be viewed on the internet.

I suggest that we need to provide a service of advice and data conversion for those for whom it is simply too difficult to do this work themselves. An example of such a service is the Resource Network for Linguistic Diversity (<http://www.rnld.org>).

³³ *Work reported on in this paper was supported by the following Australian Research Council grants: SR0566965 – Sharing access and analytical tools for ethnographic digital media using high speed networks; DP0450342 – New methodologies for representing and accessing resources on endangered languages: a case study from South Efate.*

It should always have been part of the discipline to produce good research data, but the use of digital recorders, storage and archives, together with the development of suitable standards for data and metadata construction, have all combined to refocus our efforts in this direction.

At the turn of this century, a group of Australian linguistic and musicological researchers recognised that a number of small collections of unique and often irreplaceable field recordings mainly from the Southeast Asian and Pacific regions were not being properly housed and that there was no institution in Australia which would take responsibility for them. The recordings were not held in appropriate conditions and so were deteriorating and in need of digitisation. Further, there was no catalogue of their contents or their location so their existence was only known to a few people, typically colleagues of the collector. These researchers designed the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC³⁴), a digital archive based on internationally accepted standards (DC/OAI-PMH metadata, IASA audio standards and so on) and obtained Australian Research Council Infrastructure funding to develop an audio digitisation suite in 2003.

2. Background

Researchers working with speakers of small languages (those with few speakers) typically conduct fieldwork to learn how aspects of these societies function, how the languages are structured, or how musicological knowledge is constituted, in addition to recording life stories, ethnobiological and other information. Typically these are minority endangered languages for which no prior documentation exists. This is vitally important work which often records language structures and knowledge of the culture and physical environment that would otherwise be lost (see e.g., Evans 2010, Maffi 2001, Harrison 2007). However, while it is typical for the interpretation and analysis of this data to be published, the raw data is rarely made available. The data – tapes, field notes, photographs, and video – are often not properly described, catalogued, or made accessible, especially in the absence of a dedicated repository. This means that enormous amounts of data – often the only information we have on disappearing languages – remain inaccessible both to the language community itself, and to ongoing linguistic research.

The data that we create as part of our research endeavour should be reusable, both by ourselves and by others. First because any claims that we make based on that data must themselves be replicable and testable by others, and second, because the effort of creating the data should not be duplicated later by

³⁴ <http://paradisec.org.au>.

others, but be used as a foundation that can be built on (cf. Thieberger 2009). In order to be made accessible, the data recorded by researchers must be properly collated and indexed for public presentation and archiving (see Austin 2006, Himmelfmann 1998, Woodbury 1998, 2003). However, until recently there has been no simple means for doing this and access to physical analog records can be difficult, if not virtually impossible, when they are stored in a single location.

This issue is being faced by scholars in many disciplines and is being addressed under the rubric of cyber infrastructure or e-humanities – how to build on existing knowledge and how to add new data that is being created in the course of various research projects so that the broader research community can benefit from it. This is all the more important when a linguist makes the only recordings for an endangered language—one that may no longer be spoken in the near future. Australia and its immediate neighbours are home to a third of the world's languages, most of which may never be recorded.

3. Implementation of PARADISEC

In the initial phase of the PARADISEC project we established a steering committee with representatives of each of the partner universities (initially the University of Sydney, the University of Melbourne, and the Australian National University (ANU)). The director of the project is Linda Barwick at the University of Sydney.

With invaluable technical support from both the National Library of Australia and the National Film and Sound Archive and with funds from the Australian Research Council we bought a Quadriga digitisation suite and employed an audio engineer and administrative assistant, based at the University of Sydney. We also built a vacuum chamber and low-temperature oven to allow us to treat mouldy tapes that required special care before being playable. Tapes stored at the ANU were identified and located and then permission was sought from the collectors or their agents to digitise and accession them into the collection.

In the first year of funding we had to come up with outcomes that would justify further funding grants and we aimed for 500 hours of digitised tapes in that first year (we achieved this goal in ten months). We wrote a catalog database in Filemaker Pro, aware that it would provide us with an immediately usable tool that would ultimately have to be converted to an online database. This database allowed us to refine data entry forms and controlled vocabularies without relying on a programmer. This first catalogue worked well and exported to the XML files required for inclusion as headers in Broadcast Wave

Format (BWF) files, and also exported to a static repository for Open Archives Initiative harvesting via the Open Language Archives Community³⁵ harvester.

Files generated by this system (at 96khz/24 bit) are large, around 1.5 Gb per 45-minute side of a cassette, and so require dedicated storage facilities. We established a tape backup system which ran periodically to copy files from the hard disk to storage tapes, but were fortunate when the National Computational Infrastructure³⁶ (NCI) designated PARADISEC a 'Project of National Significance', allowing us to use their mass data storage system, with considerable storage space provided to support our work. They further provided programming support by writing specialized software (called 'Babble') which provides weekly, monthly and quarterly reports on the state of the collection, as well as nightly querying the server in Sydney and copying files that are ready for archiving.

Data is organized by collector, but also by the internal logic of the collections (the same collector working on two different languages will have two collections, or a collection of video may be distinct from a collection of still images). The collection-level also speeds up a user's typing into the catalog as common fields from the collection level can be inherited down to the item level. Our naming convention is rather simple³⁷ ('CollectionID'-'ItemID'-'FileID'.'extension') and it also provides the hierarchical file structure into which files are placed and stored on the server (with directories corresponding to the collection level and subdirectories corresponding to the item level). Subsequently and with funds from the ARC, we built digitisation suites in Melbourne and Canberra, allowing us to preserve important heritage tape collections.

The primary aim of the project to date has been on preservation of unique cultural records. Including a licence, or information about how each item can be used, is critical to the establishment of a properly curated collection because without it there is no way of providing access. Each depositor must fill out a deposit form specifying any conditions that may apply to the material. We provide a default set of access conditions which any user must agree to prior to being given access to data, and depositors can choose to allow this set of conditions to govern their collection, or to determine their own conditions. We are presently investigating the use of Creative Commons³⁸ licences as a less restrictive and more standardised form of agreement.

³⁵ <http://www.language-archives.org/>.

³⁶ <http://nci.org.au/>.

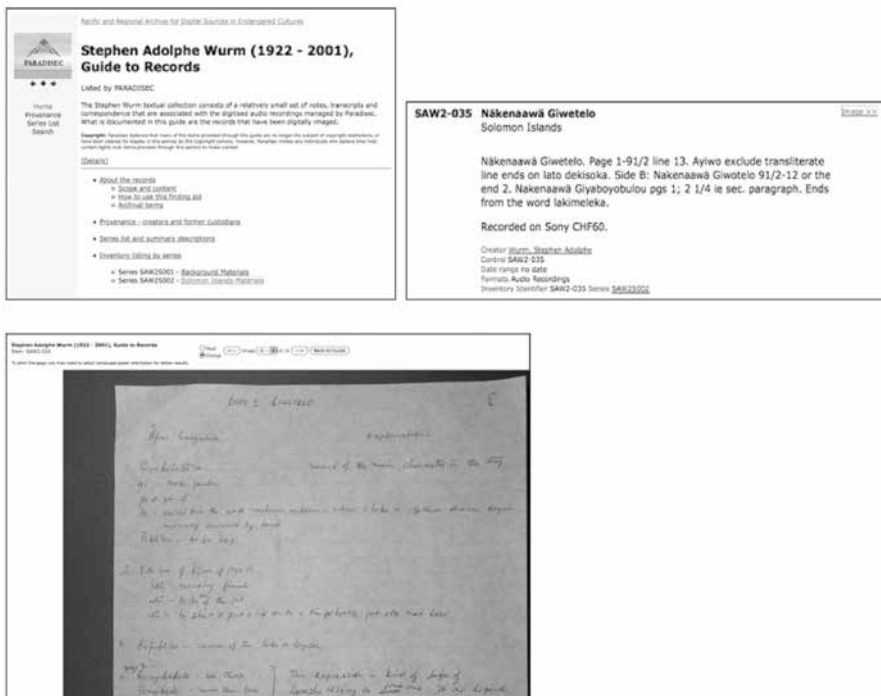
³⁷ See <http://www.paradisec.org.au/info.html> for more on *filenaming*.

³⁸ <http://creativecommons.org/>.

4. Delivery of Archival Material, Page Images and Dynamic Media

We provide material from the collection to those authorized to receive it, typically in the form of downloadable files, however we have also worked on specific methods for the online delivery of two kinds of material – page images and time-coded media. We made available images of 14,000 pages of fieldnotes (see Figure 1) from three deceased researchers using the Heritage Document Management System³⁹ with a digital camera rig that we took to the home of the estate's executor, or to the office in which the papers were stored. These notes from deceased researchers would otherwise have only been available in a single physical location. As we do not have the resources to keyboard all of these manuscripts the images are stored in the collection with sufficient contextual metadata to make them discoverable on the web. As noted earlier, the archival version of each image is stored separately from the representational version.

Figure 1. Page images from the Wurm collection of online manuscripts showing finding aids from the highest level (top left), to the item level (top right) and finally the image itself (bottom) (<http://paradisec.org.au/fieldnotes/SAW2/SAW2.htm>)



³⁹ <http://www.esrc.unimelb.edu.au/hdms/>.

5. The EthnoER Online Presentation and Annotation System (EOPAS)

While building a method for working with our own data we consider it important to create generalisable models and structures for others to use, and to engage in discussions and training sessions both in order to refine our methodologies and to impart new ideas. An example of such development is our work on the online presentation of interlinear glossed text together with recorded media (EOPAS⁴⁰), allowing material from any language to be heard in concert with its transcript and translation (Schroeter and Thieberger 2006). A number of tools for annotating language data have been produced recently⁴¹ and it is clear that more are envisaged now that several large projects are engaging with these issues in the USA, UK⁴², Germany and the Netherlands⁴³. Annotation is a basic task that is undertaken following recording, and can take several forms, the most common of which, for linguists, is interlinear text. These texts are analysed and parsed by a glossing tool that produces parallel lines of text, word translation and grammatical information, together with a free translation. These texts are then the input into EOPAS, a schema-based XML system for making explicit the relationship between parts of interlinear texts together with links to the source media, streamed using HTML5 (see Figure 2) which allows searching and concordancing linked directly to the media. EOPAS is portable (the source code is freely available⁴⁴), allowing other initiatives to capitalise on the work and potentially develop it in different directions.

The ultimate aim of this approach is to allow new perspectives on the data itself, provided by contextualised access to primary data, and then to allow new research questions to be asked, and richer answers to be provided, all in a fraction of the time that it would have taken with analogue data. the image itself (bottom) (<http://paradisec.org.au/fieldnotes/SAW2/SAW2.htm>).

Figure 2. Example of a video clip with time-aligned text as presented in EOPAS



⁴⁰ <http://www.linguistics.unimelb.edu.au/research/projects/eopas.html>.

⁴¹ Tools for annotating linguistic data are listed here: <http://www ldc.upenn.edu/annotation/>.

⁴² <http://www.eldp.soas.ac.uk/>.

⁴³ The DoBeS (Dokumentation Bedrohter Sprachen /) project (<http://www.mpi.nl/DOBES/>).

⁴⁴ <https://github.com/eopas/>.

6. Current Status of the PARADISEC Collection

Currently (mid-2011) PARADISEC contains 7,220 items made up of 48,555 files totaling 5.3 TB, with just over 3,020 hours of audio data⁴⁵. Digital video already makes up an increasingly significant part of the collection. We hold data representing 676 languages from 60 countries (see examples of the kinds of collections in Table 1) which is organized into 260 collections, some 85 of which represent new fieldworkers who have deposited material on their return from fieldwork (and one during the course of her fieldwork), thus providing a citable form of their data for their own research. Here are some examples of collections digitised and described by the PARADISEC project:

- Arthur Capell: 1950s Pacific and PNG (114 tapes and 30 archive boxes of fieldnotes)
- Tom Dutton: 1960s onwards, PNG, 295 tapes
- William Foley: 1970s, PNG, 34 tapes
- John Harris: 1960s, Kiwai, PNG, 75 tapes
- Don Laycock: 1960s, PNG, 98 tapes
- Al Schütz: 1960s onwards, Vanuatu, six tapes
- Stephen Wurm: 1970s Solomon Islands tapes (~120 tapes and transcripts/ fieldnotes)
- Bert Voorhoeve: West Papua, 180 tapes

The remaining collections are digitised from recordings made since the 1950s. The provision of this service requires ongoing support and negotiation with depositors and we have found that a key to establishing the collection has been the depositors' perception of the benefit accruing to them and to their data in having it well described. In addition, there are collections we know about and would dearly love to digitise but we do not have the resources to do this work. These include large audiotape collections at radio stations around the Pacific, many in local languages, and collections in regional cultural centres that do not have any local equipment to digitise their collections. Further, we are regularly approached by former colonial patrol officers or missionaries who have recordings, notes or photographs that they want to preserve.

We have published on our website⁴⁶ a detailed description of our workflow, developed over seven years of operation, that describes the various processes involved in locating tapes and then assessing, accessioning, digitising and

⁴⁵ The collection catalog can be searched at <http://paradisec.org.au/catalog>, or at <http://www.language-archives.org/search?q=&a=paradisec.org.au>.

⁴⁶ <http://www.paradisec.org.au/pdsc-manual10.pdf>.

describing them, managing the resulting data and metadata, and the return of original tapes. PARADISEC has been cited as an exemplary system for audiovisual archiving using digital mass storage systems by the International Association of Sound and Audiovisual Archives⁴⁷ and, in 2008, won the Victorian Eresearch Strategic Initiative prize for humanities eresearch.

Once we built the infrastructure for a research repository, including the catalog, file system and naming conventions, it has been taken up by those researchers who are aware of the need to describe and preserve their research material. Often it is only in the process of depositing with PARADISEC that a collection is first described in a systematic way – one that then allows the description to be searched by Open Archives Initiative search engines (and also google). Every eight hours the PARADISEC catalog is queried by a service run by the Open Language Archives Community (OLAC) and any new or edited catalog entries are copied and made available to their aggregated search mechanism. Similarly, because the catalog complies with relevant standards, the Australian National Data Service (ANDS) has been able to incorporate our collections into its national search mechanism. The quality of the metadata we provide ensures that targeted searches by language name can be resolved without locating similar but irrelevant forms.

7. Regional Links and Training

While the initial focus for our collection was the region around Australia (as suggested by the name we chose at the outset of the project), it has become clear that we need to accept material that has no other place to be archived. Typically, this means supporting Australian researchers whose research is outside of Australia, with the geographic spread of material we house now extending from India, into China, and across to Rapanui (Easter Island). With limited resources PARADISEC has nevertheless established working relationships with cultural centres in the Pacific region (e.g., the Vanuatu Kaljoral Senta, or the Institute of PNG Studies) which have involved providing CD copies of relevant material and, in the case of the University of New Caledonia, cleaning and digitising old reel-to-reel tapes in Drehu. A serious concern for many such agencies in the region is the lack of continuity in funding and in staffing, with the potential result that collections established and curated over time may be at risk. We would like to be able to digitise the many hours of tapes held, often in less than ideal conditions, in countries of the region. We have begun

⁴⁷ *Paragraphs 6.2.2.11 and 6.2.2.12 of the International Association of Sound and Audiovisual Archives (IASA) Guidelines on the Production and Preservation of Digital Audio Objects (IASA-TC04)*. Aarhus, Denmark: International Association of Sound and Audiovisual Archives (IASA), 2004, p. 51. *The Sub Committee on Technology of the Memory of the World Programme of UNESCO recommends these guidelines as best practice for Audio-Visual Archives.*

an occasional mass backup of significant collections of digital material from the Vanuatu Kaljoral Senta and would like to extend this as a service to other agencies.

We regularly offer training workshops in linguistic research methods, including the use of appropriate tools and recording methods and in data management for ethnographic field material. This is extremely important, as the more informed the research community can become about the need for reuse of primary data, the more likely they are to be creating well-formed data that needs no extra handling by PARADISEC to be accessioned into the collection. Such training has been offered at community language centres as well as in academic settings.

We cooperate in two further initiatives for disseminating information. The first is a blog (Endangered Languages and Cultures⁴⁸) and the second a resource website with FAQs and a mailing list (the Resource Network for Linguistic Diversity⁴⁹). Because of the rapid changes in methods for recording, transcribing, and analysing human performance no one can keep completely up to date, so these web-based resources are widely quoted and appreciated by the community of researchers.

8. Conclusion

PARADISEC is a practice-based archive, arising from a community of practice who recognised that it was part of our professional responsibility to ensure that the records we create are properly curated into the future. This is a new conception of a data repository, built into workflows and research methods of particular disciplines, but recognising the need to adhere to broader international standards. It is unique in its links on the one hand to fieldworkers and to speakers of Indigenous languages and on the other hand to the cutting-edge technologies of Web 2.0 and HTML5.

References

1. Austin, Peter. 2006. Data and language documentation. In Gippert, Jost, Nikolaus P. Himmelmann, and Ulrike Mosel (eds) *Essentials of Language Documentation*, 87-112. Mouton De Gruyter, Berlin.
2. Evans, Nicholas. 2010. *Dying Words: Endangered Languages and What They Have to Tell Us*. Maldon, MA: Wiley-Blackwell.

⁴⁸ <http://paradisec.org.au/blog>.

⁴⁹ <http://www.rnld.org>.

-
3. Harrison, K. David. 2007. *When languages die: The extinction of the world's languages and the erosion of human knowledge*. Oxford University Press, New York.
 4. Himmelmann, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36: 161-195.
 5. Maffi, Luisa (ed.) 2001. *On Biocultural Diversity: Linking Language, Knowledge and the Environment*. Smithsonian Institution, Washington, D.C.
 6. Schroeter, Ronald and Nicholas Thieberger. 2006. EOPAS, the EthnoER online representation of interlinear text. In Barwick, Linda and Nicholas Thieberger (eds) *Sustainable Data from Digital Fieldwork*. 99-124. Sydney University Press, Sydney.
 7. Thieberger, Nicholas. 2009. Steps toward a grammar embedded in data. In Epps, Patricia and Alexandre Arkhipov. (eds.) *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, 389-408. Berlin ; New York, NY : Mouton de Gruyter Mouton. <http://repository.unimelb.edu.au/10187/4864>
 8. Woodbury, Anthony. 1998. Documenting rhetorical, aesthetic, and expressive loss in language shift. In: Grenoble, L.A. and L. J. Whaley (eds.) *Endangered languages : language loss and community response*, 234-258. CUP, Cambridge.
 9. Woodbury, Anthony. 2003. Defining Documentary Linguistics, address given at the Annual Meeting of the Linguistic Society of America, Atlanta, Georgia, on January 3, 2003.

Dietrich SCHÜLLER

*Consultant, Vienna Phonogrammarchiv;
Vice Chair, Intergovernmental Council,
UNESCO Information for All Programme
(Vienna, Austria)*

Audio and Video Documents at Risk: Safeguarding the Documents Proper of Linguistic Diversity and Orally Transmitted Cultures⁵⁰

It is almost trivial to state that written text documents are inadequate tools to represent acoustical phenomena such as spoken language, dialect, and music, or optical manifestations of rituals, dances, etc. This applies specifically for orally transmitted cultures where no traditional relations are in place between texts and spoken language, or traditional forms of notations, e.g. for music and/or dance. Verbal descriptions and written texts are insufficient and subjective, and this makes historical studies in ethnolinguistics, ethnomusicology, and social anthropology at large, cumbersome and rudimentary.

However, this situation has changed with advent of audiovisual documentation technology in the 19th century: photography was available since 1839, the phonograph was invented in 1877 and cinematography emerged in the 1880s and 1890s.

The development of the phonograph was specifically associated with the scientific interest to understand the physics and physiology of human speech. Consequently, the phonograph has attracted linguists and anthropologists immediately since its practical availability in 1889/90. Systematic language and music recordings were at the cradle of emerging disciplines like phonetics, ethnolinguistics and dialectology, as well as ethnomusicology. Their histories are closely associated with the history of sound recording. Consequently, this led to the systematic establishment of sound archives, namely so-called “Phonogram Archives”, the first in Vienna in 1899, followed by Berlin in 1900, and in 1908 by St. Petersburg.

⁵⁰ *This paper is based on an earlier study (Schüller 2008) carried out within EU-funded project TAPE - Training for Audiovisual Preservation in Europe, 2004-2008. The general situation of audiovisual documents, in which the situation of research materials is embedded, has recently been discussed at the International Conference Preservation of Digital Information in the Information Society: Problems and Prospects, Moscow, October 2011 (Schüller, 2012).*

Cinematography was also introduced to anthropological fieldwork, but, because of the technical complexity and the costs of this technology, not systematically employed to considerable extent.

Over the following decades the phonograph remained in use until the 1940s, as gramophone recording, though technically superior, was hardly applicable under field conditions. Audio field recording became only widespread with advent of battery operated tape recorders in the mid-1950s, which permitted uninterrupted recording of considerable lengths and quality everywhere in the world. Similarly, video recording became popular with the advent of “handy cams” since the 1980s. Though technically inferior to cinematography, even to 8mm amateur film, this was affordable even for private researchers which made “videotaping” a widespread documentary tool in fieldwork. More recently video documents became an important factor even in linguistics, to permit research into gestures and mimics.

As a result of this technological development audiovisual collections mushroomed and became irreplaceable stocks of primary sources of linguistic and cultural diversity of human kind. This mushrooming was supported by the relative affordability of audio and video recording equipment, which led to the establishment of collections as part of research institutions, museums, and even in the possession of private researchers. Generally, however, these collections remained without specific custodial infrastructure, or any preservation strategy, let alone budgetary provisions. It is estimated that the greater part, possibly 80% of these primary sources, which are the basis of our present knowledge in many disciplines, are outside archival custody in the narrower sense. Only 20% of this heritage are professionally preserved.

This system of relative anarchy has worked until recently somehow.

However, audio and video recordings are prone to deterioration and threatened by format obsolescence. More dangerous than the instability of carriers is meanwhile the inescapable unavailability of replay equipment, as traditional analogue, and also single carrier based digital formats, became obsolete. As a consequence, the industry ceases production of equipment, spare parts and provision of service.

This situation was anticipated already in 1989/1990 which had led to shift of paradigm for audiovisual archiving; first for audio, followed since the late 1990s by video, and presently also applied for film archiving: The new strategy is to preserve the content, not the carrier, by transferring contents to digital files and migrate these files from one IT preservation platform to the next⁵¹. There is unanimous agreement that the time window left for transferring audiovisual

⁵¹ *The principles of this strategy and the practical guidelines for audio and video see IASA-TC 03, 04, and 06.*

carriers into safe digital repositories is only 15 years, if at all. After that even well preserved originals will be useless, because of unavailability of dedicated replay equipment.

Feeding analogue and digital single carriers into digital repositories is in demand of time, modern format-specific and regularly maintained replay equipment, and specialised experts in fading technologies, to keep operations running. It is important to understand that autonomous transfer operation can only viably be performed if critical masses are available, which must amount to several thousands of items per format.

However, transfer is the first step only. Professional IT repositories must be in place to take up digitised contents, which have to be migrated into the future from one technical preservation generation to the next. Long-term preservation of digital files calls for permanent engagement in terms of logistics, personnel, and financial means in previously inexperienced dimensions. Critical masses are again crucial for viable installations. Present costs are in the order 1-2 USD/GB/year for great repositories. Prices are further decreasing, but the slope will eventually flatten out.

Over the past 15 years we have seen many audiovisual digitisation projects that had suffered from a bundle of typical insufficiencies or mistakes: Most concern inadequate equipment for signal extraction from original tapes. Latest generation of equipment shall be used, because only this would capture the originally recorded quality. Mediocre or badly maintained equipment would distort original quality. Typical mistakes made in the production of digital files are the use of non-precision AD converters (as parts of cheap sound cards), the choice of streaming instead of file formats as digital target formats (CD audio instead of Wave), inadequate digital resolution (44.1kHz /16 bit instead of at least 48kHz /24 bits), the use of data reduced (“compressed”) target formats for analogue originals – e.g. DVD for analogue video originals or MP3 for audio, and, finally, the use of optical recordable discs as sole digital preservation media.

Radio and television archives as well as national archives of wealthy countries will solve preservation problems professionally by - possibly selective - transfer of holdings within next 15 years and by providing sufficient funds to keep digital files alive.

But what can be done to safeguard the small and hidden collections, which are outside custodial care, reflecting the greater part - estimated 80% - of the documents of linguistic and cultural diversity of human kind?

The first and foremost action to be taken is awareness raising. Poor standards of basic knowledge about audiovisual preservation principles are widespread, even

amongst librarians and (paper) archivists, let alone amongst the specialists in the contents of the audiovisual carriers. Most efficient, therefore, are tutorials and workshops organised at the fringe of discipline oriented conferences of specialists, such as linguistics, musicology or anthropology gatherings.

A typical barrier against safeguarding audiovisual research materials in the West is that excellence, and therewith budgetary allocations, are measured on the academic output of institutions and individuals. Optimisation in terms of publications is in higher esteem than safeguarding primary source material for systematic restudies and new interpretation by other (schools of) researchers or by later generations. This frequently leads to sub-optimal archiving standards, and to a further neglect of audiovisual preservation in case of budgetary problems of academic institutions.

Typical problems in Russia and post-soviet countries are mainly due to particular research traditions. Relatively small research units worked in separation of each other, often (part of) institutes of academies of sciences, and often even under one roof, generally without sharing archival infrastructures. In this manner great amounts of audio recording collections have been accumulated, significantly more than in the respective disciplines in the West. A specific widespread problem is the use of acetate cellulose tape of East-German origin, produced in the 1950s and 1960s, which was used all over the region until Vietnam. These tapes become very brittle with age and their replay is often a veritable challenge⁵².

As typical small collections of cultural and academic institutions are below critical masses to allow for economic viable preservation, cooperative solutions have to be organised and financed to rescue the accumulated holdings. Cooperative models can be organised at different levels:

Institutional: Universities, academies of sciences, and other institutions holding several audiovisual collections in their sub-units, join efforts by establishing a common preservation strategy. Digitisation and long-term preservation of digital contents is often organised separately. Typically, the strongest unit sets up a central transfer laboratory, while the digital files are often kept in a digital repository as part of the computer centre. A recent prominent example is the central organisation of the preservation of 560.000 audiovisual carriers at the Indiana University Bloomington, USA⁵³. A smaller, but structurally identical project is under preparation at the University of the Philippines⁵⁴,

⁵² See: Wallaszkovits 2012.

⁵³ http://research.iub.edu/communications/media_preservation/.

⁵⁴ <http://www.upethnom.com/index.php>.

National: Other successful, although yet not wide spread models work on a national level. One example is the Vienna Phonogrammarchiv⁵⁵, which holds a considerable part, probably 40-50% of audiovisual primary sources produced by Austrian scholars, which had made their field recordings with methodological advice and technological support by the archive. Over the past years, however, many institutions and scholars that had made their recording autonomously, without depositing their originals in the archive, offer now their collections, as they understand that their precious sources would otherwise be lost. The archive tries to raise sufficient funds over the coming years to safeguard at least the most important collections yet outside proper archival care.

A remarkable national system has been established by Memoria⁵⁶, a Swiss initiative to cooperatively preserve audiovisual collections of greater as well as smaller institutions and even private collectors.

On the European level, the Presto project family, after having specifically dealt with mass digitisation of radio and television archives, is presently under way to set up a network of competence centres⁵⁷ which are intended also to assist smaller collections. Direct funding of digitisation is to some extent planned within the European Strategy Forum for Research Infrastructures (ESFRI, Schüller 2011).

Last but not least, international discipline and research oriented initiatives have been established. Most of them concern linguistic material. The Max-Planck-Institute for Psycholinguistics has established a Language Archive⁵⁸, ELAR is the Endangered Language Archive of the London School of Oriental and African Studies⁵⁹, Paradisec⁶⁰, the Pacific and Regional Archive for Digital Sources in Endangered Cultures, collects languages and music from the Pacific region, while EVIA⁶¹, Ethnographic Video for Instruction and Analysis, is an international ethnographic video archive in Bloomington. The major problem of many of these and similar projects is the uncertainty of continued funding to keep digitised files available in the long-term.

Significant support has to come from disciplines themselves, which are challenged to enhance recognition of primary sources by promoting systematic re-studies of archive materials, by promoting diachronic and comparative studies, and by intensifying study of respective archival materials before new field work is started.

⁵⁵ <http://www.phonogrammarchiv.at/wwwnew/>.

⁵⁶ <http://en.memoria.ch/>.

⁵⁷ <http://www.prestocentre.org/>.

⁵⁸ <http://www.mpi.nl/research/research-projects/the-language-archive/>.

⁵⁹ <http://elar.soas.ac.uk/>.

⁶⁰ <http://paradisec.org.au/>.

⁶¹ <http://www.eviada.org/>.

The European Science Foundation as well as national research funding agencies are increasingly recognising the importance of research infrastructures, specifically in digital age, which will lead to a significant rise of financial means within the 8th European Framework Programme. On national levels, research funding should limit autonomy of researchers by enforcing and financing the deposit of raw research materials in archives for further (and alternative) evaluation, and, additionally, by earmarking a percentage of research budgets for infrastructure, e.g. for archiving.

Failure to preserve audiovisual primary sources will lead to their swift and total loss, which undermines fundamental research principles, invalidates modern research results, and diminishes considerably the resources for linguistic and cultural diversity in the cyberspace.

References

1. Bradley, Kevin. Risks Associated with the Use of Recordable CDs and DVDs as Reliable Storage Media in Archival Collections - Strategies and Alternatives. UNESCO, Paris 2006. Available at: <http://www.unesco.org/webworld/risk>.
2. IASA Technical Committee. The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy, edited by Dietrich Schüller. (IASA Technical Committee - Standards, Recommended Practices and Strategies, IASA TC-03), Version 3, 2005. Available at: http://www.iasa-web.org/downloads/publications/TC03_English.pdf. Also available in German, French, Spanish, Italian, Swedish, Russian, and Chinese.
3. IASA Technical Committee. Guidelines on the Production and Preservation of Digital Audio Objects, edited by Kevin Bradley. (IASA Technical Committee - Standards, Recommended Practices and Strategies, IASA TC-04), second edition 2009. Available at: <http://www.iasa-web.org/audio-preservation-tc04>.
4. IASA Technical Committee. Handling and Storage of audio and Video Carriers, edited by Dietrich Schüller and Albrecht Häfner (IASA Technical Committee - Standards, Recommended Practices and Strategies, IASA TC-05), forthcoming 2012.
5. IASA Technical Committee. Guidelines on the Production and Preservation of Digital Video Objects, edited by Kevin Bradley. (IASA Technical Committee - Standards, Recommended Practices and Strategies, IASA TC-06), forthcoming 2012.

-
6. Schüller, Dietrich. Audiovisual research collections and their preservation. 2008. Available at: http://www.tape-online.net/docs/audiovisual_research_collections.pdf.
 7. Schüller, Dietrich. Audiovisual Research Documents. Case study in: Research Infrastructures in the Digital Humanities (Science Policy Briefing 42). European Science Foundation, September 2011, 23-24. Available at <http://www.esf.org/research-areas/humanities/publications.html>.
 8. Schüller, Dietrich. The Role of Digitisation in the Preservation of Audiovisual Documents. In: Proceedings of the International Conference Preservation of Digital Information in the Information Society: Problems and Prospects, Moscow, October 2011, forthcoming 2012.
 9. Wallaszkovits, Nadja. Replastification of brittle cellulose acetate tape. Forthcoming 2012.

Adel EL ZAÏM

*Senior Programme Specialist, Information and Networks,
International Development Research Centre
(Cairo, Egypt)*

Cyberactivism in the Middle East and North Africa: Promoting Local Languages

Introduction

Although 2010–2011 will enter the history as the two years of dramatic change in the Middle East and North Africa, revolutions in Arab countries didn't start in 2010. One should say the revolutions culminated and succeeded mainly in Tunisia and Egypt in 2011, after several attempts during the first years of the 21st century. Revolts continue in several countries and are transforming in some others into massacres (Syria), into civil war and international intervention (Libya), and into a hidden repression (Bahrain).

In fact, the young population is uprising on all the continents. After China, Burma and Iran come Middle East and North Africa in 2010-2011 – from Morocco, Algeria, Tunisia, and Egypt to Libya, Yemen, Bahrain and Syria. Revolts and repressions are still harvesting lives and ruining economies in all these countries. People are also demonstrating in Spain and Greece as well as in Senegal more recently. In Senegal, population is demonstrating against the corruption and the president's plans to run for a third mandate and create a Vice-President position⁶².

The question about the role of the information and communication technologies, and social media in such a revolution was one of the first questions asked very early in Tunisia and Egypt. Researchers and journalists wanted to know if it can be called Digital Revolution or Revolution 2.0. More serious questions are also asked about the timing and the location of the revolution: why now, why in Tunisia, why in Egypt, what country is next? How to reform a country: by way of creating new political systems, and a new category of politicians, or political practices?

The new era of reforms and renaissance will fill a library of research reports and books, studying all kinds of impacts, benefits, history, roles, and new models. Dozens of books about the revolution have already been published in Tunisia and Egypt. Cultural products including songs, music, movies and photos are

⁶² <http://www.senrevolution.com> (among other sites).

being produced and exposed all around the world from Tunisia and Egypt to the United Kingdom and the United States of America.

And what about language? What about its usage during the uprising in MENA countries? Studying language and revolution is a very large domain which could be tackled from different perspectives. I had the opportunity to witness the revolution in Egypt and I was particularly interested by the role of language in that event. I studied the role of social media in Egypt mainly, and tried to explain how users/protestors and political movements used the Internet and the social media to communicate, mobilize the demonstrators, organize activities, and document and share the revolution events, repression, success and celebration. I came to ask the question about the impact of the revolution time on the presence of Arabic language on the net and in social media. Based on that experience, I am making the hypothesis that since the revolution started in Arab countries, more content in Arabic language is being produced and shared on the Internet and social media. This content is created by individuals, organisations and media, and is contributing to enhancing the rank of Arabic language on the Internet. Some research is corroborating my observations, but a lot is still to be done on that matter.

Usage of Social Media in Arab Countries

Social media like Facebook, Twitter and YouTube are among the most used tools of the Internet in Arab countries. Facebook was blocked and unblocked several times in different countries from Tunisia to Syria during the last 3 to 4 years for political and social reasons.

The 1st issue of the Arab Social Media Report⁶³ showed that over 21 million Facebook users are in the Arab world, as of January 5, 2011. This number jumped up to 27,711,503 users in the first quarter of 2011 (April 5th, 2011) as reported by the 2nd issue of the Arab Social Media Report published in May 2011⁶⁴, with a high 30% of increase in the same period. Countries of the Gulf Cooperation Council (GCC) dominate the top five Arab Facebook users as percentage of population. Egypt constitutes about a quarter of total Facebook users in the Arab region, and has added more users in the first quarter of 2011 than any Arab country, at close to 2 million new users between January 5 and April 5⁶⁵.

Governments are seizing the opportunity offered by this huge concentration of their citizens on Facebook and are becoming more present at that social media. Egypt and Tunisia constitute good examples, where immediately after the success of the revolts in ousting the presidents, the new government

⁶³ <http://www.dsg.ae/NEWSANDEVENTS/UpcomingEvents/ASMRHome.aspx>.

⁶⁴ <http://www.dsg.ae/NEWSANDEVENTS/UpcomingEvents/ASMROverview2.aspx>.

⁶⁵ <http://www.dsg.ae/portals/0/ASMR2.pdf>, page 9.

bodies started to reach out to Tunisians and Egyptians through Facebook. Although the e-Government applications are very well developed in these two countries, and each ministry and organisation have web sites, they created pages on Facebook, YouTube and opened accounts on Twitter to communicate their messages and try to engage the population in discussion and interaction inside the country and abroad. Even in a non-conflict situation, the United Arab Emirates (UAE) government built its profile on Facebook due to the presence of 45% of its population on that network. The government of the UAE is encouraging its employees to use social media to interact with citizens. It has trained some of them on the responsible usage and risk of Facebook, and offered a policy guidelines document for government entities⁶⁶.

Types of Social Media Usage in Arab Countries

Given the demographic of the population of the Arab countries (around 30% are young people between 19 and 25 years old), the political and economic situations, Facebook is being used in a wide variety of ways: “whether to rally people around social causes and political campaigns, boost citizen journalism and civic participation, create a forum for debate and interaction between governments and their communities, or to enhance innovation and collaboration within government.”⁶⁷ However, the main usage of Facebook was and is still the social networking among individuals and groups of buddies as intended by its creators. Despite the censorship and blocking by authorities, Facebook is the networking tool by excellence for those young people who want to communicate, meet each other, share hobbies and dreams, and endorse celebrities.

The first quarter of 2011 witnessed a dramatic change in the role and perception of Facebook in Arab countries, due to the cyber-activism and the revolts in the streets. Curiously, blocking Facebook and Internet in Tunisia, Egypt, Libya, and Syria, gave these tools more credibility and impact, and increased demand, then their usage.

As quickly observed in Egypt since January 2011, social media are used by cyber-activists for mobilising people and demonstrators, preparing political activities and movements, sharing instructions and hints, and informing the international communities about the local events. YouTube and similar networks are used to document and share the events, ranging from calls for meetings to demonstrations, attacks and massacres. Dated videos permitted the protestors to prove the pacifist character of their actions and the brutality of the authorities such as in Syria (April 2011), Libya or Bahrain. Twitter is the communication channel for rally, SOS, quick instructions, feeding and receiving

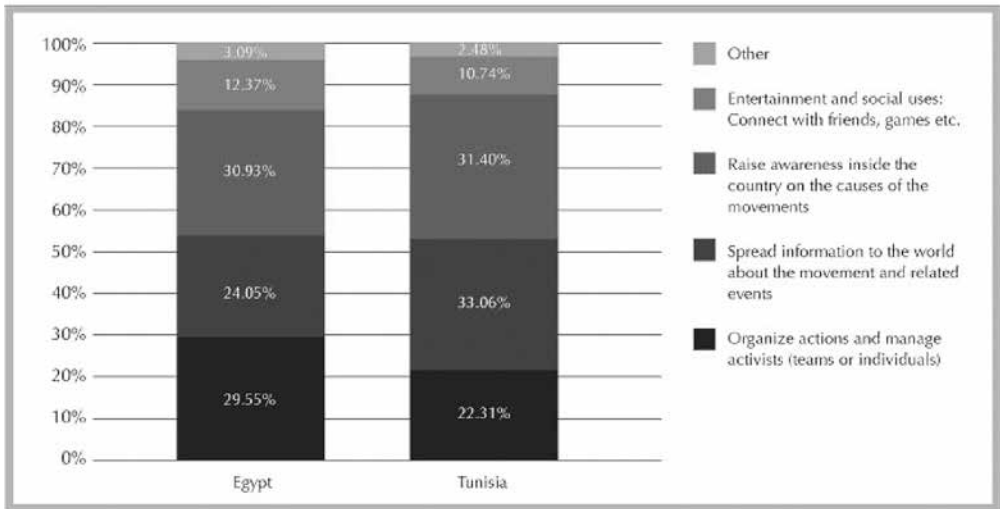
⁶⁶ *Guidelines by the government of the UAE in arabic*, <http://www.emiratesegov.ae/web/guest/83>.

⁶⁷ <http://unpan1.un.org/intradoc/groups/public/documents/dsg/unpan044212.pdf>, page 1.

news, among others. The very nature of Twitter resides in the short messages or micro blogging, which make its integration on mobile phones seamless.

Moreover, The Arab Social Media Report surveyed 126 people in Egypt and 105 in Tunisia that were asked about the main usage of Facebook during the civil movement and events in early 2011. As represented in the figure below:

“In both countries, Facebook users were of the opinion that Facebook had been used primarily to raise awareness within their countries about the ongoing civil movements (31% in both Tunisia and Egypt), spread information to the world about the movements (33% and 24% in Tunisia and Egypt respectively), and organize activists and actions (22% and 30% in Tunisia and Egypt respectively). Less than 15% in either country believed Facebook was primarily being used for entertainment or social reasons”⁶⁸.



The main usage of Facebook during the civil movement and events in early 2011

The Language Facebook Is Used In

Facebook offers its interface in tens of languages most of them localised by users themselves. Users in Arab countries surveyed for the Arab Social Media Report “vary in their preference of language interface”⁶⁹. Three main languages used on Facebook with no surprise are Arabic, French and English. The survey showed net preference for English in the Gulf countries, except for Saudi

⁶⁸ <http://www.dsg.ae/portals/0/ASMR2.pdf>, page 6.

⁶⁹ <http://www.dsg.ae/portals/0/ASMR2.pdf>, page 14.

Arabia, and net preference for French in the three Maghreb countries and the Comoros. Egypt and Tunisia are worth observing because of the changes we will see later during the revolution. In terms of preference of language interface, users in Egypt split evenly between the use of Arabic (49.88%) and English (48.98%) interfaces (similar to Jordan, Libya and Iraq). Tunisian users showed a net preference for French interface (94.60%), then English (2.72%) and finally Arabic (1.56%).

However, the language of the interface setting (only one language at a time) doesn't say much about the languages in which users are actually interacting on Facebook or other social media⁷⁰. In fact, thank to HTML and UNICODE, browsers are now able to display text in virtually all the world languages. "Facebookers practice a diversity" or a mix of languages which "challenges conventional notions of multilingualism as a combination of two or more monolingualisms"⁷¹.

Revolution Promoting Local Language

It is not a surprise that language played an important role during the social and political uprising in Tunisia and Egypt, as well as in other countries of MENA. Some slogans chanted by demonstrators made the tour of the planet and became symbols or songs and are repeated by demonstrators around the world. Remember: "Ben Ali, dégage!" in Tunisia, or "The *people want* the regime to fall" (ماضنلا طاق سا ديرى بعشلا), repeated in Tunis, Cairo, Damascus, Benghazi, and Sana'a. Signs made and handled by protestors in Cairo streets were in Arabic mainly, but also in English, French, and Hebrew⁷².

On the social media front, linguistic creativity was positively impacted by the uprising. We are making the hypothesis here that due to the revolution and the need to reach out to a larger community on burning issues social media users used local languages (Arabic), increasing the quantity of Arabic content published online both on social media like Facebook and Twitter, and on regular websites.

This hypothesis is based on our observation of 1) the number of new websites published in Arabic by newspapers, social movements, and government entities; 2) the number of social media users who are now writing in Arabic.

⁷⁰ As demonstrated in <http://www.languageonthemove.com/language-globalization/multilingualism-2-0>, blog posted on August 02, 2010 by Ingrid Piller.

⁷¹ *Ibid.*

⁷² See picture by gcarlstrom (Gregg Carlstrom) at <http://yfrog.com/h3fbsbj> with a slogan in Hebrew "Azov Mubarak", translated as "Leave, Mubarak".

Attempt to Validate the Hypothesis

This hypothesis is corroborated by the results of the Arab Social Media Report updated in the 2nd issue of May 2011. If we compare the language of the interface, and the language used by users in Egypt and in Tunisia to communicate during the civil movement of the first quarter of 2011, we see a huge difference.

The interface language distribution⁷³ is as follows:

Country	Arabic (% of FB users)	English (% of FB users)	French (% of FB users)
Egypt	49.88	48.98	0.39
Tunisia	1.56	2.72	94.60

However, questioned⁷⁴ about the primary language they used to communicate on Facebook during the civil movement, citizens of Egypt and Tunisia responded like this:

Country	Arabic (% of FB users)	English (% of FB users)	French (% of FB users)
Egypt	75.40	25.60	0
Tunisia	51.43	0.95	47.62

The difference between the language of the interface and the language of communication is accentuated since the revolutions in Egypt and Tunisia.

As an empirical way to validate that hypothesis, I tried to examine the tweets about #CairoExplosion happened on July 6, 2011. Cairo and suburbs residents heard on Wednesday, July 6, a big explosion. For hours nobody was able to guess what it was exactly. Rapidly, a hashtag was created on Twitter and users started sending messages asking or guessing or retweeting what they heard. In less than an hour, that hashtag started to trend (becoming a trend on Twitter), creating a sort of panic among the tweeps in Egypt. I copied randomly 500 tweets with the hashtag #CairoExplosion sent in two hours approximately, and classified them in terms of languages. I obtained the following distribution:

Arabic	Arabic in latin characters	English	Mixed	Not determined	Total
353	33	89	19	6 (smiles)	500

Despite the interests of those numbers, the observation is still limited and needs to be systematically examined and validated. There are some initiatives like

⁷³ Arab Social Media Report, Issue 2, May 2011, <http://www.dsg.ae/portals/0/ASMR2.pdf>, page 14.

⁷⁴ Arab Social Media Report, Issue 2, May 2011, <http://www.dsg.ae/portals/0/ASMR2.pdf>, page 7.

R-Shief⁷⁵ aimed, among other things, at mining and visualising Twitter content and the public sphere of Facebook to study the language distribution and uses.

Increasing the quantity of content doesn't imply enhancing the quality however. Twits are limited by their very nature of short messages or microblogs. Content of social media is characterised by its informal style, although more institutional pages and networks are created and contributing then to giving more formal status to social media, mainly by the fact of the presence of government entities and organisations. A huge number of international organisations like the United Nations Development Programme (UNDP), International Fund for Agricultural Development (IFAD) and Canada's International Development Research Centre (IDRC) are maintaining Facebook pages and Twitter accounts, contributing then to enhancing style and credibility of such tools. Since Obama's first campaign, almost all candidates of the developed and developing world play the social media card and try to reach their voters, or more precisely the Facebookers among their potential voters.

Conclusion

We have signals that social uprising, political and social unrest are making social media one of preferred tools to communicate, mobilize, demonstrate and voice the concerns of the population around the world. MENA countries showed a serious increase in the number of users of social media that are mainly communicating in their mother tongue, despite their interface preferences and the software offers. More research is still needed if we want to better document this phenomenon and build on it to enhance the quality of the content and, more important, the quality of the citizens' participation and engagement in the cyberspace, leading to more interaction and more benefits from the knowledge society.

⁷⁵ <http://www.r-shief.org/>.

Elmir YAKUBOV

Director, Khasavyurt Central Municipal Library System
(Khasavyurt, Russian Federation)

Developing Digital Content in the Languages of Dagestan: A Means of Overcoming the Increasing Marginalization of Local Languages

While specifying the most pressing challenges of enlarging access to the world's documentary heritage, the Lena Resolution adopted in 2008 at the International Conference on Linguistic and Cultural Diversity in Cyberspace has offered ways of efficiently tackling these challenges by means of developing multilingualism in digital environment.

For the last three years after the dissemination of this essential document strenuous efforts have been made in the Republic of Dagestan by government agencies and civil society organizations to increase the volume of digital content in local languages in response to the growing marginalization of languages of numerous Dagestan ethnic groups. Given that marginality implies a shift towards monoculture, activities aimed at overcoming this negative trend in our specific linguistically diverse region should have a special focus on the preservation of languages and cultural openness of the peoples of Dagestan.

Today the use of regional languages is drastically reduced to family and household communications in very limited areas, mainly rural, while cultural requirements and linguistic interests of Dagestan's peoples go far beyond ethnic borders. Emerging digital resources of various kinds has become the means of defining sociocultural profile of modern Dagestanis and their glorious forefathers.

Full-text collections of fiction and local literature account for a sizable proportion of the Dagestan segment of the Internet. At this stage they cannot be considered full-fledged digital libraries, as the documents displayed are of medium quality, lack metadata, navigation and search tools. However the demand for such publications in traditional libraries makes us positive about their popularity in digital environment. This is proved by constant additions to these collections and an increase in the number of resources providing relevant content. Affecting sensory perception, digital products of that kind provide for a better and deeper understanding of one's affiliation not only to a certain ethnos, but to the whole mankind.

The Lena resolution highlights *inter alia* the necessity of further support for the creation of multilingual dictionaries and thesauruses. Dagestan can boast

of successful activities in this field. Over the past two years institutions of the Russian Academy of Sciences prepared and published for the first time ever a 18,000 Avakh-Russian dictionary and a Tsakhur-Russian dictionary comprising over 10,000 native and adopted words. The dictionaries are currently being digitized by one of the enthusiasts of Dagestan web resources development with the agreement of the copyright holders. Digital versions will be accessible online in addition to the 14 Dagestan-Russian dictionaries available there.

In terms of the problem at hand, how mass media define the role and time to be provided for regional languages is of vital importance. Today a language can maintain its position as long as it actively penetrates into new communication modes. Despite new opportunities for regional languages emerging due to the development of broadcast techniques, active support by the state is required for Dagestanian languages to find a niche in digital media.

A web portal for various Dagestanian newspapers in 13 local languages has been running since 2009, providing information on social, political, economic, academic, cultural and sports life of the republic. The Dagestan State TV and Radio Company broadcasts in six local languages, and two state radio companies broadcast over four hours daily in 13 languages. Almost every town and district has its own TV production companies funded by local budgets and broadcasting for the most part in one of the local languages. The number of private radio companies has increased significantly, that include national programmes in their line-up. Some of these channels have already created their own web pages with access to valuable archives of ethno-cultural materials.

However these costly measures are not always efficient for tackling the pressing problems that regional languages are facing today due to a lack of qualified experts and the passivity of management unable to keep pace with the rapid changes of media environment. The just-completed educational project supported by the Council of Europe and European Commission was aimed at the elimination of these two problems. The results are yet to be summarized, but for now it is obvious that the project has equipped the republic's leading digital media with powerful tools for preserving and developing local languages.

According to social surveys, Russians spend over 10 hours per month in social networks, being second only to Israelis. Networks are quite popular among Dagestanis as well, especially among younger generation. It is notable, that many prefer using their mother tongue for communication on Facebook and similar social network sites "Odnoklassniki.ru" and "VKontakte.ru". Blogs in regional languages have appeared recently and are actively developing. Languages of Dagestan peoples are widely used in various chats and forums. Possible negative impacts of promoting social networking in regional languages, like

building up a cult of ethnic exceptionness or alienation from other population groups, used to raise our concern. However, now we witness the emergence of the first multilingual social network Godudu – a project supported by the Ministry of Communications and Mass Media of the Russian Federation and the Government of the Republic of Dagestan. The investment project was initiated by businessmen from the United Arab Emirates, a country showing how modern technological solutions can be combined with preserving original cultural traditions. Language barrier is the last barrier to people’s unlimited communication in the Internet era. Godudu developers provide users with an opportunity to write in their mother tongue to addressees speaking other languages. It is made possible with an online translator. For now the system knows five languages and the translation lack quality, however it is already a great alternative to a communication failure. Although Godudu does not support Dagestania languages as yet, the project’s authors pin themselves to add them in the course of the network development. They are also planning a system to let users study their mother tongues by following their own correspondence.

It is evident today, that developing digital resources allows for Dagestania languages to integrate successfully into the new social reality, ensuring their active use in every sphere of life. Technological change and ITCs do not bring harm to the indigenous languages “cultural aura”, but, on the contrary, can make up for the poor conditions, in which these languages used to exist.

Sergey BOBRY SHEV
Commercial Director, ParaType Inc.
(Moscow, Russian Federation)

Emil YAKUPOV
Director, ParaType Inc.
(Moscow, Russian Federation)

Multilingual Fonts for Visual Communications

Russia is a vast and polyethnic country, perhaps the vastest and most polyethnic in the world. More than two hundred languages are used in its territory, belonging mainly to four language families: Indo-European, Altaic, North Caucasian and Ural. For the most part, the languages of peoples of Russia are based on the Cyrillic alphabet, but there are languages that use Latin and other writings, as well as oral languages. Russia as a multinational country has accumulated a solid experience of friendly coexistence of various nations.

In such a country, the most obvious integrating element is not the territory and certainly not the language. All of us, the people of the Russian Federation, are united by our writing and its graphic tools, the fonts.

Therefore, in 2009 while initiating a project to develop a nationwide font typeface for the title languages of subjects of the Russian Federation the ParaType Inc. set the following goals:

- creating enhanced language support fonts for the languages of Russia,
- replacement commercial fonts with free/libre fonts,
- replacing western design with domestic for Cyrillic fonts,
- creating a set of fonts for a wide range of applications.

These goals conform to the Federal Law “On languages of the peoples of the Russian Federation” granting all languages of the peoples of our country equal rights for preservation and development, guaranteeing the right for education in native languages, and fixing the Cyrillic as the writing system.

At present the process of formation and regulation of national languages’ use is gaining momentum both in our country and throughout the world, regional laws on languages are being elaborated and adopted. This process should be accompanied by the creation of high-quality national fonts to be freely available.

However in order to develop fonts supporting national characters knowledge of both the composition of national alphabets and the form of these glyphs’ representation is required. Today, the Unicode standard determines most of them.

Unicode							
ä	ǎ	æ	г	ѓ	ƒ	ђ	ћ
ђ	ě	е	ѓ	ж	Ж	Ж	з
ž	š	з	й	й	і	ї	ј
ќ	ќ	к	к	к	ќ	љ	н
ң	њ	њ	ö	ө	ë	ø	п
ç	т	у	ÿ	ÿ	ÿ	ү	ү
х	ц	ц	ч	ч	ч	ч	ы
е	ә	ә	ı				

Figure 1. Representation of additional Cyrillic characters in accordance with the Unicode Standard

We collaborate with the Unicode Consortium and a number of specialised agencies of Russian entities responsible for the issues of national languages to monitor the current linguistic situation. To provide feedback we created a special page on our site www.fonts.ru/public, representing all titular languages of the RF subjects and their characters.

As a rule, regional laws on languages require duplicating texts and inscriptions in national languages in official documents, on road signs, signboards, etc. In addition, even without special legislation literature in national language should exist, in particular, textbooks and dictionaries. Therefore, major requirements for a national font are multilinguality, i.e., ability to support character sets of national alphabets, and accessibility. These are indispensable conditions for tackling a wide range of national fonts' issues.

In order to solve these problems efficiently within regions professional communities and local authorities should cooperate to settle the national characters set and forms of these characters.

The elaboration of regional language laws, cooperation with local institutions and authorities, as well as public debates in the media are essential in this regard.

Free/libre fonts are those included in operating systems or put on the Web for free access. Existing national fonts do not meet certain requirements. Fonts put on the Internet for public access are usually of very poor quality both in terms of design and technical execution. Those few fonts with an extended set of characters included with Windows, support only a very limited number of Cyrillic-based languages.

Paradoxically, the population of a huge country mostly uses fonts designed by a private American company. With all due respect for Microsoft, even knowing their careful and thoughtful approach to national traditions, we cannot expect them to be able to embrace the boundless and solve the problems of preservation and development of writing of small peoples of Russia.

The following decision is logical in this situation. The State may order a set of national fonts to be developed and made available for free downloading online. In addition, it is desirable that these fonts are included in the localized operating systems distributed on the territory of Russia (Apple Macintosh and Microsoft Windows). These fonts should support all titular languages of the entities of the Russian Federation.

These considerations and the desire to get universal modern fonts, consistent with the idea of good Cyrillic became the basis of work on the design of the PT Sans – PT Serif typeface system by the ParaType Inc. in 2009–2011 with support by the Federal Agency for Press and Mass Communications⁷⁶.

The Structure and Composition of the PT Sans – PT Serif Typeface System

In 2009, ParaType developed and presented PT Sans, the first typeface of the project. In 2010 an antique font PT Serif was developed in the same proportions.



Figure 2. PT Sans – PT Serif typeface system. Basic font styles.

⁷⁶ *PT Sans and PT Serif were designed by Alexandra Korolkova with participation of Olga Umpeleva and under supervision of Vladimir Yefimov, ParaType Art Director.*

PT Sans is a grotesque font of modern humanistic design intended for widespread use. The font is based on classical designs, but includes very distinctive features fulfilling present day aesthetic and functional requirements and making it usable for large-size headlines.

The family consists of 8 styles: 4 basic styles; 2 caption styles for small sizes and 2 narrow styles for economic setting and is primarily intended for both printed and e-document flow.

PT Serif is transitional serif face with humanistic terminals designed for use together with PT Sans and harmonized with PT Sans on metrics, proportions, weights and design. PT Serif consists of six styles: regular and bold weights with corresponding italics form a standard computer font family for basic text setting; two caption styles (regular and italic) are for texts of small point sizes.

Figure 3 shows distinctive features of both fonts.



Figure 3. PT Sans – PT Serif typeface system.

PT Serif covers a wide range of type sizes.

Captions styles of PT Sans (Caption, Caption Bold) are designed for notes, text boxes, and references. They are suited for web sites as well.

But these styles are applicable primarily for various road signs, plates, billboards and other objects of communication design in the urban environment. They are readable from a big distance in poor lighting conditions, fog, rain, etc.

Two narrow styles (Narrow, Narrow Bold) are designed for putting a large amount of text in a limited amount of space on a printed sheet. Low compression ratio (20%) allows for considerable savings without reducing readability. Texts on packaging, annotations to medicines and special conditions of insurance contracts (where narrow styles are traditionally used) are thus made readable.

PT Sans – PT Serif Characters

Each of the 14 styles of the PT Sans – PT Serif typeface system contains about seven hundred characters. In addition to the standard set of characters to work with the languages of Western and Eastern Europe, and the standard set of Cyrillic, these fonts include all characters of alphabets of title languages of the Russian Federation and are a unique and extremely important tool for the development and preservation of written cultures of the peoples living in our country.

Free/libre fonts supporting national writing systems meet the needs of education, culture, media, government agencies, businesses, and, moreover, facilitate digital communications of the representatives of the Russian entities in their native languages and the development of national resources in the Internet. Having a common font for all national languages of the country is important for international communication on the basis of a common standard, both in terms of encoding, and in terms of similar design for cognate languages and dialects. This will prevent the isolation of neighboring cultures because of differences in the tradition of writing of same letters.

Experts of the Institute of Linguistics of the Russian Academy of Sciences were engaged to design the set of characters for PT Sans PT Serif and prepare a special section with information on national writings for www.fonts.ru. The development of PT Sans Pro and PT Serif Pro fonts allowed for support for 163 languages in seventy styles (32 and 38 Sans Serif).

Analysis of the Use of PT Sans and PT Serif

The PT Sans – PT Serif typeface system has quickly gained popularity with designers all over the world.

Google Web Fonts API Stats monitors the demand for fonts analysing the number of users viewing websites using the above fonts.

To date, more than 13 million users viewed sites where PT Serif is used (it is over 5 million a month!). The demand for PT Sans is even more significant – more than 210 million users per month! It is ranked fourth in a huge list of the most used fonts. We hope to enter the top three in the nearest future.

Font Family	Requests				Trends		Rate fonts/sec
	Total	30-Day	7-Day	1-Day	4-Week	1-Week	
TOTAL	17,149,566,237	4,551,792,816	1,037,061,578	152,162,411			
Droid Sans	2,134,807,893	424,349,251	106,785,831	12,725,919	-16%	-11%	219.9317
Droid Serif	1,024,949,084	235,077,153	57,986,776	7,060,878	-14%	-17%	122.8167
Lobster	1,272,427,333	217,053,722	56,011,231	6,908,326	-3%	+15%	119.3744
PT Sans	847,297,974	213,689,645	48,686,495	6,864,220	-12%	-5%	109.8856
Yanone Kaffeesatz	904,344,885	188,569,606	42,917,811	5,245,486	-29%	-21%	86.4833
Ubuntu	575,246,890	126,287,598	29,709,930	3,746,410	-16%	-16%	66.6631
Arvo	318,578,415	150,598,686	35,634,849	3,502,827	-36%	-43%	64.9794
Open Sans	171,931,510	104,019,754	26,102,547	2,968,566	-10%	-2%	48.8272
Calligraphiti	250,643,504	81,726,098	19,616,636	2,777,915	-0%	-3%	41.6006
Pacifico	134,449,944	83,073,349	19,392,659	2,650,020	-6%	-9%	35.4639

Figure 4. Top ten of the most popular fonts by Google Web Fonts API Stats

PT Sans and PT Serif as free/libre fonts are freely available and distributed. They can be used, copied, modified, embedded in documents, etc. provided that both the original fonts and their derivatives or parts thereof are not used for commercial purposes, except for their usage as part of other commercial products. Fonts are available at www.fonts.ru/public and will be included with operating systems.

The importance and significance of PT Sans – PT Serif fonts for multilingual visual communications was proved in 2011, when Apple Inc. included PT Sans into the Mac OS 10.7. Microsoft made a similar decision for Windows 8! Moreover, the website of the British Government redesigned in 2011 uses PT Sans and PT Serif among other fonts. Downing Street relying on Russian fonts for their typographic needs? What is it if it is not undeniable success?!

Emil ASANOV

*Director, VectorSoft Company
(Bishkek, Kyrgyzstan)*

Localization of Free Software as a Means to Support Multilingualism and Information Literacy Development of the Population of Kyrgyzstan

Localization is normally understood as both the process and the output of translation from one natural language (the source language) into another (the target language) of the software product interface and/or help system, as well as its messages. As far as Russian is concerned, such terms as “Russification” and “Russifier” are widely used to mean localization. RuNet has a great variety of sites dealing with software Russification.

This report looks into the issue of the Kyrgyz translation (i.e. the Kyrgyz localization) of the most well-known and widespread free software products: OpenOffice and Mozilla Firefox. The former is a worthy rival of the Microsoft Corporation’s MS Office software suite, and the latter is a successful competitor of Microsoft Internet Explorer web browser.

In Kyrgyzstan, like in all the other CIS countries, the Windows Operating System is the most widespread and comfortable user environment. Many people have no idea that other similar systems exist; though users of smartphones are also familiar with Android OS.

Our survey has revealed that most Windows users believe the Mozilla Firefox browser to be a Windows OS application, and are completely unaware of its cross-platform nature. And vice versa, the OpenOffice suite is steadily associated with the Linux Operating System, though it is cross-platform similarly to Firefox.

So why were Mozilla Firefox and OpenOffice, of all, chosen for localization? The main *pro* is that both software products are distributed under free, non-commercial licenses; moreover, it is possible to use (optionally) their source codes.

Localization of the Mozilla Firefox Web Browser

The first Mozilla Firefox localization (of version 1.0.3) financed through the grant of the Soros Foundation – Kyrgyzstan, appeared in February, 2005. But due to the lack of understanding of problems and their solutions throughout the working process, the localization product became highly outdated by the time of issue; and localization updating is one of the key elements in the comprehensive approach. The work was probably discontinued because of

the low spread (at the time) of the browser across Kyrgyzstan as compared to Internet Explorer and Opera. The situation changed dramatically by 2009 when Firefox became more popular than its competitors.

By the summer of 2009, the browser was upgraded to version 3.5 which underwent a final localization process. The basic and central distinction of version 3.5 and later versions lies in the fact that our Firefox build has become multilingual; this means that it allows switching between Kyrgyz and English or Russian.

All other existing localized Firefox builds were monolingual, which led to certain difficulties: since it was impossible to use two versions in different languages simultaneously, you had to choose in favor of an English, Russian or Kyrgyz version. The appearance of the multilingual build offered a solution to this dilemma, and the build itself got widely spread in the Firefox user community in Kyrgyzstan. Today, the user may choose between two versions: typical installation and portable editions. The Tamga-KIT software product (to be described below) contains the Firefox 3.6 browser portable version as a mandatory item.

Localization of the OpenOffice Suite

OpenOffice is traditionally used as a basic software office suite in Linux environment for various builds, e.g. Debian, Ubuntu, Fedora, etc. Infra-Resource's Russified OpenOffice version was of great help in popularizing the suite among Windows users. After a number of technical problems with the deployment of OpenOffice spell check and hyphenation modules for Kyrgyz had been settled (in-house development), the Kyrgyz OpenOffice 2.4 localization for Debian OS came out in late 2007, and the Kyrgyz OpenOffice 2.4.2 localization for Windows OS, in early 2008. Today, OpenOffice version 3.2.1 is available on the Kyrgyz market. The portable version is also offered as a supplement to the Tamga-KIT software solution.

The instrumental and term bases developed during the work on the OpenOffice and Firefox localizations made a very prompt localization of an enormous bulk of both free and proprietary software possible.

Background of Works on Kyrgyz Localization

Before starting work on the localization, we had to address the following general issues:

1. Development and finalization for introduction of general standards guiding the usage of the Kyrgyz language in ICT;
2. Development and deployment of drivers of standard keyboard layouts;

-
3. Development and deployment of Kyrgyz language spell check and hyphenation software modules in MS Office and OpenOffice suites;
 4. Issuing general electronic dictionaries based on the available paper analogues;
 5. Developing common ICT terminology and special terminology for word-processing units and web browsers.

We consider items 1, 2 and 5 of the above essential for getting started with the work on localization of public software, for instance, office software. Items 3 and 4 are desirable, since they allow for higher speed and quality of localization.

Alongside tackling the above stated tasks, tools to computerize the localization process of software products were developed for home use. They simplify the process of upgrading localized versions, reducing it to updating and clarifying data bases for item and message translation, which takes dozens of times less time than manual localization.

Tamga-KIT Software Solution

The Tamga-Kit Software Solution was developed to provide full support of the Kyrgyz language in Windows environment. The product basis comprises:

1. keyboard layouts;
2. 8 bit <=> 16 bit conversion system;
3. spell check and hyphenation modules;
4. thesaurus;
5. dictionaries by K.K.Yudakhin.

Depending on its purpose, a specific product version may contain up to 20 different components: additional software modules and localizations, thematic and terminological dictionaries, electronic manuals of the Kyrgyz language, phrasebooks, encyclopedias and different reference materials on the Kyrgyz language and culture.

Basic components have been developed since mid-1990s, and the full version of Tamga-KIT available since November, 2002, is free of charge for both secondary education and home use. The case of its creation is unprecedented: no funds from either government or non-governmental organizations were spent on its development. But its functionality and performance speak for themselves. This software has become extremely popular and widely-used; moreover, by parliament's decision the use of Tamga-KIT in education, science, culture, as well as in public and local institutions (i.e. nearly everywhere) was made mandatory.

Dzhavdet SULEYMANOV

*Professor, Kazan State University;
Member, Academy of Sciences of the Republic of Tatarstan
(Kazan, Russian Federation)*

The Tatar Language and ICTs

Introduction

In today's globalised world, the demand for languages as cognitive and communicative tools, as well as their development and even further preservation as cultural phenomena are largely dependent on the use of information technology.

Let us study the experience of introducing the Tatar language in cyberspace (i.e. the space of human interaction with computer systems and technology).

Ensuring the Tatar language usage in computer systems can help it significantly increase its relevance and competitiveness as a means of information acquisition and human-computer interaction. It can also facilitate creating new processing and storage technologies for data in the Tatar language considering a number of its structure and vocabulary' cognitive features.

To ensure equitable functioning of Tatar and Russian as official languages of the Republic of Tatarstan it is vital to make Tatar really usable in computing systems. Accordingly, special processing programmes for the Tatar language should be created, and Tatar localizations for software are required.

In the late 1980s research and development on introducing the Tatar language in IT began in the Republic of Tatarstan with the development of first peripheral device drivers, text editors and a Tatar spellchecker necessary for publishing Tatar books, newspapers and magazines in digital form, as well as for digital recordkeeping. In 1993, a Joint Research Laboratory for Artificial Intelligence Problems was established under the Academy of Sciences of the Republic of Tatarstan (RT) and the Kazan State University to meet the challenges of the Academy's *Scientific Programme on the Application of IT for Ensuring the Usage of Tatar as a State Language of the Republic of Tatarstan* and to develop relevant software under the *State Programme on the Preservation, Research and Development of Languages of the Peoples of the Republic of Tatarstan*.

In the last 15 years the Research Institute for Applied Semiotics and the Research Laboratory for AI Problems have been active in the field of developing

standards and software for supporting proper usage of the Tatar language in ICTs as one of the state languages of the Republic.

Fundamental research and development have been made in three major areas:

1. introducing the Tatar language in cyberspace,
2. developing and adapting IT for Tatar,
3. using cognitive potential of the Tatar language for creating new information technologies.

Introducing the Tatar Language in Cyberspace

This area of research and development is directly linked to the problem of preserving the language, increasing its activity in the infocommunication space, and facilitating the use of Tatar in cyberspace as a cognitive and communicative tool, i.e. a means of information presentation, storage and transmission. It is also essential for ensuring equal use of Tatar and Russian as the official languages of the Republic of Tatarstan, and for enabling native speakers' direct communication with computer systems without any intermediate language.

This area of work includes basic and full localizations of operation systems.

By now this task has been fully accomplished for the Cyrillic Tatar language. Screen and keyboard drivers, printer drivers, and fonts have been designed by the experts of the Academy of Sciences and the Kazan State University and offered as standardized for use in information technology in Tatarstan. On this basis, on December 9, 1996 the Cabinet of Ministers of the Republic of Tatarstan adopted the *Decree on the Tatar Characters' Encoding Standard for Computer Applications*.

The Decree helped unify the device drivers which were initially created by various groups and individuals and distributed to different computers, thus fixing diverse layouts and leading to a "discrepancy." Before the introduction of a common standard, Tatar texts typed on one machine were incorrectly displayed – if readable at all – on another. The unification of device drivers has prevented chaos in recordkeeping.

Under the agreement with Microsoft, device drivers were developed and brought to the Windows NT and MS Office 2000 in compliance with the adopted standards. Currently, the TATWIN driver package, included in the TatSoft 2 software package supporting Tatar, allows for recordkeeping in Cyrillic Tatar in all Windows 95/98/2000/XP/Vista/7 applications, as well as for websurfing. The Microsoft website contains relevant information

about this package. Due to our experts' efforts Tatar has become the second Turkic language (after Turkish) available in the Windows environment at any workplace.

The Academy of Sciences' cooperation with the Microsoft Office in Moscow began in 1995 with the Tatar localization of Windows 95 and found a promising sequel. In 2005-2010 Tatar localization of Microsoft Windows XP, Windows Vista, Windows 7 and MS Office applications was completed. The Research Institute for Applied Semiotics and the Research Laboratory for AI Problems developed a Tatar OS interface. The Tatar language, along with such international languages as English and Russian, has thus become a "native" language of the Windows OS and popular MS applications (Word, Excel, Power Point).

Tatar localization of the MS Windows OS and its applications stimulates the introduction of Tatar in ICTs, its development and expansion in the information space. It is clear that only in case the Tatar language (just as other peoples' languages) is actively used in ICT, in information storage, processing and transfer, and in human-computer interaction it can become a real state language of the Republic, a language of culture and science, an efficient means of communication in cyberspace.

A test version of a driver package and fontware for the Latin Tatar language has also been developed.

Upon the recommendation of the Academy of Sciences of the Republic of Tatarstan the *Decree on the Encoding Standard for Latin Tatar Characters and on Basic Computer Software Products* was adopted on September 27, 2000. The TATLAT Software Package was created, allowing for recordkeeping and publishing in Latin Tatar in all applications of Windows 95/98/2000/NT4. These software packages were being actively implemented for several years after the adoption of the corresponding Law of the Republic of Tatarstan, presidential and ministerial decrees, State Implementation Programme, and a number of educational initiatives. At present, however, the Russian State Duma's decision has frozen this process.

Developing and Adapting Information Technologies for the Tatar Language

Activities in this area include developing software packages for the Tatar language, tools for digital recordkeeping, newspaper and magazine publishing, spell checking Tatar texts, as well as for creating computer workstations. Research has been carried out on theoretical and applied aspects of computational linguistics with regard to Tatar, its grammar, lexicology and

lexicography, and its various manifestations in speech. The key goal is building pragmatic oriented linguistic models to serve as a basis for the creation of automated systems of processing data in the Tatar language. Special attention is paid to the decisive issues of Tatar terminology in cyberspace.

So far, a full-featured computer model of the Tatar morphology has been created. There are several options to choose from given the structural specificity of the Tatar language and applied problems to be solved. The **Generative morphology model** based on the inflection rules, while operating a bit slower, provides complete analysis of word forms taking into account the agglutinative nature of the language and allows to recognize word forms of potentially unlimited length. The **Paradigmatic model** provides rapid detection of word forms and their correctness analysis with up to 95 percent accuracy. The model is used in the search engine of the RUSSIA University Information System (by the Centre of Information Technologies of the Lomonosov Moscow State University), as well as in the MS Windows and MS Office applications. The recognition speed reaches 100 words in 0.014 seconds.

In addition, within a joint project with the Bilkent University (Turkey) a *two-level morphology model* has been developed working under the well-known PC-KIMMO shell programme. It is used as a part of the Tatar-Turkish machine translator.

A *structural functional model* of Tatar affixational morphemes has also been created, allowing for the construction of various pragmatic oriented morphological models. It served as a basis for the integrated “Tatar Morpheme” software data set. In fact, it is a computer workstation for developing various linguistic processors, and for educational and research activities in the field of Tatar linguistics. The “Tatar Morpheme” can be successfully used as a research tool for other languages as well.

The Tatar-Russian machine translator of Tatar proper names is especially essential for automated systems of Civil Registry Offices and Passport and Visa Services. The programme is also used to automatically generate names in reliance on the Tatar names component model.

ABBYY FineReader OCR software tool has been successfully localized to Tatar. Due to the built-in Tatar morphology component Tatar texts are recognized with the same speed and accuracy level as Russian and English ones.

We are currently working on the creation and support of a digital Tatar corpus, i.e. an Internet-based national corpus with the following components:

- Digital raw (unformatted) texts (newspapers, magazines, books, documents, etc.);

-
- Annotated texts, dictionaries, thesauruses;
 - Software units: linguistic processors (machine translators, TTS engines, text and speech recognizers, etc.); computer workstations (for teachers, editors, linguists, etc.); multilingual search engines.

Creating a *Tatar electronic corpus* is a fundamental scientific and practical problem. By solving it we will ensure quick and easy computer access to a variety of large-scale linguistic resources. This project can help establish an appropriate infrastructure (Tatar content and proper tools for working with it) for the Tatar language to be fully represented in cyberspace.

Another interesting and useful product, developed in cooperation with the ABBYY company and the Language, Literature and Arts Institute of the Academy of Sciences of the Republic of Tatarstan is a multilingual Lingvo X3 dictionary, containing 154 dictionaries in 12 languages, including Tatar. This electronic dictionary is of value for Tatar-speaking users, as through the Tatar-Russian dictionary of about 60,000 entries one can find equivalents in all dictionaries for other languages (English, French, Spanish, German, Chinese, Turkish, etc.).

WordCorr, a morphological spell checker of Tatar texts for MS Word, is an essential tool for recordkeeping and publishing. The programme allows to find misspellings in Tatar texts and correct them choosing from a list of word forms. It works for Windows 95/98/2000/XP/Vista/7 and applications.

Since 1990s, we have been working actively on the creation of the Tatar language e-learning tools, and educational programmes for studying various subjects in Tatar. A number of recent developments are available on the Internet, such as the Tatar Teller Zaman, a multimedia electronic Tatar textbook (<http://ttz.fosslabs.ru/>); Tatar-Online, a multimedia online Tatar textbook (<http://dev.tol.tatar.ru/>); and a multimedia fifth grade Tatar textbook for remote education (<http://distat.stage.metastudio.ru/>).

The Tatar Teller Zaman Programme contains over 2000 Tatar words, more than 2,500 images, voiced dialogue on various topics, and 11 exciting language games. Three types of different exercises allow to assess a student's language proficiency. One can improve Tatar pronunciation by way of repeating after the recording. Multilingual interface (Russian, Tatar (Cyrillic and Latin), and English) enables both Russian and English speakers to study Tatar.

With the assistance of the Ministry of Education and Science of the Republic of Tatarstan and the Magarif Publishers in cooperation with the Prosveshcheniye-Media Company (Moscow), Tatar versions of electronic teaching aids "Chemistry-8" and "Physics-7" have been developed. Multimedia

options (video, animation, sound, high-quality illustrations, and hundreds of interactive activities) make the learning process interesting and efficient. Technologies and software for creating these teaching aids can be also used for creating education facilities for other subject fields in various languages.

TatSoft 2 software package supporting the Tatar language in ICTs includes all the above-mentioned tools for working with the Tatar language and can be easily installed by any user.

Our long-term activities are the following.

Creating an intellectual multilingual search engine. This initiative, facilitating the creation of an electronic Tatar corpus, is conditioned upon the current linguistic situation in the republic and upon new emerging linguistic and intellectual technologies for multilingual search based on a thorough word sense disambiguation. Regarding that in some developed countries several languages enjoy official status, the project can become in demand for further commercial use.

Developing speech recognition software for Tatar. This action line is especially important as speech technologies are expected to be among the top trends of IT development in the coming years and ASR is to be widely implemented in major economy sectors.

Creating machine translators. A Tatar-Russian machine translator will enable access to English online databases through Russian equivalents, thus supporting equality of Tatar and Russian as official languages of the Republic. Machine translators for Tatar and other Turkic languages will facilitate the convergence of kindred languages speakers. The task of creating such programmes is rather easily solved due to the affinity of languages.

Using Cognitive Potential of the Tatar Language for Creating New Information Technologies

This area of research is related to the crucial task of developing intelligent operating systems and software tools using the potential of natural languages, their semantic and syntactic structures, as well as vocabulary.

Four factors are essential for computer technologies, namely *information processing time*, *memory capacity* for data storage, *active character of knowledge* and the *ability to give fuzzy instructions* (unambiguous in a certain context). The latter two properties are of critical importance for intelligent systems and technologies. Research in this field is a burning issue. In case

we identify structures, circuits, and formulas, implementing these properties in natural languages we can use them while creating artificial languages and programming systems, as well as other means of information description, storage and processing.

It is commonly known that operating systems, programming languages, information processing tools and almost all software used today are English-based. Therefore, they are based on Western mentality.

English as a fusional analytic language has almost zero morphology (compared with agglutinative languages). Complex meaning is communicated with phrases, which requires a much more complicated analysis and, consequently, an increase in the amount of memory and time required to process information.

The only way out of this situation is eliminating wider context and complex structures and, ultimately, simplifying the meaning and semantics. Thus, the basis of the English language itself brings computing systems to a deadlock, causing them not to get “smarter”, but to increase system performance and memory capacity, i.e. develop functional characteristics rather than “intelligence”.

Even the language structure and its syntax discourage the active character of knowledge. English is an SVO language, and it is not the information that dictates what action to take, what methods and algorithms to use for its processing. On the contrary, the means, the circuit, the algorithms force us to format, structure, and modify information.

Unlike Indo-European languages, Turkic languages refer to the SOV type and in this case it is information that comes first. Our research shows that due to regular morphology and natural complexity agglutinative Turkic languages, including Tatar, can be efficient tools for creating intelligent information processing systems. They provide for ultimate solutions in terms of information storage and processing. The meaning of a text is much more easily conveyed on the lexical level due to these languages’ ability to encode meaning synthetically, i.e. in a word form. Other types of languages, including English and Russian, have to convey the meaning by using several phrases or even groups of sentences.

Such features of the Tatar language as its agglutinative character, strong algorithmic patterns, minor exceptions, powerful meta-system, rather rigid syntax and active character of knowledge allow us to consider it to be a good basis for creating an intermediate language and even for developing new operating system based on a new ideology.

References

1. Heintz, J. and Schonig, C. Turkic morphology as regular language // Central Asianic Journal (CFJ), 1989. – P. 1-24.
2. Suleymanov, D. S. Natural cognitive mechanisms in the Tatar language // Cybernetics and Systems 2010 – Proceedings of the Twentieth European Meeting on Cybernetics and Systems Research / Trappel R.(ed.) – Austrian Society for Cybernetic Studies, Vienna, 2010. – P. 210-213.
3. Suleymanov, D. S. Natural possibilities of the Tatar morphology as a formal base of the NLP // Proceedings of the First International Workshop on Computerisation of Natural Languages (Varna, Sept. 3-7, 1999). – Sofia (Bulgaria): Information Services Plc, 1999. – P. 113.
4. Сулейманов Д. Ш. Естественные когнитивные механизмы в татарском языке // Компьютерная лингвистика и интеллектуальные технологии: труды Международного семинара «Диалог-2002» (Протвино, 6-11 июня 2002 г.): в 2 т. / Под ред. А. С. Нариньяни. – М.: Наука, 2002. – С. 500-507.

Sergey VASILIEV

*Deputy Director, Olonkho Institute,
North-Eastern Federal University
(Yakutsk, Russian Federation)*

Yakut Epic in Cyberspace: Olonkho Information System and Olonkho. Info Multilingual Portal

The preservation, research and promotion of cultural heritage require adequate information technology capabilities. Putting related information resources together and making them available to individuals and organizations involved in knowledge sharing was in 1989 set forth by UNESCO as one of the world community's priority tasks.

On November 25, 2005, the UN agency proclaimed the Yakut heroic epic Olonkho a Masterpiece of the Oral and Intangible Heritage of Humanity. This title is awarded to endangered expressions and practices of non-material cultural heritage in order to raise awareness on their value and encourage states and communities to safeguard them.

Yakutia's 2007-2015 State Target Programme to Safeguard, Study and Promote the Yakut Heroic Epic Olonkho provides for the development of an IT infrastructure for preserving, studying and promoting this and other treasures of intangible and oral heritage.

The Olonkho system designed as part of the programme is an innovative academic project aimed at creating a modern environment for education and scientific research. It can be implemented only as a collaborative effort between government agencies, epic scholars, linguists, IT specialists, translators, archivists, and so on.

Olonkho Information System (IS)

The Olonkho information system is aimed at gathering, processing, preserving, actualizing, spreading, representing and using relevant content. It should provide tools for formalizing, structuring and sharing knowledge used in research, education and archival work, as well as for standardising frequent work procedures and maintaining user-developer interaction. In this context, archival manuscripts, Olonkho texts, recitation audio and video, photographs, scholarly publications, and other related materials are seen as data arrays.

Olonkho IS has been designed to meet specific user interests and needs within its target audience. Researchers will be above all interested in visiting the IS online library, with complete Olonkho texts, scholarly publications, digitized audio archives, graphic images, manuscripts of Olonkho and other folklore pieces, specialized dictionaries, encyclopaedias, and reference books; of primary importance to them is to be able to spread the results of their research and to find necessary contacts and links. Museum and archive personnel may find particularly useful the possibility of creating digital facsimile archives of their collections and making them available to the public. To students and teaching staff, Olonkho IS offers an exciting opportunity to create a modern learning infrastructure. The Olonkho system aims to define the vector of professional growth and competency development in the academia, stimulating research among teaching staff and students alike.

Only if united into a single system, information resources could acquire new qualities (this is what philosophers call the notion of emergence, with emergent properties defined as effects that are not sums of the effects of each causal conjunct).

The recitation of an epic is, by its very nature, a multimedia event. Various elements enter into play here, such as sound (speaking and singing), visual presentation (mimics, gestures and posturing), and environment (e.g. chiaroscuro on the auditorium walls, the breath of people in the audience, their spontaneous reactions, etc).

In cyberspace, audio, video, original texts with translations, synopses, scholarly commentaries, dictionary entries on epic characters and archaic vocabulary, photos, graphics, links and footnotes can and should all be presented in a common hypermedia environment. Put together, these and other information sources will produce a cumulative effect unattainable if each is used separately.

Olonkho IS will create a common information and communications environment for various educational, scientific and cultural institutions on the basis of specific agreements that take into account appropriate technological requirements, protocol regulations and, crucially, the property and copy rights of the parties involved.

The system is designed to accumulate information in designated storages and record all necessary dimensions in databases. Digitized resources can be accessed through the Internet or an Intranet portal.

The system has several key segments.

Its ***Information Gathering & Processing subsystem*** consists of the following units:

- Olonkho: Archives, responsible for digitizing manuscripts, drawings, music scores, and other graphic images;
- Olonkho: Audio, Olonkho: Photo, Olonkho: Video, to be used in transforming original information resources into the digital format;
- Olonkho: E-Lib, to scan, identify, and collect textual documents for the IS electronic library;
- Olonkho: Documents, to support an organization's electronic document turnover (agreements, standards, official correspondence, technical specifications, reports, etc.).

The IS units' output will be presented in files of an appropriate format. Phenomena, objects and processes of the real world will acquire digital copies in a multidimensional model of a specific subject area.

Each particular file – and, if necessary, its components – will come with metadata, or specially arranged information about the file and its content, as well as the formal attributes and a scientific description of the digital objects carried. Metadata are needed to make the availability of digital content visible to a potential user, as well as to administer the saving of documents and register their reliability, technical specifications, mode of access, user responsibility, context, timeline, and conservation purposes. Metadata shall be carried in the extended Dublin Core format; digital objects are to be described using the XML language and texts saved in Unicode for an adequate script representation.

The Information Storage & Arrangement subsystem consists of two main units: Olonkho: Storage (a storage of digital objects) and Olonkho: Database (a database to be consulted by IS users while searching for objects they need). The base is to stock relevant information from the Information Gathering & Processing subsystem, metadata on a subject area model, and so on.

This subsystem's functions include:

- storing and processing textual documents in the original language, as well as video and photographic images;
- storing alternative recordings and/or versions of documents (digital objects);
- processing multiversion documents and their attributes;
- modelling new entities with non-predetermined sets of attributes;

-
- modifying existing connections and creating new ones between IS entities;
 - searching for and retrieving information with inter-entity connections in mind.

The Information & Technical Servicing subsystem is responsible for technical support of the project and for software and information servicing of corporate users (units Olonkho: Software & Information Support; Olonkho: Technical Servicing).

The IS information resources are built using licensed or open-source software. Yakut language software applications will ensure correct operations of the system's e-library, search engines, and multilingual database. Technical and semantic integration of the various constituent information resources will enable the entire system to operate more effectively.

There is a need to develop thesauruses and curated dictionaries for information resource metadata compiling, as well as to find appropriate technology for building digital collections.

Technical servicing will provide proper maintenance for corporate users' computer and office equipment, with their own maintenance departments often operating ineffectively, if at all. The protection of computer networks and PCs from malware is one of the priority tasks to be performed within the Technical Servicing subsystem.

The Information Presentation subsystem's function is to make information available to IS users. This segment consists of the following units: Olonkho: Hypermedia, Olonkho: Science, Olonkho: Education, Olonkho: Internet Portal, and Olonkho: Television.

Olonkho: Hypermedia is responsible for creating multimedia information resources and representing them correctly in the Web and as CD and DVD editions.

Olonkho: Science and **Olonkho: Education** are intended for retrieving relevant content from the database and organizing it into specialized theme packages for scientific research or for school/university curricula. The description formats SCORM (general module) and LOM (academic module) are helpful in building teaching resources. The employment of these formats will facilitate the search, study, assessment and use of educational objects by students, teaching staff and (in the long run) by teaching software applications. International standards will enhance the mobility and transferability of courses, facilitating content sharing between partner organizations.

Olonkho: Television allows to broadcast TV programmes through cable networks or the IP network (in case of Internet broadcasting).

Portal of Olonkho Information System

The Olonkho.Info Portal (www.olonkho.info) links IS users to the system's resources with the help of standard navigation tools. It also functions as an Intranet portal for corporate users. Network access is provided on the basis of user authentication and authorization, in compliance with applicable access policies. The IS developers are committed to ensuring the availability and protection of cultural heritage within the effective regulatory framework.

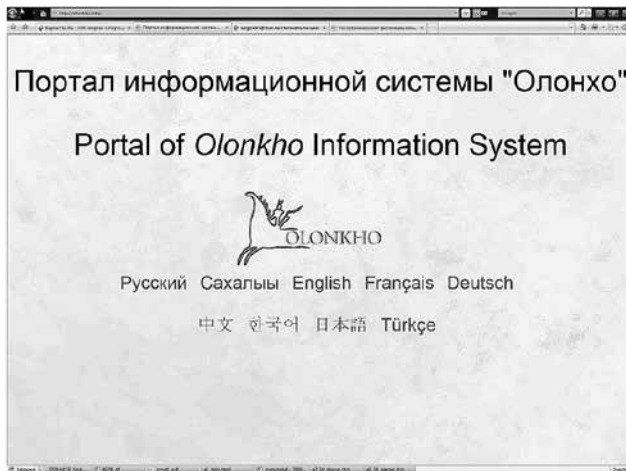


Fig. 1. Pages of the Portal of Olonkho Information System

At <http://olonkho.info>, information is available through interfaces in languages such as English, German, French, Japanese, Korean, Chinese and Turkish. Most of the textual content comes in Yakut and Russian. The website's philosophy proclaims the principle of language equality, meaning that information resources can be accessed through any of the existing language interfaces.

For example, a Yakut-language text will be presented in each of the language subsystems with its title in the original and in translation, as well as with a note about the original language, a reference to the source, a hyperreference, and, if possible, a translation of the full text or its synopsis.

Multimedia files should come with descriptions in all languages; metadata for all the information resources should also be multilingual. The user will thus be able to receive information in all available languages without having to leave his or her specific linguistic environment on the Web.

The Olonkho.Info Portal has a developed CMS-content management system, which enables authenticated users to correct already existing resources and contribute new ones.

Such a website is possible to build only in the environment of a university that can offer a powerful IT infrastructure as well as competent personnel, including computer programmers, folklorists, translators, and bibliographers.

The Olonkho.Info Portal has significantly enhanced its capacities thanks to the 2010-2011 Development Programme for the North-Eastern University, previously known as the University of Yakutsk.

The content it features includes scholarly and non-scholarly publications; Olonkho texts; profiles of Olonkho reciters and academics specializing in epic heritage; digitized manuscripts; archival audio recordings; video footage of Olonkho contests, including among young narrators; voiced dictionaries; class pages; press; teaching manuals; methodology literature; links to related websites and organizations, etc.

Information Technology for Preserving and Promoting Cultural Heritage

In our efforts to preserve linguistic and cultural diversity, we should apply information technology as extensively as possible. Subprojects implemented as part of the Olonkho IS project provide various examples of how IT could serve the purpose.

Computer technology may be an especially powerful motivation tool for young Olonkho reciters. Through the feedback it provides, they will be able to see that their heritage popularization efforts resonate with a wide audience.

Here are some of the Olonkho-related projects sponsored under the North-Eastern University's Development Programme.

The ***cultural heritage television channel Olonkho*** broadcasts through Yakutsk's Helios cable network. It airs educational programmes and culture content 16 hours per day.

To be launched soon are the channels Olonkho HD (high-definition video) and Olonkho 3D (three-dimensional stereo).

Internet television (<http://olonkho.info/InternetTV/>) and ***Internet radio*** (www.olonkho.info/internet_TV) allow to watch and listen to epic and other folkloric narrative recitations (recent and archival alike) any time of the day or night -- something that traditional media cannot possibly provide.

All events held as part of the Epic Heritage Archiving project are now filmed in FullHD. This format provides footage of excellent quality, with images five times as large as ones obtainable with analogue television systems such as PAL or SECAM.

Used in teaching tools, digital technology may dramatically enhance the learning effect. Interactive maps and drawings stimulate our cognitive instinct more than conventional ones do, enabling us to learn more easily and effectively.

3D visualization technology allows us to better feel the energy of a live event, such as a folk festival, and to identify with it.

Reinforced with the multi-channel 5.1 or 7.1 sound, a ***3D video*** makes a much more powerful experience than two-dimensional footage.

3D scanners and printers make it possible to produce a digital three-dimensional copy of an artefact and to then print it out for academic use.

The North-Eastern University has already acquired some practical experience in using 3D technology in educational projects.

Indeed, modern life requires that research and academic materials be broadly represented in cyberspace. Olonkho IS, along with its Olonkho Info portal, should try to meet this challenge. For the Olonkho state target programme, this is a strategic life-support system as well as a modern networking hub, which can facilitate and enhance efforts to preserve, study and promote the Yakut epic.

Lyubov BADMAYEVA

*Head of the Department of Linguistics,
Institute of Mongolian, Buddhist and Tibetan Studies,
Russian Academy of Sciences (Siberian Branch)
(UlanUde, Russian Federation)*

Yulia ABAYEVA

*Research Fellow, Department of Linguistics,
Institute of Mongolian, Buddhist and Tibetan Studies,
Russian Academy of Sciences (Siberian Branch)
(UlanUde, Russian Federation)*

Modern Linguistic Technology to Help Preserve Buryat Language: Corpora & Database Developments⁷⁷

Modern information technology provides linguists with broad opportunities both for doing research and building textual corpora and databases online. Such opportunities have been efficiently used within the framework of corpus linguistics for several decades now. The accessibility of open-source content on the Internet can facilitate and speed up the material-gathering routine while bringing out new aspects to be explored.

Nowadays, corpus technology offers excellent opportunities for representing any language online, for purposes such as research, education and, of course, preservation.

In our age of globalization, smaller languages often find themselves sidelined, and Buryat – one of Russia’s minor languages – is no exception. Bringing such languages online in open-access corpora and databases may prove an effective solution.

The promotion of language diversity in cyberspace could also help preserve native speaker communities’ distinctiveness and identity, so developing online linguistic resources for small and community languages should be seen as an important state-level task.

Adopted by Buryatia’s government in August 2010, the republican programme “Preservation and Advancement of the Buryat Language in 2011-2014” hails the idea of building a Buryat textual corpus, emphasizing the feasibility of supporting suchlike initiatives.

The Mongol, Buddhist and Tibetan Studies Institute’s Linguistics Department, which is leading the project, has already compiled a corpus of Buryat-language

⁷⁷ Work has been done with support from the Fundamental Linguistic Research Foundation (<http://www.flli.ru>), Project S-43.

texts totalling some two million word uses. The texts span all major styles of modern Buryat, including fiction, academic writing, and news.

In keeping with the standards of the Text Encoding Initiative (TEI)⁷⁸, all the corpora should go through corpus-header markup.

A newly-launched pilot version of the site <http://corpora.imbtarchive.ru/index.php> features selected materials, such as concordances⁷⁹ to Buryat fiction. These have been built using original software developed by the project team. Contexts bring out the meanings of designated word forms, facilitating their subsequent breakdown across grammar categories for analysis. The corpus is to be regularly supplemented, updated and, if necessary, corrected/modified.

The project's preliminary results have been reported in scholarly paper collections, symposia proceedings, and on the Internet.

Representing a language on the Web and making related resources accessible to the research community should be a key priority with modern-day linguists.

One of the main aims behind the Buryat corpus project is to integrate the language into the global information environment. This aim is outlined in the Russian Academy of Sciences Presidium's Corpus Linguistics programme of fundamental research (Direction 3. Creation and development of corpora resources on Russia's languages; <http://www.corpling-ran.ru/n3.html>). Being implemented as part of this programme and with its support, the project is aimed at modelling a morphological description of the Buryat language, which would pave the way for the development of a morphology parser.

The Institute of Mongol, Buddhist and Tibetan Studies is also involved with experimental developments in Buryat phonetics. Its current projects are aimed at building a database from information collected ever since the Linguistics Department's Experimental Phonetics Research Laboratory was set up. These developments could subsequently be incorporated into the Buryat corpus as an oral speech subcorpus.

Speech databases as a major type of linguistic resources are, per se, of much research interest. Such bases are essential to scholarly tasks related to the analysis and description of oral speech. Building large, wide-ranging and informative (multitier) speech databases, along with an easy-to-use and reliable set of tools for their development and employment is an increasingly important task, of relevance for computer applications and for fundamental phonetic research alike.

⁷⁸ *The TEI's aim is to develop standardized methods for marking textual resources. [Editor's note.]*

⁷⁹ *Concordance is a list of examples of the use of a particular word in context, as sourced from a textual corpus, complete with links to the source. [Ed.]*

Our efforts to create a speech database for the Buryat language are being made with due account for its regional varieties. Most of Russia's Buryat speakers live in the Republic of Buryatia, the Trans-Baikal region (the Aginskoye area), and in the Irkutsk Region's Ust-Ordynsky area; there are also large Buryat communities in Mongolia (specifically in the Dornod, Khentii, Selenge, and Khovsgol provinces, known locally as "aimags") as well as in northeastern China (Hulunbuir, in the Inner Mongolia Autonomy). The Buryat ethnicity's spread across vast territories in three countries, as well as its lack of homogeneity, has resulted in the language's broad dialectal variation of every level: segmental, suprasegmental, morphological, and lexical.

The project's ambition is to preserve the distinctive regional features of the Buryat language, designing strategies for the creation of oral speech databases, and systematizing and putting into circulation the accumulated audio content.

Samples of Buryat speech featured on the database should be used for further exploration of its phonetic and prosodic structure and its morphological and lexical characteristics.

Corpus methods appear the most appropriate for the purpose as they allow to comprehensively represent a large, versatile array of data – with due account for the various characteristics of speech fragments, ranging from acoustic to discursive.

The would-be database is to include separate words, sentences with varied communicative purport, and coherent texts. The speech signals will each come with a transliteration and a phonetic/prosodic transcription. There will also be notes on idiosyncratic or unusual pronunciations and on emotionally coloured speech fragments, along with some background information on the speaker.

The project involves recording speech samples and arranging them in the form of audio files. The technical groundwork will consist in the digitization of audio recordings and their multi-layer segmentation (into phrases, syntagmata, words, and sounds), along with textological decoding.

As a result, each recording should be provided with an audio file carrying various segmentation markups, as well as with textual files that are transliterations or phonetic conversions of the recorded material.

Fellows of the Mongol, Buddhist and Tibetan Studies Institute's Linguistics Department have by now assembled ample audio content on the standard Buryat language and its dialects, as well as on other Mongol languages, such as Daghur, Baerhu, and Khalkha Mongolian. Systematized, homogenized and arranged in a database, that material will allow to preserve the distinctive speech character of regional Buryat communities, which is now being erased by the growing influence of media language as well as by the shrinking use of the Buryat language itself owing to extralinguistic factors.

Anatoly ZHOZHNIKOV

*Director, New Information Technologies Centre,
North-Eastern Federal University
(Yakutsk, Russian Federation)*

Svetlana ZHOZHNIKOVA

*Programme Supervisor, New Information Technologies Centre,
North-Eastern Federal University
(Yakutsk, Russian Federation)*

Creating a Web Portal on Indigenous Cultures of North-Eastern Russia

In an age of globalization, the world is turning into one big mechanism, with all its parts interconnected. Innovative information and communications technologies push this process forward; they have a unifying role to play in building a new world order, opening up, as they do, new possibilities for the preservation of the languages and cultures of all ethnicities populating Planet Earth, including the small indigenous communities in Siberia and the circumpolar North.

Lots of factors may put the existence of a language at risk: grave natural disasters wiping out entire communities; faulty public education systems providing no, or scarce, opportunities for schoolchildren to study in their native languages; or the lack of an adequate writing system, to give just a few examples.

One more factor to have emerged in recent decades has to do with information and communications technology. The Internet offers plenty of opportunities for exercising one's right to free speech and getting access to information and quality education.

One serious problem, though, is that information and services in cyberspace are still available only in a limited number of languages (just 400 of the world's 6,700 currently spoken languages are represented online so far).

At the outset of the 21st century (on November 2, 2001, to be precise), UNESCO adopted its Universal Declaration on Cultural Diversity, postulating that respect for diversity, tolerance, dialogue and cooperation in a climate of mutual trust and understanding are the best safeguards of peace and international security. Being a source of exchange, innovation and creation, cultural diversity is as important for the human race as biodiversity is for wildlife.

The new century and millennium are seeing the formation of an integral socio-cultural system that could serve as the basis for dialogue and interaction between cultures and faiths. This system is expected to shape a worldview for the generations to come and to determine their sustainable development patterns.

The international conference on cultural and linguistic diversity in cyberspace that took place in the Republic of Sakha's capital, Yakutsk, on July 2-4, 2008, adopted, for its turn, a final document that became known as the Lena Resolution. The Resolution urges to foster linguistic and cultural diversity on the Web in every possible way and to continue efforts to record, preserve and advance various languages, especially smaller ones, with the help of modern ICT.

A recent survey of the content available on the Internet about the indigenous communities of Russia's Siberia and circumpolar North has shown that the existing sites are disparate and not informative enough while fully systematized, comprehensive resources aren't there yet.

Quite an exhaustive analysis of related content has been carried out by A. Burykin in his essay "Internet Resources on the Languages of Small Indigenous Peoples of Russia's North, Siberia and Far East: Content Overview and User Enquiries" [1]. The problem of supporting multilingualism in specific Russian regions as well as nationwide has been the focus of numerous publications by the Russian Committee of the UNESCO Information for All Programme (IFAP) and the Interregional Library Cooperation Centre [2, 3]. In practice, though, little progress has been made. Hence the idea to build a Portal on North-Eastern Russia's indigenous communities. Such a project should seek to ensure the linguistic and cultural diversity of those communities is adequately represented on the Internet, as well as to create an online environment where they could communicate in their native languages.

This prospective Web site will provide wide-ranging information on the languages and cultures of all indigenous communities populating Russia's North-Eastern regions. The choice of that particular geographic area arises from its being under the jurisdiction of the Amosov North-Eastern Federal University, the driving force behind the project.

Not all the indigenous language scripts are yet part of computer operating systems. Efforts will therefore be made to bring all missing scripts onboard, introducing them first as part of a universal keyboard layout. In a longer term, though, language-specific layouts are to be created. The online launch of scripts for all indigenous languages spoken in the designated area is expected to raise their profile on the global information scene while also offering possibilities for communicating in native languages online. This will

facilitate both the preservation of endangered community languages and their advancement.

The new Web portal is to be presented in indigenous languages as well as in Russian and English. Each of the relevant communities will have a separate section devoted to it, carrying informative content on its language and culture. In addition, a forum for communication in indigenous languages will be created.

The project's implementation has got underway as part of the Amosov North-Eastern Federal University's "Programme for Preserving and Advancing the Yukaghir Language & Culture on Digital Carriers and in Cyberspace in 2011-2014" (Project 4.1, Event 2.35). The reason for picking out the Yukaghirs is their being an endangered culture on the brink of extinction, with very few native speakers and tradition carriers left.

Basic information on the Yukaghir is now available at www.arcticmegapedia.ru. It comes in the following 16 categories:

- General Information,
- History,
- Yukaghir Language,
- Myths & Folklore,
- Customs & Beliefs,
- Matrimony & Family Raising,
- Traditional Economic Activities,
- Housing & Home Utensils,
- Clothing,
- Music & Decorative Arts,
- Traditional Cuisine,
- Transportation,
- Yukaghirs Today,
- Scientific Research,
- Literature,
- Personalities.

An English version of the site is already in place, along with an overview in the Yukagir language. A Web forum will be created in a while, for Yukaghirs to be able to communicate in their native language online.

The Yukagir section is the curtain-opener for this new Web portal, which will soon expand to include information on all other indigenous communities of North-Eastern Russia.

References

1. Языковое разнообразие в киберпространстве: российский и зарубежный опыт. Сборник аналитических материалов / Сост.: Е. И. Кузьмин, Е. В. Плыс. – М.: МЦБС, 2008. С.111-129.
2. Многоязычие в России: региональные аспекты. – М.: Межрегиональный центр библиотечного сотрудничества, 2008. – 136 с.
3. Развитие многоязычия в киберпространстве: пособие для библиотек / Авторы-составители: Кузьмин Е. И., Паршакова А. В. – М.: Межрегиональный центр библиотечного сотрудничества, 2011. – 128 с.

SECTION 2. INSTITUTES TO PROMOTE LINGUISTIC AND CULTURAL DIVERSITY

Jan Henry KESKITALO

*Vice-President Indigenous, University of the Arctic;
Member, Commission of Norway for UNESCO;
Executive Chair, World Indigenous Nations Higher Education Consortium
(Alta, Norway)*

University of the Arctic's Contribution to Maintaining Linguistic and Cultural Diversity in the Arctic: Improvements Through Networking and Partnership

What is the University of the Arctic?

The University of the Arctic, or as referred to in popular form: UArctic, is a cooperative network consisting of universities, colleges, and other organizations in the Arctic committed to higher education and research in the North. The UArctic members through their network share resources, facilities, and expertise to provide post-secondary education that is relevant, and accessible to students as well as communities of the North.

For this particular presentation, I will combine general information about the University of the Arctic; its vision, goals, priorities and strategy with more specific focus on UArctic priorities regarding the indigenous peoples of the Arctic. This will cover how indigenous peoples in the Arctic involve with the network, and what UArctic services and activities could possibly support the further development of the priorities of the indigenous peoples. The linguistic and cultural integrity of the indigenous peoples, and how this is served by the higher education system, is what contributes to a true linguistic and cultural diversity from the viewpoints and needs of the indigenous peoples.

UArctic is founded partly by initiative of, and strongly in principle supported by, the Arctic Council. UArctic already from the early initiative of establishment works in close partnership with indigenous peoples recognizing their integral role in northern education, in training, in knowledge generation and transfer. UArctic practices involvement of indigenous peoples in its governance, in planning, and seeks to engage indigenous principles, perspectives and participation in its operations. Our overall goal is to create a strong, sustainable circumpolar region by empowering northerners in general, and indigenous

peoples, through education, training and shared knowledge. UArctic also seeks to promote excellence in knowledge generation and knowledge application in areas relevant to the North.

UArctic Members and Strategy

The UArctic membership body consists of members from the circumpolar area throughout all 8 Arctic Council member states, as well as associate members outside the Arctic that has strong interests in education and research in the Arctic.

Currently UArctic has more than 130 members; together members have approximately 1 billion students and more than 70 thousand academic faculty. Even if not all students or faculty are directly engaged in UArctic activities, the network as such represents a huge potential for regional development in the Arctic.

With regard to the indigenous profile, many UArctic members have strong ties to indigenous communities. Many of them offer relevant academic programmes and have designated departments and research programmes specifically dedicated to service their indigenous constituency. Besides this, among the UArctic members we also find smaller institutions that specialize on serving the needs of indigenous peoples highly focusing specifically on human, social and cultural development.

With the current vision and goals UArctic represents a huge potential for further development of the cultural and linguistic diversity in the North, when carefully planned and carried out in practice. Visions and strategies however need implementation into institutional cultures and individual practices and the results need to be identified at the receivers' end. It is only when students, researchers and the northern communities together experience the fulfillment of their goals and aspiration, that we have a good indication of success of the added value of network.

As already initially indicated, University of the Arctic as a network was established *inter alia* to create an improved and expanded platform for post-secondary education for indigenous peoples of the Arctic.

This is clearly stated by the University of the Arctic Governing Bodies, and the strategic plan for the 2009–2013 provides evidence in this regard. It states a clear goal, and a clear strategy: UArctic shall look to increase leadership of indigenous peoples in the operation and governance, as well as programmatic activities of UArctic. Let us look at the following quotation from the UArctic strategic plan 2009–2013:

“Vision 2013

By 2013...

- UArctic will have increased relevant training, higher education, and knowledge generation and application in the North with clear socio-economic benefits – particularly to remote communities and indigenous peoples.
- UArctic’s innovative programs will have a significant impact on increasing the level of education in the Circumpolar North, and generate highly qualified people in Northern communities by providing career bridging opportunities.
- UArctic, through its members, will have a global leading role on Building Human Capacity in the North, Adaptation to Climate Change in the North, and Energy in the North from technical, cultural and economic as well as environmental perspectives.
- UArctic’s member institutions will be committed to the implementation of a common set of activities as outlined in the UArctic Charter.
- UArctic will have enabled increased capacity in education, training, knowledge generation, and knowledge application for member institutions through its collaborative framework.
- UArctic will be recognized as the body that carries forward the Arctic IPY Training and Higher Education Legacy.
- UArctic will have, through partnering with other stakeholders in the Circumpolar North, ensured a stronger voice for the North globally”.

Further, among the specific goals in the strategic plan we find the following examples to be fulfilled by 2013:

“Specific goals for 2013...

- Indigenous peoples and northerners will continue to have a well-defined prominent role in the leadership and development of UArctic.
- Opportunities will be created to facilitate online and local access to UArctic curriculum for indigenous and other students in Northern communities.
- UArctic will continue to operate in close partnership with national and local governments, including indigenous peoples’ governments and organizations, and the private sector.”

These short examples help demonstrate how UArctic formulates its mandate and obligation to serve the indigenous communities through its operations.

We are glad that UArctic education is circumpolar, holistic, and diverse in nature, and draws on our combined members' strengths to address the unique challenges of the region.

What do we mean by circumpolar? UArctic promotes a northern voice in the globalizing world that reflects a shared regional identity across all eight Arctic states and among all Northern peoples and cultures.

What do we mean by holistic? UArctic promotes understanding among different learning systems, bringing together perspectives from the arts and both social and natural sciences with traditional and indigenous knowledge.

What do we mean by diverse? UArctic promotes cultural diversity, language plurality, and gender equality

UArctic Governing Bodies

The University of the Arctic has the following governing bodies.

The Board of Governors is the highest decision making body of UArctic responsible for strategic development, institutional priorities and finances. It consists of an internationally representative Board elected by the Council.

The nomination and election process is based on overall UArctic policy to represent diversity and this secures the blend of representativity that covers also the indigenous populations of the Arctic.

The UArctic Council consists of representatives from all UArctic members (over 130). It decides on membership of UArctic and nominates members of the Board, makes strategic decisions on UArctic' programme development and gives academic guidance for programme implementation and delivery.

The council meets once a year to run a business schedule combined with break-out sessions focusing on themes of importance for the development of the network. The council also elects members to an Indigenous Issues Committee among its members. The committee oversees that UArctic practices the visions, strategy and goals of indigenous involvement in programmes and other activities. This particular committee is connected directly to the Vice-president Indigenous of the network, who acts as the secretary of the committee and is the link between the UArctic leadership and the committee.

University of the Arctic also has a forum for the leadership of the academic member institutions. The ***UArctic Rectors' Forum***, held once a year, brings

together university and college Presidents, Rectors, Provosts, Chancellors as well as Vice-Presidents around specific themes. In 2008 the rectors signed the UArctic Charter. The idea behind the charter was to have a mechanism for UArctic members to show increased commitment to UArctic activities.

By signing the charter the institutions agreed among other things to the following:

“UArctic recognizes the integral role of indigenous peoples in northern education, and seeks to engage their perspectives in all of its activities. UArctic and its member institutions further respect the needs of the indigenous peoples, and commit themselves to actively include the needs of the indigenous peoples and indigenous communities of the Arctic in education and training.”

This statement is very promising with regard to having UArctic as a strong engine to create good possibilities to maintain the cultural and linguistic diversity in the Arctic.

During this strategic period 2009–2013 UArctic will refine the Rector’s Forum as a venue for the leaders of the UArctic institutions to engage, jointly, in facilitation of development in the north.

Strategic Areas

UArctic is organized into seven strategic areas, each consisting of one or several programs.

These include:

1) Shared Focus –Thematic Networks

The thematic networks are a mechanism for building partnership among members. Networks provide a structure for facilitating student and faculty mobility and collaboration. A Thematic Network is a group of UArctic members working together on subjects of shared interest to create learning experiences for students, faculty and communities. This includes student and faculty exchange and curriculum development.

2) Shared Learning –Undergraduate Studies

The Undergraduate Studies programmes at UArctic focus on building human capital, and knowledge about the North among its residents and others. They are characterized by the expansion of offerings across members to develop fully operational Circumpolar Studies Program (BCS); constant revision and review of core curriculum; high-quality programming by continuing to engage external academic control mechanisms; development of relevant Advanced

Emphasis; close cooperation with Graduate Studies Strategic Areas to increase BCS graduates in UArctic graduate studies.

3) Shared Competence –Graduate Studies

UArctic Graduate Studies focus on the development of high-quality, joint and parallel graduate and postgraduate programme activities, which provide UArctic members with opportunities to develop and support relevant thematic networks.

Programs include:

- Master Studies
- UArctic Field School
- Graduate Networks.

4) Shared Knowledge –Research and Development

Shared Knowledge means that:

- UArctic members are engaged in research, education, and development activities with each other.
- UArctic activities are coordinated with Arctic research.
- UArctic activities are relevant to Arctic Council working groups, Indigenous peoples organizations, and Arctic science and development organizations.

5) Shared Experiences –Mobility Programmes

North2north student exchange programme allows students at UArctic Institutions to visit different northern regions, and share experiences face-to-face, through study at other UArctic institutions. Mobility grants are provided for 3–12 months of study.

Mobility programmes ensure the facilitation of student exchange, focus on best practices, increased funding of mobility, enhancement of knowledge about northerners and building of shared northern identity, motivation of all major scholarships to support north2north exchanges.

6) Shared Resources –Services to Members

Shared Resources means that UArctic builds capacity within members by providing services through the added value of the network, promotes UArctic members as study destinations, and promotes the North as a subject of study.

Services to members include:

- UArctic Information Service

-
- UArctic Studies Catalogue
 - GoNorth Programme
 - UArctic Press
 - UArctic Institutes.

7) Shared Organization –UArctic Organization

UArctic administration is light and decentralized, daily operations are distributed between:

- International Secretariat at University of Lapland, Rovaniemi, Finland
- UArctic President's Office, GRID Arendal, Norway
- Vice-President Indigenous office at Sámi University College in Guovdageaidnu/Kautokeino, Norway
- UArctic Russian Information Center, at the North-Eastern Federal University, Yakutsk, Russia
- UArctic International Academic Office, Northlands College, Saskatchewan, Canada
- UArctic Program Offices around in different locations in the Arctic.

New Initiatives

The University of the Arctic has recently accepted the invitation to establish the UArctic Research Office located at Northern (Arctic) Federal University in Archangelsk, Russia. This research office will coordinate the UArctic international research cooperation connected to UArctic international Arctic partners in research and research by the thematic networks and UArctic institutes.

Programs are coordinated by offices hosted by member institutions and located around the Circumpolar North. UArctic plan to arrange with the North-Eastern Federal University to establish the office of the leadership of the UArctic undergraduate studies. The University of the Arctic hopes that organizing the undergraduate studies leadership will give a good opportunity to expand with thematic networks and Bachelor of Circumpolar studies and other activities to future associate members further east.

The Members as the Core Resources

The University of the Arctic, as a network, always depended on its member institutions commitment to succeed as a network for the benefit of indigenous

communities. The success of the programmes, research and service depends on members' cooperation with indigenous communities and stakeholders. The next decade calls for improvements in this regards. UArctic 2011 Rectors' Forum and Student Forum declarations carried important messages about the need for continuous development. Both directly and indirectly the declarations reflect the need for enhanced activities that will foster more capacity to work with safeguarding and further development of the regions including indigenous languages, culture and knowledge. This underlines what the previous Rector's Forum declarations already stated, and is in line with the overall UArctic goals and strategy.

The next step is to follow up on the operations level. This includes converting what there was into robust organizational structures. This challenges the network, but even more the single members themselves. Even if the network agrees of strong visions and strategic choices, the fulfillment of those depends on well developed institutional cultures and the practices of each one facilitating research and training. It is what happens at the delivery end of the line, as experienced improvements by students and community that counts as added value.

UArctic aims at getting an NGO status and also realizes that to achieve such status there is a need for our member institutions and UArctic combined activities to add even more cooperation with UNESCO.

We will integrate e-Learning/flexible learning platforms (e.g. telemedicine, social networking) into health and education to increase access to information and formal/informal education.

High-speed internet is a prerequisite for establishing and maintaining interconnectivity between Arctic communities. Once all communities across the circumpolar north have access to high-speed internet, UArctic will be virtually situated to support education and research. Northern societies need access to community-based educational programmes that will enable them to determine their own futures. Especially the small indigenous minority languages need robust infrastructure and state of the art technology adapted to the diverse linguistic situation.

During the UArctic council meeting in Yakutsk 2010 the VP Indigenous arranged a break-out session focusing on Ways to Improve Higher Education Services to Indigenous Peoples of the Arctic by Simultaneously Strengthening their Language and Culture.

We will give selected examples from the report to show what it focuses on.

“2. Teaching of language and culture and livelihood

- teacher training in teaching indigenous languages: so elements of language and culture can be passed along to younger generations
- development of teaching resources, including digital resources
- development of capacity within communities: teachers who are prepared and can contribute to transmission of language and culture
- laws and policies on language learning and education (this varies from jurisdiction to jurisdiction and involves varying levels of support (including non-existent) for indigenous language education)
- local, regional, national and international components: all must work together cohesively and towards the goal of language teaching
- role of local communities and parents: development of community-based approaches to indigenous language learning
- more demand for indigenous language learning than there are programs for (Yakutia example)
- linguistics can assist: need to communicate the research in meaningful ways to indigenous communities

3. Encouraging applied research on language and culture

- example: recording languages
- for some groups language and culture is on the brink of being lost; for others the languages are still very alive

8. Technologies

- connect development of indigenous languages/cultures to the development of digital technologies,
- access to new technologies: access to internet.”

In future UArctic developments, attention should be given to the interaction between language, culture and livelihood. Many of the traditional indigenous livelihood connected challenges are caused by no connection to modern virtual technology: technology is not adjusted to their real needs (like language compatibility), technology back up lacks, etc.

As we experience, there are a lot of challenges reported that need solutions before cyberspace can in an advanced manner service the needs of small languages. Another break-out session at the 2011 council meeting discussed possible indigenous principles for the future development and will soon be available on the UArctic web-site. The idea is to arrange at the annual council

meetings an indigenous forum where UArctic members can discuss and suggest possible ways of solving implementation gaps in UArctic activities to better fulfill the needs of the indigenous communities.

Arctic Social Indicators 2010

The 2010 Arctic Social indicators report (a project under the auspices of the Arctic Councils' Sustainable Development Working Group) presents a broad definition of culture that leads to a multidimensional understanding of cultural well-being and vitality. They identified for their purposes the following dimensions of culture (p. 92):

- Language (its use and retention),
- Knowledge (and its transmission),
- Communication (including education and performance),
- Spirituality, such as religion and ritual,
- Sociocultural events and media,
- Economic and subsistence practices,
- Social organization, institutions, and networks.

Such an approach connects valuable domains that together form the potential composite of knowledge for a people. It is important to facilitate the opportunities for the indigenous communities of the Arctic to maintain such dimensions as a community and a society. It is not an easy task; this is what represents a part of the linguistic and cultural diversity of the Arctic.

A holistic approach means to reach a level of balance between the preservation of domains that traditionally form the identity of the indigenous peoples and the expansion of modern technology and economy. However, such balance demands to form an integral platform for indigenous peoples to utilize possibilities in an integrated manner.

References

1. Arctic Social Indicators. TemaNord 2010:519. Copenhagen: Nordic Council of Ministers.
2. Draft Minutes of the 13th Council Meeting of the University of the Arctic. Yakutsk, Russia, Break-out session on Indigenous Issues, June 2-4, 2010.
3. UArctic Strategic Plan 2009-2013. UArctic Publication Series 3, UArctic International Secretariat. Rovaniemi: University of Lapland.

Sergei BAKYKIN

*Executive Director, Interregional Library Cooperation Centre;
Vice Chair, Russian Committee,
UNESCO Information for All Programme;
Council Member, Russian Library Association
(Moscow, Russian Federation)*

**Promoting Linguistic and Cultural Diversity in Cyberspace:
Activities by the Russian Committee
of the UNESCO Information for All Programme
and the Interregional Library Cooperation Centre**

The Russian Committee of the UNESCO Information for All Programme (IFAP) acts as an expert board for the Commission of the Russian Federation for UNESCO. It contains prominent librarians, archivists, museum specialists, scholars, politicians, public officers, representatives of non-governmental organizations and private companies. The Committee was established in 2000 to promote IFAP in Russia and to voice in UNESCO Russia's views on information policy.

Over the years of its existence, the Committee has held many events of the national and international level in every priority action line of the IFAP, i.e. information literacy, information preservation and accessibility, information ethics, and information for development. A number of reports on these issues have been prepared, over 60 books on the issues of knowledge society building are published.

The Russian Committee of the UNESCO Information for All Programme is not a legal entity, hence to provide its efficient operation, it needs a working body to elaborate specific projects and to look for the means of their implementation. Interregional Library Cooperation Centre (ILCC), a non-governmental organization with a legal status of "an interregional public organization", has become such a body.

ILCC was established in 1995, and since then it makes efforts in several directions. ILCC has taken active part in the elaboration and implementation of the national policy of reading promotion and is the developer of the National Programme for Reading Promotion and Development in Russia. Within this action field, since 2007 about 50 regional workshops and training sessions have been held, over 20 books covering philosophical and sociological aspects of the problem have been prepared and published, methodological recommendations for regional authorities, libraries, educational institutions, and mass media have been proposed.

The annual All-Russian Conference “National Programme for Reading Promotion and Development: Problems and Prospects” is held at Moscow’s most prestigious President Hotel. In 2010 we carried out an all-Russian monitoring of the related activities of the governmental authorities and major public libraries in all Russia’s 83 regions. The questionnaires comprised about 100 questions on the efforts towards reading development, stakeholders, existing problems, results achieved, etc. An analytical report was drawn upon based on the monitoring findings.

ILCC also participates in elaborating and implementing national library policy and collaborates closely with the Library and Archives Department of the Ministry of Culture of the Russian Federation.

ILCC acts as a coordinator of Russia’s National Programme for Analog Library Collections Preservation. We work in partnership with all major Russian federal and regional libraries. Numerous books on these issues have been prepared and published.

In 2010 the Ministry of Culture of the Russian Federation and the Federal Agency for Press and Mass Communications requested for analytical reports on the problems of digital information preservation in Russia. The reports presented by the ILCC cover organizational, legal, technological and personnel issues of this tremendous problem, that the whole world is facing today. At present, on the order of the Russian Ministry of Culture we are working on the National Programme for Russian Digital Library Collections Preservation.

We have delved into the issues of information accessibility and act as the coordinator of the Programme for Building the all-Russia Network of Public Centres for Legal and Socially Important Information. Approx. 7,000 such centres based in libraries operate in Russia today. In 2010, we carried out the monitoring of the network status and published an analytical report.

ILCC organizes major international and all-Russian conferences, workshops, and round tables in relevant fields. During the recent 6 years, over 100 such events have taken place. We have prepared and published over 60 titles of books, with 25,000 free copies delivered to scientific and public libraries in Russia, Ukraine, Belarus, Kazakhstan, Kyrgyzstan, Azerbaijan, and other countries.

Almost all projects are joint projects of the Interregional Library Cooperation Centre and the Russian Committee of the UNESCO Information for All Programme, as these two organizations are closely connected and augment each other’s efforts.

The activities of the Russian Committee of the UNESCO Information for All Programme are recognized in many countries of the world. That is confirmed by

the fact that in 2010, Evgeny Kuzmin, Committee Chair and ILCC's President, was elected the Chairman of the Intergovernmental Council of the UNESCO Information for All Programme. Today ILCC acts not only as the working body for the Russian Committee of the UNESCO Information for All Programme, but also as the Secretariat for the Intergovernmental Council Chairman.

Promoting linguistic diversity in cyberspace is a cross-cutting issue of the UNESCO Information for All Programme, and the Russian IFAP Committee and ILCC have been working in this field for about 5 years. On the order of the Commission of the Russian Federation for UNESCO we prepared national report on measures taken to implement the UNESCO Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace. We investigated the state of the art in the area in Russia and worldwide and recognized that not a single attempt had ever been made in Russia to conduct a systematic study and search for solution of the problem of language promotion in cyberspace, and especially from the political viewpoint, despite the fact that we have numerous highly experienced linguists and efficient politicians tackling the issues of cultural diversity.

We have translated into Russian and published the works by prominent world experts in the area of multilingualism and cultural diversity development, and have prepared a number of original publications:

- “Measuring Linguistic Diversity on the Internet” (the book by John Paolillo, Daniel Prado and Daniel Pimienta translated into Russian).
- “Comment assurer la présence d’une langue dans le cyberspace?” (Russian translation of the book by Marcel Diki-Kidiri comprises practical recommendations on facilitating the process of increasing the number of languages presented and used in cyberspace through developing linguistic and information resources, cultural components and supporting user communities).
- “Multilingualism in Russia: Regional aspects” (the book covers linguistic policies in several multinational and multilingual Russian regions, as well as activities by major regional libraries related to multilingualism promotion).
- “Preservation of Linguistic Diversity: Russian Experience” (the publication in English examines practical experience and efforts made at various political levels and by various institutions to support multilingualism).

-
- “Human Language Technologies for Europe” (the European Commission’s book translated into Russian investigates the current status, problems of machine translation and prospects for Europe).
 - “Representing the Languages of Russia and the CIS countries in the Russian Internet Segment” (the book comprises the papers presented at the international seminar held in 2007 by the Russian IFAP Committee and ILCC).

We started with seminars and later went on with the more representative events to promote this theme and managed to gather a pool of highly professional experts.

For better understanding of the efforts being taken in Russia to develop multilingualism in cyberspace we have carried out a target study. We developed two questionnaires of 40-50 questions each and distributed them among Russia’s leading universities and the Russian Federation constituent administrations. As a result, we have got a vast data array for us to analyze and learn who is doing something to support multilingualism on the Internet, where and what exactly. The findings are presented in the publication “Language Diversity in Cyberspace: Russian and Foreign Experience”.

The problem of multilingualism promotion concerns many parties – authorities, universities, libraries, and archives. However, they all need methodological support, analytical materials and information in Russian. For that reason, we have been gathering such information and posting it on the website of the Russian IFAP Committee <http://ifapcom.ru/en>.

We came forward with the initiative to hold an international conference on language and cultural diversity in cyberspace in Russia and got support of the Ministry of Culture of the Russian Federation, Yakut Government, and UNESCO Moscow Office. The Conference was held in Yakutsk in July 2008, with participants representing 15 countries and all the continents. The proceedings were published both in the Russian and English languages. The financial support for the publications was provided by the North-Eastern Federal University and the Ministry of Culture of the Russian Federation. The President of the Russian Federation also devoted a grant for these purposes.

The Conference became Russia’s contribution to the United Nations International Year of Languages. It helped us win the recognition and invaluable experience, get acquainted with many prominent Russian and foreign experts who are today our partners, participants in this conference, and some of them – the coorganizers. I am speaking about Adama Samassekou and Daniel Prado, heads of the MAAYA World Network for Linguistic Diversity and Latin Union.

The Conference final document –“The Lena Resolution” – has been widely recognized in the world.

After three years, we have gathered for the second international conference under the same name. We are happy to inform that, in accordance with the Lena Resolution recommendations, Centre to Advance Multilingualism in Cyberspace was established in 2010 and works efficiently under the North-Eastern Federal University. Three years ago we could hardly imagine that. Nevertheless, the second conference has got two groups of organizers – in Moscow, and in Yakutsk, where the main load of organizing the conference in Yakutsk is taken by the Centre.

Expecting the 2nd International Conference “Linguistic and Cultural Diversity in Cyberspace”, the book “Developing Multilingualism in Cyberspace: Guidelines for Libraries” was published with the support by the Ministry of Culture of the Russian Federation and UNESCO Moscow Office. The publication examines linguistic situation in Russia and efforts to support language diversity; potential action lines for libraries, global approaches, efforts made by international organizations in the sphere are described.

The 2nd Yakutsk conference gathered representatives of 33 countries and offered opportunities for finding new partners to initiate and implement new interesting projects.

The Russian Committee of the UNESCO Information for All Programme and the Interregional Library Cooperation Centre invite everyone for cooperation. Detailed information on our work and full-text versions of our publications are available at the web-sites: www.ifapcom.ru and www.mcbs.ru.

Claudia WANDERLEY
Researcher, UNICAMP University
(São Paulo, Brazil)

Multilingualism and Digital Libraries in Local Languages

Introduction

The Multilingualism in Digital World project has been held from 2005 within a network of 11 higher education institutions in 8 Portuguese speaking countries. After six years of a very intense experience on promoting multilingualism in an academic ambience of monolingual (lusophone) tendency, we present some of our main obstacles and some of the possible horizons that we could glimpse. Our practical solution by now is to work with free operational systems, free softwares, digital libraries and translations. We hope that our experience might be useful to people interested in building a network like ours, to policy makers, to start a broader debate on the construction of inclusive societies, or in short: to start conversations.

1. Big Picture of the Project

Working in a team with researchers from Angola, Mozambique, Cape Verde, Bissau Guinea, Saint Tome e Principe, East Timor, Macau (China), Portugal and Brazil it is quite easy to think of post colonialism and on what does it mean to develop common content in Portuguese that enable us to treasure, to respect and to maintain the bonds with our local languages and cultures. These countries and the region of Macau have Portuguese as their official language due to previous Portuguese colonization.

I'll focus on the situation of Brazil: Brazil or officially the Federative Republic of Brazil has a territory of 8,514,877 km², and a population of 190,755,799 (census 2010, IBGE), nowadays it has 39 linguistic families – the bigger diversity of the continent - and around 200 living local languages/cultures. And if we say that in 1500, at the time of the “discovery” of Brazil by Pedro Álvares Cabral, our linguists estimate around 1,300 languages/cultures living in the territory of what would become Brazil some hundreds years later, therefore more than five sixths (5/6) of these languages are gone.

If we gather the local languages present today in all these territories which we work with, we will reach around 700 living local languages, and surely a big agenda to think how to manage to include these languages in digital world,

and how to promote social/digital inclusion of these communities. Still, if we consider that Prof. David Crystal (2006) estimates the loss of two thirds (2/3) of the living languages of the world in the next three generations, we have to hurry. We have a big picture of the languages that we are connected with, and the countries concerned, and I'll draw some lines on the position that our project holds. It is very important to keep in mind that it is a research project by individuals feeling concerned with Multilingual issues, who are academics as well.

I understand that a collective responsibility based ethics could be a good way to approach Multilingualism in Digital World issues, but to do that researchers from Latin America and Portuguese Speaking countries have to participate in international debates and know each other's work. We need to become a research network on Multilingualism.

2. At the Academy

“Here the heritage of colonialism and the operation of neocolonialism can only be confronted by systems of collective responsibility-based ethics, [...]”

(Spivak, 85, 1999)

As a linguist and as a researcher it's been a new experience to bring multilingualism to the campus at UNICAMP University, at the city of Campinas in São Paulo, Brazil. As we all know, the promotion of multilingualism and inclusive societies do not fit in a specific disciplinary field of knowledge in a traditional academy for many reasons, and I'll cite two of them that have surprised me: first of all, for its extreme interdisciplinary nature, and second because it has a link with local cultures and local communities that are not always welcome in educational institutions that traditionally were thought for local elite.

There is lot of work ahead, many layers of change that would be necessary to have ideal conditions to promote multilingualism in digital world in our region and with our partner countries. Most of the barriers that we have found are historical, political and not at all in synchronicity with our project, or are being changed in a way that enhance such a proposal of digital and social inclusion like ours. So, we are working against the grain.

2.1. Languages and Academic Tradition

Today, among 192 countries that are part of the UN, approximately 20 national states have more than one language as its official language. This does not mean that other societies or countries (like Brazil that have only Portuguese as its official language) are not bilingual or multilingual, it only shows the lack of

political, juridical and educational support and recognition of these spoken languages and living cultures in national territories. It shows an old habit that was part of the construction of the state-nations that permit particularly in our case: the old Portuguese Empire and the Portuguese identity, a cultural and linguistic paradigm that has had its functionality and that now we need to better comprehend to move on. Fortunately it was good surprise to learn about the initiatives of the Russian Federation that together with countries like India, Canadá or South Africa might help us to approach efficiently our Multilingual issues.

The understanding of what is valuable for a nation changes through time and specifically in 2005 Brazil signs the Universal Declaration on Cultural Diversity. This declaration suggests, among other things, that cultures and languages shall be considered treasures of the nations, and underlines the need to support initiatives of recognition, respect, inclusion and circulation of these groups and the access to the knowledge generated by these groups in national territory as well as in the World Wide Web. In our region it is a brand new concept.

The change of attitude towards minor languages concerns a perception of our socio cultural and historical bonds. There are many blind spots that we have to deal with to make this move. To work and to reflect on multilingualism social and digital inclusion when the paradigm since colonization has been to affirm monolingualism is a difficult job, and our local education goes towards major or dominant languages while minor languages were forbidden until approximately ten years ago. Some theoretical apparatus is necessary to deal with such contrast in academy – the interest to be part of hegemonic culture versus the need to recognize the value of local culture. Otherwise we might develop an academic-schizophrenic profile as Kosambi and Spivak point out: “[...] “one cannot truly know the cultures of other places, other times,” and then proceed to diagnose the hegemonic readings into place.” (Kosambi/Spivak, p.50, 1999)

The first theoretical approach that I found to be useful and productive to comprehend Multilingualism in Brazil and in our partner countries is postcolonial reason criticism. One of the reasons might be that the idea of the foreclosure of the native informant is very present and quite visible in my field. Another possibility is for the need to comprehend the role of women in third world countries that is also something that catches my attention in postcolonial criticism works - being a woman in a third world country university – and by “third world” I’m referring to traditionally colonized academic culture and not specifically to Brazilian economy. Anyway, it is based on this theoretical approach and methodological tool that helps to deconstruct some hegemonic readings, that I’m able to present some perspective for the incoming project.

From this perspective I could understand that the linguistic bond [the Portuguese as official language] that at first gathered the network of this project in a very naïve perspective, has been historically overestimated, and, as PEREIRA (2009, 155) says, “became a monument of the complicity between colonized and colonizer, complicity which does not guarantee the end of the reproduction of the colonizer–colonized [...]” violence. One possible question here is what language bonds cover, silence, pasteurize, and what can be done in terms of promoting inclusive societies in such ambiance.

2.2. Post Colonial and Peripheral Perspectives

At school we learned about indigenous groups as part of the discovery of Brazil by the Portuguese in 1500, what immediately made them become creatures of the past. And what a surprise it is to find out that these communities survived and that they are alive in 2012, some of them with internet, wikis, blogs, some (the great majority in fact) needing help to get on digital world and to access information.

Indigenous societies in Brazil are not studied in Sociology or History. Of course there are indigenous languages studies, but unfortunately they live apart from linguistics (western linguistics), language and literature groups. It is a very specific field, called in seminars linguistics of non-western languages, ethnic-linguistics, anthropological linguistics, etc. So it is not part of the general culture about our country or identity, for as Brazilians we are officially part of the western civilization, we are monolingual, and they are not.

Talking to some indigenists in Brazil I figured out that they made a choice during the dictatorship to stay in national ground to work with these communities, learn their languages etc., and although it has propitiated a strong bond of the researcher with the local communities, it has left them no opportunities to participate in international debates. This is not a specific historical issue in Brazil, it happened in general in Latin America. If we look for organizations, academies or networks concerned with Multilingualism in this region, we will find a blank to be filled. More than that, if we look for funding, support, infrastructure, museums, collections, libraries, observatories on Multilingualism in our region... basically it is for our generation to build them.

Our scope to comprehend Multilingualism in this project is quite open: in Latin America, Africa and Asia we have very different regional and historical aspects, but we have in common the link to Portugal, that somehow establishes patterns for silencing local cultures. To guide us in Multilingual issues we have clear policies of silence, policies for censorship, considering

the historical and economical background of each group, but to reach local cultures, local languages, to develop and preserve it is yet something to come. Such a comparative study would be very interesting, not to victimize but to enable us to have a memory, to understand what happened and to resist these effects in better “knowledge conditions”. With Europe we have a very clear counterpoint for they made an option for multilingualism which is different from the former scenarios. Our dialog with European countries is very important regarding their experience on multilingualism development, their academic history or even considering the funding possibilities. A very good partner to work with Multilingualism in our region is UNESCO team, because they give us feedback – which is rare in our local reality concerning this theme – and are capable of contextualizing the obstacles that we cannot surpass locally.

3. Actors of the Project

In dialog with Frances Albernaz (UNESCO) who coordinates the Network Humaniredes⁸⁰, and reading Luis Felipe de Alencastro – *O Trato dos Videntes* – I have chosen to start working on the Multilingualism in Digital World project with Portuguese speaking countries: Brazil, Bissau Guinea, Mozambique, Cape Verde, Saint Tome e Principe, Angola, and East Timor. Brazil’s role is a bit differentiated in this group, for it’s been officially “independent” for 190 years, while the other countries have an average of 36 years of “independent” life. Nevertheless, our common bond in international ground is the presence of the Portuguese language as the official language. It enhances an exchange of information considering Portuguese a vehicular language.

Also, there is a strong need to reflect on the role of academic institutions, inspired by postcolonial criticism. In this aspect we are specially interested in Humanities local authorship development, recognizance and knowledge circulation. In short: south-south cooperation.

3.1. Memory of the Actions

This is our starting picture, and to develop such a proposal together with Prof. Frances Albernaz we have talked to the representatives of the eight countries about the network who have put us in touch with their national higher education institutions. We have had from the Brazilian Ministry of Culture the possibility to create and support points of culture (*pontos de cultura*). Basically we would have to: 1) promote a field research in at least

⁸⁰ *Network, established in 2005, involves mainly higher education institutions in Portugal and countries that were former colonies of Portugal.*

one local language and to register it in media to share with the network and 2) elaborate a proposal to include this local community online. The project of points of culture would provide the cameras, server and computers to make the edition of the content and to publish it online. We sent a call from the Humaniredes network (2006) and fourteen universities sent us proposals. In 2007 we had the approval for a *UNESCO Chair Multilingualism and Local Content Production in Digital World* at the University of Campinas in São Paulo, Brazil. However the projects sent to the Ministry of Culture were denied, for the Brazilian Ministry of International Relations (Itamaraty) had forbidden to send Brazilian funding abroad. So, we had the projects ready, a UNESCO Chair, and no funding. We made some attempts to engage Brazilian research agencies, but multilingualism was out of the scope of any national call. We also tried to ask for funding from the European Union who understood that the project was good, but saw no objectives in the simple construction of a south-south network, no possibility of and no reason for opening dialog and content exchange among us.

Another interesting initiative in partnership with the European Research Council in 2007 was to foresee the needs of research in multilingualism and to produce a call for research customized for the ongoing projects. We have worked on that possibility and as far as I know, in the ERC internal meeting to deliberate about funding the representative from Portugal at ERC denied the proposal and so Portuguese speaking countries were kept out of the range of the call.

At UNICAMP University in 2008 my former laboratory understood that multilingualism was out of the scope of the lab, and I was transferred to the Center of Memory, and in 2011 to the Center of Logic, Epistemology and the History of Science which is where I work now in this very project.

3.2. *Border Line*

Facing such difficulties was a surprise for us. In fact, it seemed natural to propose such a network, and as every agency not only agreed but also manifested interest to support it, none of these “failures/rejections” were expected. The Brazilian agencies and ministries discourses were centered in the idea of an interesting proposal that should have all institutional support, but by different reasons wouldn’t have fund provisions, and at the university we had the recognition of an important project that wouldn’t have institutional support by now.

Multilingualism in Digital World – although thought as a simple proposal – touched pseudo-ethical issues, that were not in perspective at that time. The context that I’d call postcolonial did not permit the infrastructure to enhance the emergence of

the discourse of the subalterns, of local languages, of local content production and of dialog. Nevertheless it seemed interesting to explore real possibilities for this enterprise as we have reached the bottom line of the politically correct discourse and found nothing concrete. It is a dead end for the project as conceived, but an interesting starting point for reflexion and for finding the right partners.

Therefore with patience and without funding, I have consulted UNESCO about possibilities to work with this network under such conditions, and the suggestion was to work with the Greenstone digital library project <www.greenstone.org>. And we developed the following idea.

4. Digital Polyphonic Library: Open Proposal

The notion of democratic access to information in this project will be developed through the creation and diffusion of local content in local languages. The idea is that speaking subjects experience their speaking capabilities through a variety of repertoires of linguistic character that exist simultaneously and in a gradual manner, without necessarily the presence of linguistic frontiers well delimited and well defined. The objective of a polyphonic digital library is to create a polyphonic knowledge data base that permits the comprehension and an integrated digital experience of this myriad of repertoires and to facilitate the navigation among them. This digital polyphonic library therefore is not centered in a mere translation process or in the transportation of knowledge from one language to another, but in an imaginary web of knowledge in which many linguistics repertoires are imbricated. This web finds its possible actualization in a digital polyphonic platform that permits the presence and the simultaneity of various phonies to be serialized, sequentialized, and shown in multilateral relations that do not affect their individuality.

Before colonial era and the expansion of Europe in the world, the majority of populations lived in phonic regimes – speaking praxis, and sometimes, writing praxis – that were not based on the idea of languages as well determined totalities belonging to a specific territory (with its specific speaking population). Our point here is that the majority of people had an open linguistic repertoire, with internal and gradual differences, using them in different contexts and different goals, as well as linguistic formulas and vocabularies, that today are perceived as belonging to many languages.

The idea of this polyphonic knowledge base has the following criteria:

- To promote in its various aspects the permeability among many repertoires;
- Consider computational system and its users intrinsically polyphonous;

-
- Permit to many repertoires to interact with the minimum of barriers, and allow them to constitute themselves mutually as communicational spaces and digital knowledge;
 - Allow the knowledge of these many repertoires to be shared at its most possible extent;
 - Allow available collections historically associated to many repertoires to be freely shared;
 - Allow the digital inclusion of repertoires that do not possess substantial written collections and the creation of their relations others, in multimedia base;
 - Allow permeability among repertoires to generate rizomatic knowledge creation (i.e. in a web design and without neuralgic points of control), without disciplinary approach (i.e. through knowledge exchange and/or through linguistic territorialities thought as historically rooted in different spaces and different time flux, that immediately will demand translation among them to communicate.); and
 - Use digital space as an instrument of de-territorialization.

5. Best Conditions

We have also asked the network members to share opinions on the opportunities of this project, and the mains obstacles.

The list of current main local opportunities concerns roles for the university that were never available in our postcolonial set. And, unless we can trick history, there is no easy possibility to promote such insurrection of local content within the place that have systematic and consistently silenced it. Here is the positive list.

- Possibility to develop local academic and artistic authorship, concerning subjects of local reality;
- Possibility to develop digital inclusion with multicultural perspective;
- Possibility to develop research on multilingualism with local researchers (south-south cooperation);
- Participation of universities empowering communities in digital editing and publishing, with content quality priority;

-
- Participation of local universities in historical and cultural patrimony preservation;
 - Participation of local universities in multicultural preservation;
 - Possibility to develop humanities networks (Humaniredes) in Portuguese speaking countries;
 - Human Sciences reaching needs of local society and promoting wider debates on multiculturalism and multilingualism;
 - Humaniredes and Network of higher education institutes contributing to the UNESCO Information for All Program.

Current main obstacles include:

- Lack of stabilized electric energy in Higher Education Institutions;
- Lack of telephony, or expensive telephony in Higher Education Institutions;
- Low infodensity and little access to the Internet in Higher Education Institutions;
- Lack of human science institutes to promote research concerning local reality;
- Local content production crisis, know-how to conceive media content not being taught in digital inclusion process;
- Local content circulation crisis, strong circulation of ready-made global content;
- European academic models to research funding not fitting real profile of possible academics in poor countries that could be pioneers in a local research field;
- Insufficient English proficiency for young researchers to write a project to end for a funding agency abroad;
- Dependency on European or American academics to give a “brand” on the quality of the research, rather than promoting networks of similar proposals that would enrich their comprehension of the matter together;
- Low libraries presence, difficulties with publishing and distributing books;
- Low social recognition of the academic work;

-
- Funding institutions with difficulties to innovate subject and approaches, unless they're considered important abroad;
 - Brain drain of academic staff usually to Europe and to United States.

This list shows that we were not ready to start working at the level proposed in terms of basic infrastructure and sometimes human resources. It requires of us, as a network, a change of culture to work to propose academic South-South cooperation.

Departing from this picture we made suggestions for new requirements (institutional and funding formats) for international and cutting edge research on Multilingualism:

- Invest locally in Junior Researchers;
- Consider research projects on national languages or official languages. This would be interesting to strengthen partnerships with local institutions, so they would be forced to get in touch with new research trends;
- Invest in research network for humanities, sharing libraries and promotion of professors and researchers exchange;
- Invest in similar pairs evaluation (someone in Angola is much more able to evaluate the real condition for research production in Mozambique, than a specialist that doesn't share the same research conditions).

6. For the Construction of Inclusive Societies

In our region there is a contradiction on the necessary presupposition that there is a dominant knowledge, expressed in academy, and that we – as scholars – are in a position to re-inscribe local languages and local cultures without any support. In contact with the group that works with multilingualism in the 2011 Yakutsk conference and with the UNESCO Information For All Programme group dedicated to multilingualism I figured out that from the academic perspective we need to organize the memory of the initiatives already taken, and to create an observatory of the projects in course.

It is important to give a place in academy for studies on multilingualism in general and for Multilingualism in Digital World. We should be more flexible with multiple areas of knowledge to consider Multilingualism studies. It is possible to work with multilingualism from many angles. But to enable the construction of inclusive societies within local universities in our region, it is necessary to have a minimum of good examples, reports of national experiences, similar to the Russian publication “Preservation of Linguistic Diversity:

Russian Experience” (2008), and – why not? – at least one academic review on the subject. Of course we are not good examples of preserving local culture, that is why it is so important to have access to information about Multilingualism, and to be in touch with major actors. We have a lot to learn.

References

1. Crystal, D. *Language and the Internet* (Cambridge: CUP), 2nd edn, 2006.
2. Eco, U. *A busca da língua perfeita na cultura européia*, 2nd ed. Bauru, SP, EDUSC, 2002 [1993].
3. Ethnologue, *Languages of the World*, available at <http://www.ethnologue.com/>.
4. Instituto Brasileiro de Geografia e Estatística (IBGE), available at <http://www.ibge.gov.br/>.
5. Pereira, P. S. “Fermento da República, Bolor do Império: civilização Ibérica, excepcionalismo e o legado luso-brasileiro do lusotropicalismo” in *Estúdios Portugueses* 9, 2009.
6. Ribeiro, D. *O povo brasileiro: a formação e o sentido do Brasil*. São Paulo: Companhia das Letras, 1995
7. Spivak, G.C. *A Critique of Postcolonial Reason: Toward a History of the Vanishing Present*. Cambridge, Mass. & London: Harvard University Press; Calcutta: Seagull Press, 1999.

Zhanna SHAIMUKHANBETOVA

*Director General,
National Academic Library of the Republic of Kazakhstan
(Astana, Republic of Kazakhstan)*

The Role of Kazakhstan's Libraries in Shaping the Country's Linguistic Landscape

Language is one of the most important cultural phenomena as well as a multifunctional tool. It moulds us as personalities, passing on information about our heritage.

In his prose poem “Kara Soz” (“Words of Advice”), the Kazakh poet Abai Kunanbayev says that the command of a language opens up a window onto the achievements of world civilization: “Knowledge of a foreign language and culture puts one on a par and at ease with the native speech community, and if receptive to their concerns and aspirations they struggle for, that person is bound to join in.”

By exploring other cultures, people can also develop self-critical thinking and expand their outlook.

Kazakhstan: Linguistic Situation

Nowadays, as so many of the world's languages disappear (according to UNESCO statistics, one language becomes extinct every other week), the Kazakh government remains committed to its policy of supporting all the languages spoken in the country. “Language policy” may sound rather formal as a term, but behind it in Kazakhstan are the fates of the many constituent ethnic communities. So maintaining peace and harmony in the country's multilingual, multi-faith society should always be a key priority.

Despite the many historical upheavals they have been through, the Kazakh people have not lost their greatest asset, the language, which is key to comprehending the nation's essence.

“It makes me happy to see Kazakhstan's people become more respectful and proud of their national language and more willing to study it,” President Nursultan Nazarbayev said in his address to the nation on January 28, 2011.

And speaking at the 13th session of the Assembly of the Nation on August 20, 2007, Mr Nazarbayev said: “...we should make every effort to further promote the Kazakh language, which is the main unifying force for all of Kazakhstan's

citizens. That said, we also need to create favourable conditions for members of the country's other ethnic groups to speak and receive schooling in their native languages."

This principle became the foundation for the National Programme of Language Use and Promotion, intended for the 2011-2020 period. The programme was elaborated in keeping with Articles 7 and 93 of the Constitution of Kazakhstan, the July 11, 1997 Law on Languages, the Doctrine of National Unity, and the Language Policy Framework.

Work on the Programme will proceed in three areas:

- broadening and enhancing the socio-communicative role of the national language;
- maintaining the role of the Russian language in the cultural domain;
- promoting other languages of Kazakhstan.

There are three stages to cover:

- Stage 1: Taking steps to improve the regulatory and methodological base for language use and promotion (2011-13);
- Stage 2: Implementing practical measures to introduce new technology and methodology in the area of language study and use, as well as in linguistic diversity preservation (2014-2016);
- Stage 3: Consolidating the results through systemic monitoring of national language use in every area of social life, as well as of its standard usage and proficiency level, with the positions of Kazakhstan's other languages maintained (2017–2020).

At President Nazarbayev's initiative, a national cultural project, "Language Triad", has also been launched. Every person living in modern-day Kazakhstan is well aware that the command of at least three languages – Kazakh, Russian and English – is crucial to his or her success. In our increasingly globalized world, speaking several languages is an indispensable asset because we all need to know how to orient ourselves in a multilingual environment.

Objectives:

- qualitative (an improved linguistic environment, more self-motivation to learn the Kazakh language among the country's non-Kazakh ethnic groups; wider research activities for Kazakh and other languages spoken in Kazakhstan);

-
- quantitative (the proportion of Kazakh citizens speaking the national language may grow to 95%, up from today's 60%). These are official projections, to which we could add another 10% – representatives of ethnic minorities, such as Uigur, Tatar and Azeri – who speak their respective mother tongues and have a basic-level command of the Kazakh language. By the year 2020, the share of Kazakh citizens speaking Russian is expected to reach at least 95%, against today's 89%. English-language speakers should also grow in numbers, to eventually account for some 20% of the population.

Kazakh Language and the Internet: Analyzing the Current Situation

The Internet has become a truly cosmopolitan zone by now. And although the Web's linguistic landscape is still dominated by English, the world's nations seek to create online resources in many locally spoken languages.

What is the situation like in Kazakhstan? September 2011 marked 17 years since the nation went online. These days, about a hundred new websites appear here monthly. Not all of them have a Kazakh-language version, though, with Kazakhstan's developers often giving priority to Russian and English these days. However, in the past five-seven years, there has been a positive trend toward the expansion of Kazakh-language Web content, prompted by the launch of a large-scale national programme to develop an online government.

The recently created portal National Language of the Republic of Kazakhstan (<http://til.gov.kz>) is a major reference source on the government's language policy. Kazakhstan's government agencies, news organizations, mass media, large national companies, and culture, education and health institutions now all have a Kazakh-language version of their websites.

As for business-oriented sites, English and Russian still prevail there. This can be attributed to Kazakh-language sites not being attractive enough commercially, as well as to the popularity of Russian in Kazakhstan as a language of business communication, both off and online.

Some Kazakh-language websites latinize their content in an effort to expand the audience, because, as is known, of the 5-odd million Kazakhs based outside Kazakhstan, about 50% use the Latin alphabet. But for such resources to develop, appropriate financing should be provided.

Another factor holding back the expansion of Kazakh presence online is the lack of technical support for the Kazakh script (the standard Cyrillic alphabet plus nine additional characters). It has taken quite a long time to develop a

single standard for encoding these characters and to spread fonts that have them around.

The effective language promotion programme gives reason to hope that a system of government support for Kazakh-language websites will be created before long. Already, active measures are being taken to promote Web content in the national language. And there is a whole number of success stories to inspire further effort.

The national Internet award AWARD.kz (<http://www.award.kz>), established in 1999 by an eponymous public association, spotlights the best of Kazakh-language websites, including in specific categories, such as Best Website in National Language and Best Bilingual Website.

Earlier this year, the public foundation WikiBilim launched a project to raise the Kazakh profile in Wikipedia (<http://kk.wikipedia.org>). As of July 2, 2011, there are 44,702 Kazakh-language entries, submitted by 9,090 contributors, including 5 with administrator status. The number of corrections totals 321,737. There are plans to hold a nationwide campaign under the motto “200,000 articles for the country’s 20th independence anniversary.” Further expansion of Wikipedia’s Kazakh-language content will foster its everyday use, thus consolidating its role in Kazakhstan and beyond.

The project Kazakh National E-Library (www.kazneb.kz), run by the National Research Library in partnership with its smaller counterparts, also seeks to promote the Kazakh language online. These days, Internet users can have access to 2,000 complete digitized copies of books in Kazakh, featuring Kazakhstan’s history, culture, science and literature.

With ICT being one of the most vibrant sectors in Kazakhstan’s modern-day economy, the number of Internet users in the country is on the rise. In 2011, for instance, it exceeded 5 million, or 34% of the population. And Kazakh-language Web content is likely to grow further, provided that the government offers more incentives and public organizations provide their support.

Kazakhstan’s Libraries and Their Role in Creating a Language Environment

Being information and communication centres for their respective communities, Kazakhstan’s libraries see it as their professional duty to promote the creation of a tolerant language environment as a factor of national unity.

In their daily work, Kazakh library staffs are guided by UNESCO’s Universal Declaration on Cultural Diversity, signed in Paris in November 2001, and, specifically, by its points about “safeguarding the linguistic heritage of

humanity and giving support to expression, creation and dissemination in the greatest possible number of languages” and “promoting linguistic diversity in cyberspace and encouraging universal access through the global network to all information in the public domain.”

The government programme for the promotion of languages in 2011-2020 has prompted libraries to step up their efforts to popularise Kazakh as the national language, Russian as a language of inter-ethnic communication, English and other foreign languages as means of international business communication, and languages of other ethnic communities living in Kazakhstan as a prerequisite for the preservation of multilingualism and mutual intercultural enrichment.

Kazakhstan’s libraries, who boast a wealth of print and digital sources as well as wide opportunities to employ ICT, have accumulated extensive experience in meeting the information needs of the country’s multi-ethnic population, with due respect for the principles of linguistic diversity. The success of that work is ensured by an original strategy, consisting of the following 7 points:

1. Building a multilingual information resource;
2. Contributing to the implementation of targeted government programmes and projects;
3. Promoting social partnerships with various community organisations on the ground;
4. Forming a level playing field for the various communities to express themselves culturally;
5. Doing research into issues of library and information support for the country’s multi-ethnic population;
6. Using ICT in the creation of a qualitatively new model of library and reference services;
7. Providing methodology consulting on topical issues of library and reference services for the country’s multi-ethnic population.

The Kazakh libraries’ main partner, without any doubt, is the Assembly of the Nation (<http://www.assembly.kz>). Enthusiasm and love for one’s language and culture are key to the assembly’s success. This is a one-off public institution, committed to harmonising interethnic relations in Kazakhstan. No other country boasts such an organization. This is the first Kazakh president’s “know-how,” which has been unfailingly proving its efficiency ever since the country gained its independence two decades ago.

The assembly holds a depositary at Kazakhstan's National Research Library. The collection provides information support for the country's ethnic policy activities while also consolidating Kazakh society's spiritual culture, crucial to maintaining national unity.

It features books providing a profile of the assembly and telling of activities by its units – various ethno-cultural associations, an expert council, and a public foundation. The collection is also accessible online, at the National Kazakh E-Library (www.kazneb.kz).

Thanks to its partnership with the assembly, this e-library has become a platform for various programmes in fostering intercultural dialogue and promoting the national language, as well as the languages, cultures and traditions of all of Kazakhstan's ethnic groups.

The country's libraries realize perfectly well that information and communication technology is one of the most efficient tools for building a favourable linguistic environment. This is why, along with conventional forms of library and reference services, readers are also offered resources and services based on ICT, such as e-libraries and electronic reference units, websites, online displays, and e-posters. Some positive experience has already been accumulated in this area, but we still have a long way to go in promoting Web content in Kazakhstan's national language and languages spoken by the non-Kazakh population – a major priority with the country's public libraries.

Irina DOBRYNINA

*Director, National Library of the Republic of Karelia
(Petrozavodsk, Russian Federation)*

Creating Content in Minority Languages: Enhancing Users' Capacity

Linguistic and cultural diversity as a part of the world's heritage is commonly known to be as crucial for mankind as biodiversity is for nature. UNESCO documents also highlight the importance of diversity, emphasizing the respect for linguistic and cultural diversity as one of the core principles of the modern society development.

Over the last decade issues of support for national languages and cultures, maintenance of local traditions and preservation of the region's individuality have been of great topicality and importance for the Republic of Karelia – a unique cross-border region with indigenous national culture. It is mainly explained by the increasing globalization processes, rapid development of technologies and communication, intensification of migration flows, and the interconnection and interdependency of cultures.

The Government of the republic recognizes culture as an essential resource for developing the territory and encourages the region's cultural identity. The importance of facilitating the preservation, maintenance and interpretation of the tangible and intangible historical and cultural heritage and its integration in the social and economic development is enshrined in the Concept of Socio-Economical Development of the Republic of Karelia for the period until 2016.

In the course of the state national policy implementation in the republic special consideration is given to the sustainable ethno-cultural development of indigenous peoples. It is the key aim of the programme under development on "Preserving the Unity of Karelian Peoples and Ethnic Communities for the Period 2012–2016" ("Karjala is Our Home"). The programme follows up two previous programmes – "State Support for Karelian, Vepsian and Finnish Languages in the Republic of Karelia for the period 2006–2010" and "Harmonizing National and Religious Relations and Developing Civil Consent in the Republic of Karelia for the period 2007–2011" ("Karelia, the Territory of Consent").

Today we are witnessing Karelia's development as a multinational republic, inhabited by more than 100 nationalities, aspiring to preserve their mother tongues and their unique cultures. The indigenous peoples of the republic are Karelians, Vepsians and Russians. Karelian nation is divided into three groups:

north Karelian, Livviks and Ludiks. Karelian language belongs to the Baltic-Finnic branch of the Finno-Ugrian languages, and the term in its broad meaning includes all three dialects. Vepsians belong to the indigenous small-numbered peoples of the North, Siberia and the Far East of the Russian Federation⁸¹. According to the 2002 All-Russian census, there are less than nine thousand Vepsians in Russia, 4,870 of them living in the Republic of Karelia (0.7% of the republic's population). Karelia is a specific territory of the Russian North preserving unique culture of the Pomors – an ethnographic group of Russians with its own subdialect, customs and traditions formed in the course of the territory expansion and development.

These ethnoses have left rich historical and cultural heritage for their offsprings that serves as an inexhaustible source of our region's development. The heritage includes world famous cultural artifacts and landmarks, namely the Church of the Transfiguration on the Island of Kizhi; the Sheltozero Veps Ethnographic Museum; the Vodlozero National Park – the largest national park in Europe and the Paanajarvi National Park known as the “Pearl of the North”; religious buildings and structures of Zaonezhje and the Karelian petroglyphs, unique memorials of monument art of the late Stone Age (IV-II millennia B.C.). Karelian historical and cultural heritage is concentrated mainly in little towns and rural areas. There are over 150 towns, villages and settlements in Karelia with great history and significant cultural potential. They differ in building patterns, planning, natural surroundings, traditional arts and crafts, folk songs and dances.

Preserving and promoting linguistic and cultural diversity and supporting the region's identity are therefore major lines of action for the research, education and cultural institutions of the republic, as well as its NGOs.

Karelia is facing demographic problems today. According to the all-Russian census in 2002 the republic's population totaled 716,300 (0.5 % of the Russian Federation population), with Russians accounting for 76.6 %, Karelians for 9.2 %, Vepsians for 0.7 %. Preliminary figures of the 2010 census indicate a significant population decline down to 647,000.

According to a 2010 survey results 60% of Karelians and Vepsians living in the republic read newspapers and magazines in their mother tongues, 75 % – watch TV programmes and listen to radio broadcasts in mother tongues. 80 % of Karelians and Vepsians deem their right to use mother tongue is regarded. There are, however, issues for concern. One in five Karelians and one in three Vepsians polled expressed their concern over the sustainability of their mother tongue. About 40% of Karelians and 30% of Vepsians mentioned their

⁸¹ *Unified Register of Indigenous Small-Numbered Peoples of the Russian Federation (Decree of the Russian Government Nr 536-r of 17 April 2006)*

language being endangered. Only 41% of the respondents were positive of the modern school's increased opportunities for native languages preservation and promotion, while 40% considered it necessary to continue further work on preserving and supporting native languages.

The survey did not cover the issues of ICT and Internet use by indigenous peoples, however even a shallow analysis of the social networks like "Facebook" or "VKontakte" shows that Karelian and Vepsian languages find representation there. Statistics demonstrates a rise in the frequency of use of full-text resources created in the republic in these languages. More users visit republican web sites presenting resources in Karelian, Vepsian, Finnish, as well as information in Russian on the indigenous peoples of Karelia, their languages, traditions and territory.

Interactive multimedia projects (including Internet-based) are being actively implemented in the republic, aimed at the creation of content in national languages. Use of modern technologies facilitates the creation of information resources and products in various formats. Such "live" projects and complex forms of work cause constant broad public interest – both among adults and children, thus allowing raising the prestige of minority languages and strengthening their social and functional role.

For several years, the National Library of the Republic of Karelia has worked in partnership with the NEVOND Information and Legal Centre for Indigenous Peoples in the framework of the project on training in computer technologies and methods of information retrieval on the Internet provided for representatives of indigenous peoples living in Karelia.

Today the library is implementing various activities in close cooperation with the State Committee for National Policy and Liaison with Public Associations and Religious Organizations of the Republic of Karelia; the Institute of Language, Literature and History of the Karelian Research Centre of the Russian Academy of Sciences; with republican cultural institutions: museums, the National Archives, the National Centre for State Protection of Cultural Heritage, creative groups of the republic; higher education institutions of Petrozavodsk; Republican youth public organization «Nuori Karjala»; "Periodica" publishing house producing books and periodicals in the Karelian, Vepsian, Russian and Finnish languages; and with writers unions of Karelia.

The results of this fruitful partnership and inter-regional and international cooperation are reflected in the content of web-projects implemented by the National Library of the Republic of Karelia.

The Finno-Ugric Libraries of Russia project (<http://fulr.karelia.ru/>) was launched in cyberspace in 2004. This bilingual (Russian and Finnish) enterprise portal is a single point of access to the sites of participants, the story of their resources and services, e-books collection in the Finno-Ugric languages. Eight major libraries of the Finno-Ugric territories are engaged in the project. Participants are provided with opportunities for remote posting of local news.

Over the past few years the design and structure of the portal have undergone some changes, new partners have joined the project, namely the Estonian organization “Fenno-Ugria” and the Finno-Ugric Cultural Centre of the Russian Federation in Syktyvkar. Materials from Finnish and Norwegian libraries, the Estonian Academy of Arts, the National Museum of Mari El and Karelian museums are being published. Electronic collection includes copies of publications in Finnish, Karelian, Vepsian, Mansi, Komi, Cheremiss (Mari), Mordvin and Hungarian. New sections have been added to the portal, namely “Catalogues of publications in Finno-Ugric languages”, “Creators of national literature and folklore”, “Kaleidoscope”, “Finno-Ugric world”. Seven photo galleries represent the world of the Finno-Ugric peoples (http://fulr.karelia.ru/cgi-bin/flib/fin/view_kaleidoskop.cgi?id=119&material=142). Today we can boast of visitors from about 90 countries of the world.

The Indigenous Peoples of Karelia project (<http://knk.karelia.ru/>), has been implemented since late 2009. It is a sort of a blog site aimed at providing information on our homeland, amazing and magical Karelia, on its past, present and future, on the people living here. We use Movable Type software as an open source platform, permitting the application of blogosphere technologies. We work on the creation of interesting content in cooperation with researchers, journalists and publishers, with Karelian republican and regional museums and the National Archive, with municipal libraries and cultural centres of different districts of the republic, with national public associations and NGOs. Users leave their comments, address our authors and each other, moderators are eager to answer any questions and often initiate personal messages sharing, and as a result the project somehow performs functions of a social media. Questions and commentaries help us adjust and update our plans for project development. We are expanding the range of web tools for publishing materials on the website. Users can watch videos, listen to recordings and work with flash diagrams. Various photo galleries are used for viewing images.

Audio-visual materials in indigenous languages of Karelia are published in the “Media Library” section. Our key goal is the preservation and promotion of audiovisual cultural heritage and creation of our own audio and video resources. We also aim at providing support for national languages and ensuring

their promotion online. The “E-book shelf” section contains bibliographic and reference lists, articles and publications on the history of our region.

We always indicate the authorship and information sources. It might be for that reason that the Russian library community considers our project as a new format of bibliographic resources on the Web. Our website has already attracted visitors from over 80 countries.

Our new project – the Digital Library of Karelian Authors (<http://avtor.karelia.ru/>) – is freely available since June 2011. We wanted to create a virtual meeting place for contemporary Karelian authors and their readers, to assist both authors and growing army of e-readers. We started with publishing 28 electronic editions of 14 authors invited to cooperate. Publications are grouped into nine categories, including prose, poetry, translations, literary criticism, etc. After placing links in the Karelian web-space and presenting the project in the local media we began to receive proposals from new authors through the feedback.

We have made agreements with owners (authors and publishers). We are proud of having digital publications labeled “first published”. It means our project has earned the trust of authors. We hope that it will catch the fancy of Internet users. Of course, we plan to prepare publications not only in Russian but also in Finnish, Karelian and Vepsian languages with Russian translation provided.

Analyzing the demand for similar Internet projects, their standing in the local scientific and creative community, we come to the conclusion that success is only possible by combining various creative initiatives under the roof of the National Library.

Anastasia PARSHAKOVA

*Deputy Director, Interregional Library Cooperation Centre;
Project Coordinator, Russian Committee,
UNESCO Information for All Programme
(Moscow, Russian Federation)*

Elaborating Practical Guidelines for the Stakeholders of Multilingualism Development

The Declaration of Principles of the World Summit on the Information Society stressed that the policy of information society building should be based on respect for cultural identities, cultural and linguistic diversity, traditions and religions, should inspire and promote dialogue among cultures and civilizations. The creation, dissemination and preservation of content in diverse languages and formats are accorded high priorities.

The Geneva Plan of Action calls for the exchange of knowledge, experiences and best practices on policies and tools designed to promote cultural and linguistic diversity at regional and sub-regional levels.

Potential stakeholders of the multilingualism promotion process – those who can and should contribute to the maintenance and strengthening, equipping and development of language – are numerous.

Of course, government policies and activities are of paramount importance. Efficient policy includes a set of interrelated measures aimed at strengthening and improving the activities by all other major stakeholders.

Let us name these stakeholders:

- education;
- research institutions;
- memory institutions;
- cultural institutions;
- book and media publishers;
- bookstores, newspaper and magazine stands;
- digital media;
- ICT industry;
- NGOs;
- private sector.

Institutions of primary, secondary and higher education undoubtedly play a crucial role in multilingualism promotion. They should cooperate with the legislative and executive bodies to support and develop minority languages.

Research institutions can perform scientific and applied research in the field thus providing the scientific and theoretical basis for governments and other social institutions to support languages.

Memory institutions aim at collecting, storing, and promoting all major evidences of the history of a given ethnic group, ensuring their availability to the public as well as developing various methods and forms of providing access to cultural heritage, including written.

Media can also contribute to supporting and improving the status of minority languages serving as a tool for the exchange of spiritual values and the promotion of cultural and linguistic diversity.

Book publishing and book trade can also do a lot to support multilingualism, since a language's lack of access to the book publishing sphere can pose a threat of its speakers to be largely excluded from the intellectual life of the community.

Various projects on the creation of multimedia content in minority languages can be initiated and implemented by all the above mentioned institutions in collaboration with other institutions of culture, science and education.

We are all aware of the fact that developing linguistic and cultural diversity in cyberspace grants us an opportunity to preserve languages and cultures, to expand domains of use of under resourced national languages. That is why it is extremely important to understand the problem clearly and define the set of actions to be taken to study, develop, use and promote languages.

It's not only political measures that is required. Development of legal norms and standards is fundamental, however multilingualism support cannot be limited to the adoption of regulations or solving technical problems. All stakeholders should be aware of the burning challenges as well as possible solutions. This requires developing specific guidelines for each institution. Of course, promoting best practice of leading experts in a particular field is important in this regard.

Participants of activities aimed at multilingualism preservation should be provided with information on the current situation in the field, on the state of the art in their specific professional area in the country and the world. It is important to analyze the possibilities of potential actors, as this will help to

outline the scope of collaborative efforts by representatives of different areas to support and promote linguistic diversity.

We have attempted to offer a document containing specific methodological and practical guidelines for one of the institutions supporting multilingualism.

The publication “Developing Multilingualism in Cyberspace: Guidelines for Libraries” has been prepared by the Interregional Library Cooperation Centre and the Russian UNESCO IFAP Committee. The book is intended primarily for library heads and specialists, but can be also useful for other organizations and institutions working in the fields of culture, science, education, information, communication, ethnic and cultural policy. We aimed at representing the global approach to solving this problem, the activities of international organizations, and the current linguistic situation in Russia.

Why are we focusing on libraries? UNESCO considers libraries to be key partners in expanding access to diverse cultural and linguistic resources, and actively cooperates with the International Federation of Library Associations to promote the information capacity building in the library sphere.

Libraries are regarded as educational, cultural and information centers that maintain, develop and represent various cultures, provide access to educational materials and programmes, acquire, create, systematize and provide access to information to meet the needs of all communities.

The IFLA Multicultural Library Manifesto underscores the importance and essential nature of libraries in learning, as they facilitate access to a variety of cultural and linguistic resources that open horizons to different experiences. In order to fulfill this mandate, libraries must meet the varied needs and interests of the communities they serve, especially marginalized or minority groups that may exist within any given community.

Library and information services in a culturally and linguistically diverse environment include both the provision of services to all types of library users and the provision of library services specifically targeted to underserved cultural and linguistic groups. Special attention should be paid to groups which are often marginalized in culturally diverse societies: minorities, asylum seekers and refugees, residents with a temporary residence permit, migrant workers and indigenous communities.

Our publication represents best practices of libraries of Russia, CIS and foreign countries implementing activities in support of linguistic and cultural diversity, in particular in cyberspace.

Of course, we could not cover everything in one book and had to give only some examples.

While describing foreign experience, we focused on Finland and New Zealand – countries that are among the world leaders from the point of living standards, the development of information society, and of library services.

In selecting examples of the Russian experience, we took into account the leading role of largest academic and public libraries of Russian multilingual regions in supporting multilingualism in cyberspace in our country. Major efforts are taken in Russia by national and regional libraries of Russian constituent entities, libraries of major universities, central public libraries in big cities.

In addition to general guidelines for libraries and description of potential action lines the book provides recommendations on the creation of network resources in local languages using ICT. We tried to analyze the difficulties that creators of library websites and blogs dedicated to the languages and cultures of minority peoples may face, and the requirements for such resources (openness, scalability, personalization, control, consistency, etc.). The digital version of “Developing Multilingualism in Cyberspace: Guidelines for Libraries” is available online at the website of the Russian UNESCO IFAP Committee www.ifapcom.ru.

Valentina SAMSONOVA,
Director, National Library of the Republic of Sakha (Yakutia)
(Yakutsk, Russian Federation)

Preservation of Cultural Heritage of Indigenous Small- Numbered Peoples of the Russian North, Siberia and the Far East

The outcome document of the field session of the Federation Council Committee on Northern Territories and Indigenous Minorities of the Russian State Duma *On the Use of Modern Information and Education Technologies to Preserve and Promote Languages, Culture and Spirituality of the Peoples of the North (on the example of the Republic of Sakha (Yakutia))* welcomes the activities of the National Library of the Republic of Sakha (Yakutia) as a leading institution which uses modern information technologies in order to preserve native languages, as well as cultural and spiritual heritage of indigenous peoples of the North.

The National Library of the Republic of Sakha (Yakutia) is the largest library in the Russian North East and one of the largest ones in Siberia and the Russian Far East. Its holdings include over 1.5 million items of documents, the most outstanding being the collections of documents in the Yakut language and the languages of small peoples of the North, as well as ethnological literature. Since its early days, the Library has focused special attention on building up such stock. In 2005-2009, the number of Yakut language books stored in the National Library increased by 13.3%, and those written in the languages of the peoples of the North – by 12,5%, thus reaching 54,356 and 1,957 items correspondingly.

The establishment of the *Sandaly-Bichik* Memorial Foundation for Documented Memory of the Peoples of the Republic of Sakha (Yakutia) is underway. It contains the volumes that came out in the territory of the republic, as well as literature about Yakutia irrespective of language and place of publication. The Foundation comprises 29,975 items.

Evgeny Kuzmin, Chairman of the Intergovernmental Council and the Russian Committee of the UNESCO Information for All Programme believes that “Yakutia is a unique place of the world. Here it is not only the language of a small titular nation that is preserved and promoted - the titular nation itself preserves the languages of the indigenous small peoples of the North too.”

The Library holdings comprise only 625 titles of books written in the languages of the small-numbered peoples of the North, including in Even (294 titles), Evenki (157 titles), Yukaghir (68 titles) and Dolgan (21 titles).

These numbers, however, could be considered quite high since only 403 book titles have come out in the territory of the republic over the last 75 years. The current publishing policy does not facilitate language preservation either. Today, the books of small peoples of the North are published in short runs of 500–1000 copies.

In 2008-2010, only 50 titles were published in the languages of the Northern peoples or dedicated to them, of which eleven items - in Evenki, eight - in Even, seven - in Yukaghir, and one - in Chukchi. These books include *Encyclopedia of Nature, Principles of Organization of the Phonetic System of the Evenki Language*, written by T. Andreeva in the Evenki language, and *Types of Evenki Heroic Sagas* gone out of print in the *Ethnic Heritage of the Indigenous Small-Numbered Peoples of the Russian North, Siberia and the Far East* series; *Grammar of the Even language in Tables and Charts*, an e-book by S. Sharina in Even; etc.

The Library is proud to house donated collections of personal libraries of such world-known Northern ethnology researchers as Glafira Vasilevich (1,675 items), Klavdia Novikova (318 items), as well as collections of books of such famous Northern ethnology researchers as Kh.I. Dudkin (200 items), Zhanna Lebedeva (150 items) and others.

We believe that success and consistency of our work in this sphere was made possible largely due to the creation in 1975 of a section of literature of the northern peoples within the Library's department of national and ethnological literature staffed by a senior manager and a librarian, which at present is unique in Russia. Once established, the section facilitated the launch of target-oriented work on compiling holdings of literature in the languages of the small-numbered peoples of the North and its promotion, and setting up library services to these peoples living in ethnic pockets.

In the years that followed the creation of the section, a project was run to study the literature read in the languages of the Northern peoples within the All-Russian experiment on library and bibliographical service to the Northern peoples. The results of the experiment were used to produce recommendations on how to optimize library service to the small-numbered peoples of the North. It was recommended to the publishers to increase, in cooperation with the library communities, the number of books going out of print in the languages of these peoples, coordinate their publication and distribution throughout the northern territories of the country, etc. In those years, the *Talking Book* project was carried out and was highly appreciated among the indigenous peoples of the North, as well as the country's library community. Under this initiative, the "narrators" were chosen among the authors who then recorded their

works in their native languages –Evenki, Even and Yukaghir. This project has given today’s users the opportunity to hear and listen to the original voices of famous writers of small-numbered peoples of the North. Availability of “talking books” for all the categories of the population irrespective of their age and level of instruction, as well as the possibility to listen to audio files virtually everywhere - be it in public, in tents of reindeer breeders, or in a nomadic school - made them popular among the inhabitants not only of Yakutia, but of the Magadan, the Kamchatka and other regions too. According to the Library’s statistics, some 3,000 people have already made use of the “talking books”. New life was brought to the project in 2007 when audio tapes were digitized and were used to create 14 multimedia disks with the original voices of writers, taletellers, scientists and other people speaking literary language of the small peoples of the North. The recordings were published at the *Knigakan* web-site of the small-numbered peoples of the North.

The activities to preserve the documented heritage of the small peoples of the North entered a new phase in 1990. That year, the State authorities increased their attention to the problems of small-numbered peoples of the North, developed the framework for national and language policy in the Republic of Sakha (Yakutia) and granted the status of official language to the languages of the peoples of the North.

Coordination Council on Promotion of Culture of the Russian North was established at the Second All-Russian Meeting on Promotion of Culture of the North, which took place in Yakutsk in 1996. Upon the decision taken at this meeting, the National Library of the Republic of Sakha (Yakutia) became the major organization responsible for the creation of information resources in the Russian northern territories, as well as for the provision of access to and mutual exchange of information on the challenges of the North. The creation of an information center on the basis of the National Library was supported by the library community.

Since then, the Library has started to work systematically and consistently on increasing the activities aiming to create and manage the use of e-resources on the small peoples of the North on the basis of a program-based and targeted method. These efforts resulted into the establishment of the following resources:

1. two reference databases, i.e. the *Repertory of books of indigenous peoples of the Republic (1932-2010)* containing 1,200 titles in Evenki, Even, Dolgan, Chukchi, Yukaghir, Yakut and Russian, and the *Books in the languages of indigenous peoples of the Republic of Sakha (Yakutia) (1932-2008)* with 434 titles in Evenki, Even, Dolgan, Chukchi and Yukaghir;

-
2. the *Repertory Catalogue of books in languages of small peoples of the North, Siberia and the Russian Far East* with 3,525 titles in 25 languages of the minorities of the North compiled in collaboration with the Russian National Library. Those databases have provided the basis for the Reference List publication and the Catalogue of Yukaghir books;
 3. full-text databases of the literature of indigenous peoples of the Republic of Sakha (Yakutia) such as the Dolganika, Chukotika, Evenkinika, Evenika and Yukaghirika databases;
 4. full-text database of the national and regional legislation on small peoples of the North.

Currently, the library website (<http://nlib.sakha.ru>) provides full access to library rebase in Russian and Yakut, as well as foreign and the North minority languages. The total of the available references is about half a million. There is also the Repertory Catalogue of the libraries in the Republic of Sakha (Yakutia) and a consolidated database of articles of the Sakha Sire containing references to publications from periodicals, serials and collections, including newspapers of small peoples of the North, since 1991. The Knigakan: Documented Memory of the Peoples of the North project sponsored by the Ministry of Culture and Intellectual Development of the Republic of Sakha (Yakutia) and the Department of Peoples and Federal Relations is an important stage in the efforts of the National Library to preserve the documented heritage of small peoples of the North. To date, 32.6% of the archives in the languages of small peoples of the North have been digitized, which has helped compile the Knigakan e-collection as a part of the e-library of the National Library. The Sakha e-library has been visited by 2,500 people since this August. 340 documents have been browsed over with 175 of them from the Knigakan collection. Now we can say that our e-library is visited not only by Russians but also users from the US, Korea and Germany. We have also registered requests for books in the Even, Yukaghir, Evenki and other languages.

However, it should be noted that collecting and preserving documented heritage of small peoples of the Russian North as well as making it available to public is yet another challenge. Indeed, the identified collection in the languages of the small peoples of the Russian North contains 1,900 titles, with 63% of them being in our possession and the other 37% kept by various Russian regions. Today, 40 small peoples of the North, with a total population of 244,000, reside in their communities in twenty eight Russian entities. For instance, Evenkis inhabit 11 Russian regions, Evens and Chukchis reside in six and three regions accordingly. In this context, a round-table held in the National Library of the republic during a field session of the Federation Council Committee on

Northern Territories and Indigenous Minorities related to the “Preservation and Promotion of Literature in Languages of Indigenous Peoples of the North and Translation at the Digital Age” discussed how to deal with the preservation issue and ensure access to the documented cultural heritage of small peoples of the Russian North.

Taking into account the discussions, the Committee supported the proposal of the round-table participants to create an inter-regional information center of documented cultural heritage of indigenous populations of the Russian North under the auspices of the National Library of the republic as well as set up its offices in local and regional libraries in twenty two Russian entities inhabited by indigenous peoples of the North, Siberia and the Russian Far East.

The center of documented cultural heritage of indigenous populations of the Russian North will coordinate joint efforts of libraries, research institutes, educational institutions, public associations in the Russian Northern territories in collecting, preserving and using documented memory of peoples of the Russian North.

The main areas of the center responsibilities will include: coordination of joint efforts of networked libraries in the Russian North; establishment and protection of information resources and their availability; translation; methodological support of unification and harmonization of the technology used to create and organize e-resources; and promotion of international cooperation with a view to setting up topic-specific databases.

The inter-regional center will be focused on developing concepts of e-resources and their use in order to:

- a. identify, collect and preserve documents, i.e. official papers, printed, audio and video materials on indigenous peoples of the Russian North;
- b. produce a repertory catalogue of books in 25 written languages of indigenous populations of the Russian North;
- c. set up issue- and topic-specific databases such as *Reindeer breeding, Fishing, Hunting, Northern territories and their environment, Nomadic school* etc;
- d. set up a database on small peoples with no writing systems; and record the speech of peoples speaking endangered languages;
- e. translate works of fiction.

To make information resources available, a relevant platform will be needed, i.e. a multipurpose portal with its specific IP-address to provide universal access to the following information and interactive resources:

-
- the Repertory Catalogue of books of small peoples of the Russian North;
 - digital documents and issue- and topic-specific databases;
 - live communication options, such as conferences, forums, discussion boards and surveys;
 - remote training module to help learn language and raise the awareness of peoples of the Russian North;

The centre will make it possible to work with new sources of information, develop innovative information services, give access to network resources of other libraries, museums and archives and protect cultural heritage of small peoples of the North and ensure its easy availability.

The Foundation for Siberian Cultures and the Use of Information Technology for the Preservation and Teaching of Indigenous Languages in Siberia

The Foundation for Siberian Cultures was established in 2010. The idea for this foundation emerged from many years of research with the peoples of the North and from initiatives for the preservation of their cultures. More recently the focus has been on Siberia and the Russian Far East.

Beyond the preservation and further enhancement of art and craft traditions of indigenous peoples, some of the main aims of the foundation are the preservation of indigenous languages and the knowledge (especially ecological knowledge) expressed in them. Furthermore, exhibitions, cultural exchanges by means of tours of artists and workshops, as well as conferences in Germany and Russia will serve to enhance mutual understandings of peoples with different cultural backgrounds and encourage valuable and productive dialogues between them.

A digital library and ethnographic collections on the world wide web provide above all indigenous communities open access to relevant scholarly resources and research materials. Recent or current projects are presented at regular shows on the web in the form of alternating photo-video shows. This provides a forum through which indigenous communities can participate and be informed about how their traditions are presented and received abroad.

Main Issues from the Foundation's Programme

Endangered languages and traditional knowledge

The main goals of the Foundation for Siberian Cultures include the documentation of endangered languages and, especially, the local ecological knowledge expressed within them. But it also aims to take decisive steps toward preserving those languages and that knowledge by means of assisting local communities to produce relevant learning tools. To realize this mandate, the Foundation is playing an active role in international discourse on these issues. During workshops and seminars local traditions and experiences are placed into wider global contexts and provide a base for next appropriate methodological steps to be taken in order to realize this programme most effectively.

Indigenous worldviews, mythologies and landscapes, and their significance for cross-cultural dialogue in the arts

In addition to anthropological interpretations of indigenous worldviews and mythologies, the focus in this programme will be on their contemporary representation in the fine arts, music and choreography. The Foundation for Siberian Cultures supports a range of creative exchanges between artists from Russia and those from other countries. Examples of such fascinating and vivid artistic dialogue was the exhibition “Shamans of Siberia” (http://www.kulturstiftung-sibirien.de/vir_21_E.html) as well as tours by Youth dance ensembles from Kamchatka in Europe (http://www.kulturstiftung-sibirien.de/ver_42_E.html). Currently in preparation are a joint German-Russian project on video art, as well as other projects on lyrics and musical compositions that reflect Siberian landscapes and find expression in a variety of western and indigenous styles.

Historical sources and their contribution to modern discourse

Another project is the publication of comprehensive travel accounts by German explorers and scientists since the mid-18th century. Their exhaustive descriptions and detailed reports are still considered some of the most valuable documents on the ethnography of the indigenous peoples of that part of the world. These works inform us about living conditions and particular ways of natural resource use at various times and provide us with valuable background information for current assessment.

This anthology, which will be published in series format as part of the “Bibliotheca Kamtschatica” by the Foundation for Siberian Cultures (edited by Erich Kasten and Michael Dürr), is enriched through essays by scholars from various historical, ethnological and natural science perspectives. In cooperation with the Institute of Geography (KBPIG, FED RAS) in Petropavlovsk-Kamchatski, a database will be produced that brings together historical accounts, recent ethnographic recordings and relevant texts in native languages, and information from the natural sciences. All will be accessible via the Internet even by the general public, and as it is intended here in the first place, be used by colleagues in Russia and indigenous communities. (http://www.siberian-studies.org/publications/bika_E.html).

Projects on the Preservation and Teaching of Indigenous Languages in Kamchatka

DVD learning tools for the preservation of the languages and cultures of the Itelmens, Evens and Koryaks

Learning tools and teaching materials focusing on indigenous communities may help to counteract forces that lead to the loss of cultural diversity and

the dissolution of local and ethnic identities. Relevant materials have been, and will continue to be, produced in cooperation with local experts and using modern technologies.

For this, the publishing arm of the Foundation for Siberian Cultures has established the DVD series “Languages & Cultures of Indigenous Peoples in Kamchatka”, edited by Erich Kasten. The DVDs are first of all aimed at the school curriculum and at cultural programmes in Kamchatka and they present specimens of the related spoken languages in monologues and dialogues. These can be used as well in international research and in university courses. The DVDs have English and Russian subtitles. Booklets contain both the translations and the transcribed original texts. For samples and video-clips see: http://www.kulturstiftung-sibirien.de/materialien_E.html > more.

The series is subdivided into:

- Itelmen Language & Culture (ILC);
- Even Language & Culture (ELC);
- Koryak Language & Culture (KLC).

Individual publications of these sub-series address to the following themes:

- The remembered past;
- Traditional ecological knowledge;
- Clothing & decorative arts;
- Ritual practice & world view;
- Human-environment relations as expressed in tales, songs and dance;
- Conferences, workshops, festivals.

Itelmen Language Classes in Petropavlovsk-Kamchatski:

This project results from an initiative of Tjan Zaotschnaja and aims to support the preservation of the Itelmen language in Kamchatka. It is being conducted in cooperation with the Munich-based group of the Society for Threatened Peoples (“Gesellschaft für bedrohte Völker – Regionalgruppe München”).

The Itelmen language is one of the most threatened languages in the world. There are only about two dozen remaining speakers of the older generation who speak the language fluently. At the same time, among Itelmen youth and in particular among those who live in the capital of Petropavlovsk-Kamchatski, there is a growing interest in maintaining their native language. For them, it embodies not only specific indigenous knowledge. It also links them to their ancestors, one of the most ancient peoples of Kamchatka.

In 2010, a first project – organized and driven by Tjan Zaotschnaja – was launched by the Foundation for Siberian Cultures in cooperation with the Munich-based group of the Society for Threatened Peoples. Funding was provided to cover travel and accommodation costs in St. Petersburg of a future Itelmen teacher to take his exams there and the publication of his first Itelmen textbook.

“Echgan” – a quarterly periodical in the Koryak and Russian languages

“Echgan” is a quarterly periodical that serves as a teaching tool in schools and other institutions of culture in Kamchatka. It is aimed at assisting in teaching indigenous themes, such as traditional ecological knowledge and arts & crafts in conjunction with the Koryak language.

This series is published by the Foundation for Siberian Cultures in collaboration with one of its partner institutions in Kamchatka, the “State Koryak Center for Arts & Crafts” in Palana. The editors of this series are Erich Kasten and Aleksandra Urkachan.

Besides its printed version for distribution in indigenous communities in Kamchatka, a digital version of each volume is also available on the web: http://www.siberian-studies.org/publications/echgan_E.html.

Shows on the Web

Recent or current projects are presented on the web in the form of alternating photo-video shows. This provides a forum through which indigenous communities can participate and be informed about how their traditions are presented and received abroad.

- January 2012 – Learning tools: DVD “Traditional knowledge in the world of Koryak fishing”.
- October 2011 – Seminar “Endangered languages and local knowledge”.
- July 2011 – Learning tools: “Itelmen language and culture”.
- April 2011 – Exhibition “The art of flying – the flight in mythology and in the art of dancing of Siberian peoples”.
- January 2011 – Learning tools: DVD “Traditional knowledge of Koryak Reindeer Herders”.

-
- October 2010 – Exhibition “250 years of German-Russian research on the nature of Kamchatka and the cultures of its indigenous peoples”.
 - June 2010 – Learning tools: DVD “Koryak Songs and Dances, Lesnaya”.
 - February 2010 – Exhibition “Shamans of Siberia”.

Planned Future Initiatives

Currently, a more comprehensive programme is in preparation by Erich Kasten and Tjeerd de Graaf on “Vanishing voices of the Asian North Pacific rim” that will build on and expand to previous activities on the preservation of indigenous languages and cultures in Kamchatka to neighboring regions and peoples as well, such as to Nivch, Nanai, Chukchi and Yukaghir.

This programme is related to an initiative by the Foundation for Endangered Languages to create Regional Interest Groups for specific areas in the world, where the languages and cultures of the local minority peoples are studied.

Katsuko T. NAKAHIRA

*Assistant Professor
Nagaoka University of Technology
(Nagaoka, Japan)*

Usage of Mother Tongue in Education in the Case of Multilingual Activities with ICT

Introduction

UNESCO and some other international organizations are regarding the use of mother tongues in the real world as an important problem from the viewpoint of the linguistic diversity. Meanwhile also in the cyberspace, accompanying the popularization of the Internet, various problems around the linguistic diversity and the use of mother tongues are increasingly occurring. It is especially important in the educational context, as education becomes borderless with the popularization of e-learning, from the viewpoint of importation and exportation of educational service, Guidelines on Quality Provision in Cross-border Higher Education (2005, UNESCO/OECD) refer to the language usage in higher education: “Higher education institutions are responsible for the quality as well as the social, cultural and *linguistic relevance of education* and the standards of qualifications provided in their name, no matter where or how it is delivered. In this context, it is recommended that higher education institutions/providers delivering cross-border higher education (a) Ensure that the programmes they deliver across borders and in their home country are of comparable quality and that they also *take into account the cultural and linguistic sensitivities of the receiving country*. It is desirable that a commitment to this effect should be made public.” The report suggested that this consideration should be applied not only to the face-to-face education but also to the remote education including e-learning.

The research on linguistic diversity in e-learning is being proceeded with an application of translation tools, for example, Language Grid. Many of the translation tools are developed for the multilingual communication with different mother tongues, evaluation for their translate function can be found in many places but there is few verification of their effect to the users. In this report, we will investigate the effect of translation tools to the users from the two case studies of e-learning education using ICT in Nagaoka University of Technology.

Case Study 1: The Language in Class Contents

First, we consider the relation between the language in class contents and knowledge acquirement of learners.

In the Nagaoka University of Technology (NUT), there are about 200 exchange students. We extract 10% exchange students (the total number of students is 27, the largest mother tongue holder were Spanish with 10.4%, Malay with 8.3%, Chinese with 4.2%, Vietnam with 3.1%, Mongolian and Indonesian each with 1.4%) and take questionnaires for language use. Seven of ten students are enrolled at NUT with the primary objective of receiving professional education (engineering), not to learn language. For the most of exchange students, the competence in Japanese (non-mother tongue) is minimum for context-embedded language. Moreover, our university has no care of teaching language to bachelor course exchange students – they have to use Japanese for classes. For master course, there are several classes given in English, but students need to be good at Japanese or English to attend a class. It means they need to master Japanese or English if they would like to make high quality study or learning. From these facts, we can point out that exchange students in our university face a deep estrangement between special subjects they hope to acquire and the context-reduced language needed to achieve their hope. However, as the exchange students are small in number and their mother tongues are diverse, it is very difficult to support the mother tongue of each foreign student. Therefore, we have developed the multilingual learning support system with ICT and researched reactions of exchange students to the system. Figure 1 shows the structure of the multilingual learning support system.

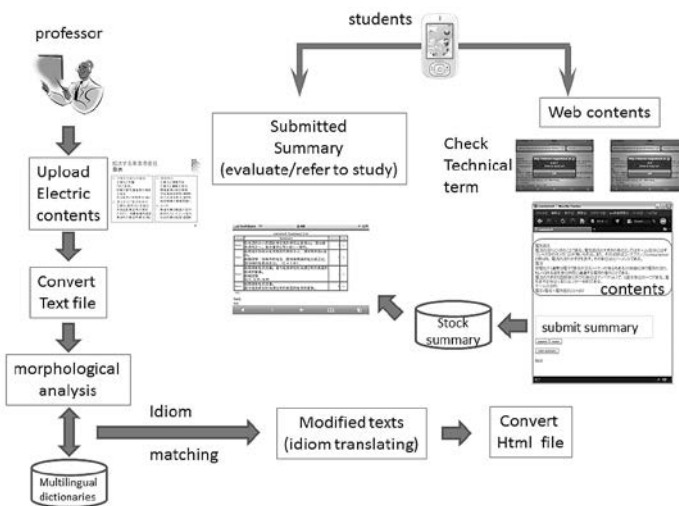


Figure 1. The structure of the multilingual learning support system

We did the research of exchange students' comprehension of contents with and without their mother tongues. We prepared 3 types of contents: A – electronic field (the rate of technical term t is 63.0%), B – mechanical field (t is 43.7%), C – information and communication field (t is 37.5%). Learners viewed the contents both with mother tongues support (type II) and without it (type I) and they were asked to evaluate their comprehension of contents with three degrees: "E1: I understood the content", "E2: I understood the content a little", "E3: I could not understand the content at all." Table 1 shows the result of the rate of understanding contents.

Table 1. The rate of understanding contents

	Contents A		Contents B		Contents C	
	Type I	Type II	Type I	Type II	Type I	Type II
E1	13%	27%	13%	20%	13%	27%
E2	20%	66%	20%	73%	20%	73%
E3	67%	7%	67%	7%	67%	0%

Seven out of ten students answered E3 in the case of type I contents, but in the case of type II contents seven out of ten students answered E2. This means that for subjects including many technical terms, especially highly specialized subjects, learners' comprehension can be clearly improved with mother tongue translation support for technical terms. Moreover, applying summary submit and evaluation function has the effect not only to encourage learners' comprehension but also to train to learn technical terms.

Case Study 2: Problem Based Language Activities in Cyberspace

The students' learning activities is not only to learn knowledge via contents. The internal thought alienation or knowledge acquired by the communication with others is also important as an achievement of educational activities. In particular, growing attention is paid to collaboration important for the promotion of learners' independent actions, student participation learning, adoption of cross-cultural communication to respond the globalization that bring drastic change to university education. We implement the Problem Based Learning (PBL) in cyberspace as an effective method of providing education smoothly (Kanematsu et. al.(2011), Nakahira et. al. (2010), Taguchi et. al. (2011), Nakahira et. al.(2011), etc.). PBL is a learning method developed by McMaster University in Canada. The learning method is designed for one or a small group of learners to be able to acquire systematic knowledge on a subject by their independent activities (such as research on the previous works, discussion and so on) to find a method to solve problems.

This method was used so far for the purpose of developing the skill to find a way to solve a problem mainly through discussion, at the place of succession of the technology. Noticing that PBL is usually supposed to be applied in a face-to-face environment, there arise two new aspects to consider when we aim at multicultural communication:

1. geographic constraint. There is a geographic limit on PBL in real world due to the practical difficulty in gathering participants from different countries.
2. linguistic constraint. In many cases we could not share the same language even if we could gather participants from the foreign countries.

To solve these problems, we construct PBL environment in cyberspace. We adopt SecondLife⁸² for developing platform and set up a classroom in cyberspace as shown in figure 2.

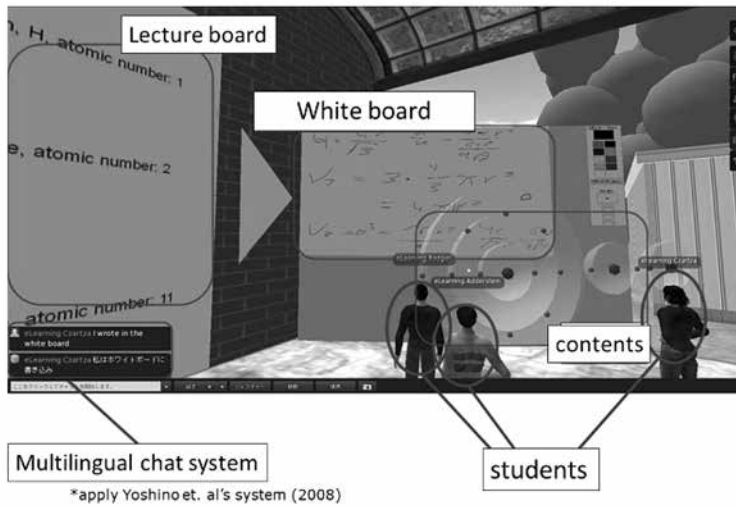


Figure 2. The classroom in cyberspace

In the classroom, we prepared a shared whiteboard and a screen for presentation. For communication in the lack of common language, we also equipped a text base multilingual chat system which adopts Language Grid⁸³ for multi-language translation. Using these tools, the learners perform learning activities that include their interaction. Figure 3 shows the PBL environment system architecture.

⁸² <http://www.secondlife.com/>.

⁸³ See Reference 3.

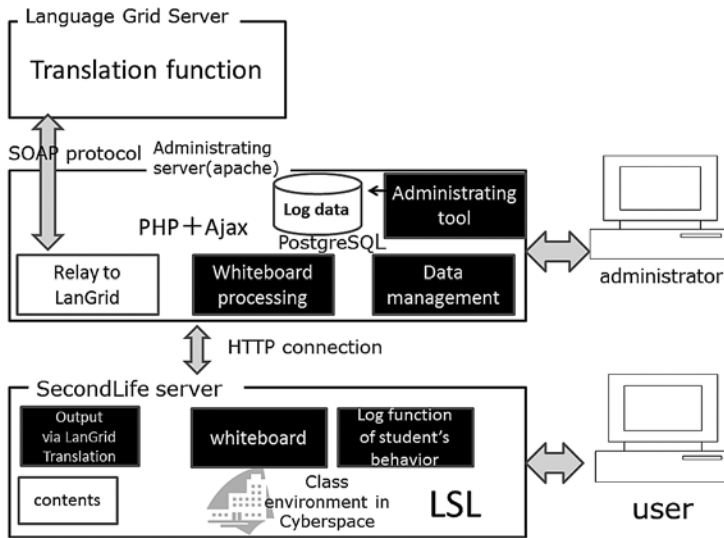


Figure 3. PBL environment system architecture

Under these preparations, we did the following experiments to evaluate PBL environment in the 3D virtual world. We collected two examinees respectively from Japanese, from German (who are familiar with both the mother tongue and English), and from Malay students, and divided them into two groups so that each group should contain speakers of respective languages. We refer to 90 minutes as one period, which is divided into two parts, a 30 minutes lecture and the succeeding 60 minutes discussion. In the lecture, a professor taught basic knowledge about a subject using chat function, and then students discussed on the subject using multilingual chat for 60 minutes. We prepared problems to ask basic knowledge of chemistry for undergraduate students as a PBL subject. The combination of theme and language is shown in Table 2.

Table 2. Environment of Evaluation Experiment

Groups	Experiment day	Learning contents	Language
Group 1	1	Cristal Structure of Metal	Only English
	2	Energy Level	Their mother tongue via Translator
Group 2	1	Energy Level	Their mother tongue via Translator
	2	Cristal Structure of Metal	Only English

Evaluation is done mainly focusing on the process of answering questions. The main points of evaluation are 1) how communication works in a multilingual chat and 2) how the shared whiteboard is used to help communication. For 1), we did a quantitative evaluation using a bibliometric method for English texts, Flesch Kincaid Reading Ease (FKRE) and the amount of conversation. For 2), we did also a qualitative evaluation by tracking whiteboard usage and communication between avatars to the level of comparison of their contents.

1) How Does Communication Work in a Multilingual Chat?

FKRE, the ratings scale for communication – measures of a multilingual chat – have the following characters : the smaller FKRE⁸⁴ value shows complex communication.

In the experiment, we define “complex communication” as “frequent use of technical terms” namely, they make discussion along the theme. Calculating these values to the contents of multilingual chat we evaluate the quality of the communication. The results are shown in Table 3, where L_e means the number of efficient chat sentences, W_m means the number of mean words per sentence, E means using English, NT means using mother tongue, Δ is (E-NT).

Table 3. The value of FKRE, L_e , and W_m .

	Group 1			Group 2		
	<i>E</i>	<i>NT</i>	Δ	<i>E</i>	<i>NT</i>	Δ
FKRE	96.8	79.5	-17.3	96.8	88.4	-8
Le	16	65	+49	21	31	+10
Wm	1.23	5.46	+4.23	3.29	4.58	+1.29

First, we focus on the values of FKRE. FKRE score of communication in mother tongues supported by translation function is clearly reduced compared with that in English. Especially the group 1 shows reduction by 17.3 points. Words and sentence structures used in communication in mother tongues supported by translation function are clearly more advanced than those in English.

Figure 4 shows the relation between the elapsed time and line number of conversation of each group either in English or in mother tongues.

⁸⁴ See Reference 2.

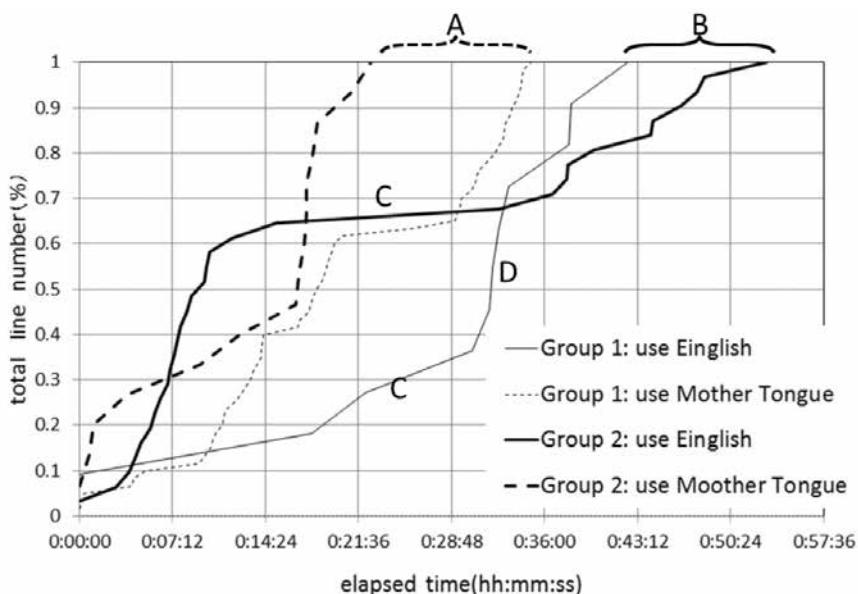


Figure 4. The relation of elapsed time and total line number to multilingual chat.

Usually when someone utters a sentence regularly in a certain interval, the chart shows linear growth. Temporal more frequent conversation pushes the graph to the upper-left, while less conversation pushes it to lower-right. We can observe from the chart that group 1 showed less communication when English was used and more frequent communication when mother tongues were used. The group 2 shows that there was delay of conversation (area B) caused by large silence (area C) when English was used. When the mother tongues were used, the graph shows a steep inclination that means a rapid interchange of conversation, and students needed just a half amount of time to solve the problems compared with the case in English (area A). From these facts, we can conclude that discussion is delayed by the language barrier when English is used, ending without getting mutual enhancement, while the use of the mother tongues makes discussion smoother, leading to some excitement by collaboration of the group.

2) How is the Shared Whiteboard Used to Help Communication?

Figure 5 shows an example in which the biggest movement was observed, displaying the communication with multilingual chat and the contents on the shared whiteboard along one time line. The timeline includes each avatar's talking and use whiteboard.

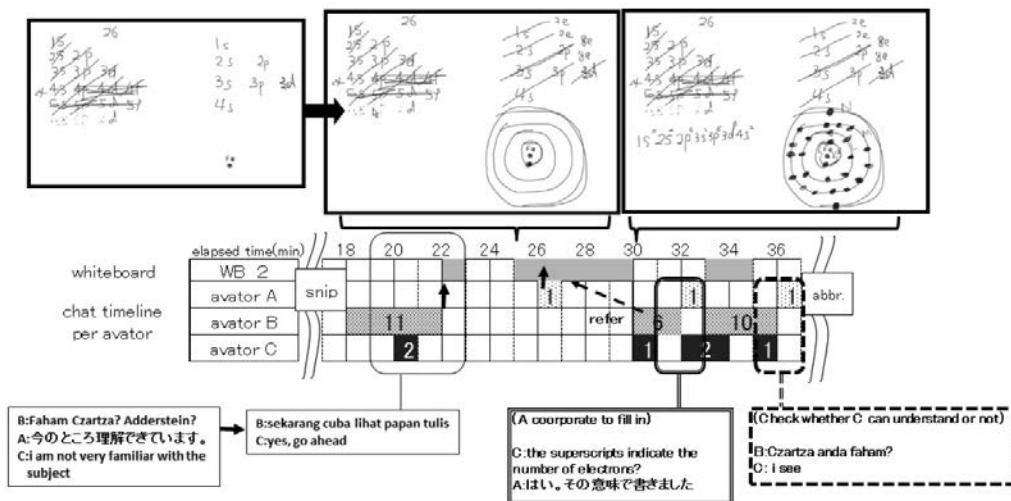


Figure 5: Example of learners' activities.

This example shows how the group 1 solved the problem on the energy level. The uppermost line in the table shows the whiteboard number used and the following three lines are colored, respectively, when avatars A, B, C did a chat. The numbers on the top of the table mean the elapsed time of discussion. The figures above the table are showing the change of contents on the whiteboard. The charts below the table show the important parts extracted from the conversations in the picked-up period.

Students shared the notion of energy level and its relation to the electron spin. We can observe how avatar B, behaving as a leader, explained them to avatar A and avatar C. Responding to avatar B, avatar A said "I can understand," and avatar C said "it is difficult to understand." The table in figure 5 starts when avatar C said "it is difficult." Answering avatar C, avatar B declared "I will explain it using whiteboard" and explained his idea using the chart. In the process avatar A also got engaged in the explanation and avatar C asked questions to avatars A and B. We could judge that they started having smooth communication. We can observe that avatar C also understood the subject from the analysis of the chat. By contrast, the group 2 did not have a leader, and there was an avatar who could not follow the discussion on the same problem that was given to the group 1. The conversation among the members of the group 2 was sparse and the number of sentences in it is just a half of the group 1. We also point out that the contents described on the shared whiteboard were not clear. From these facts, we can conclude that to promote students' learning, it is not sufficient just to prepare the discussion environment in mother tongue – an avatar acting as a leader or a mentor is also needed.

Discussion and Summary

Through these two experiments, we figure out the relation between the language that the learners use and the learning activities/acquirement of knowledge as follows.

From case 1 we find that translation of non-mother tongue technical terms contained in the class contents into their mother tongue is useful for acquisition of knowledge. Usually, however, non-mother tongue and mother tongue substitution is achieved using dictionaries. Hence, as a precondition for this translation to work, learners should have sufficient knowledge about technical terms in the mother tongue as a context-reduced language.

From case 2 we think that there are two important aspects in developing collaboration study between students having different mother tongues: (1) training common language (English) and (2) training general communication skill so that students can express their own opinion. In the context of PBL the main focus is on the second point, i.e. our educational goal is to enhance communication skills. For this purpose the use of mother tongues is more appropriate than obliging students to study common language. This is proven by our experiments: dramatic increase of the amount of conversation, appearing a leader in the group, and effective use of support tools.

To summarize, to provide collaborative education smoothly in multilingual environment it is desirable to support mother tongues for enhancing both learning activities and acquirement of knowledge. Although this is too difficult to achieve in the usual face-to-face environment, we demonstrated that we can support the use of mother tongues in multilingual environment by introducing ICT. It was proven to be one of useful ways to connect an online dictionary or an online translator with class contents and a chat tool. However, such tools have performance boundary, especially if students' mother tongues lack required technical terms. In that case, to support those students, a little more extra effort is required: one possibility is to neoterize these technical terms. However, such new technical terms tend to get widely spread quite rapidly in accordance with the permeation of education. In future, we will be able to automatically enrich the online dictionary/translator function using the feedback of observing new phrases or knowledge through, say, web crawling.

Acknowledgement

We are thankful to the Center for e-Learning Research and Application (Nagaoka University of Technology) and Hideyuki Kanematsu (Suzuka National College of Technology) for providing many valuable data. We also

thank Prof. Yoshiki Mikami, the Director of Language Observatory and Dr. Shigeaki Kodama for useful discussion.

References

1. Sahar Farjami, Ryosuke Taguchi, Katsuko T. Nakahira, Yoshimi Fukumura, Hideyuki Kanematsu. Multilingual Problem Based Learning in Metaverse. Proceedings of the 15th International Conference on Knowledge-Based and Intelligent Information and Engineering System, in printing, 2011.
2. R. Flesch. A new readability yardstick. In: Journal of Applied Psychology, 32, pp. 221-233, 1948.
3. T. Ishida. Language Grid: An Infrastructure for Intercultural Collaboration. IEEE/IPSJ Symposium on Application and the Internet, pp. 96-100, 2006.
4. Katsuko T. Nakahira, Rodrigo N.R., Ryosuke Taguchi, Kanematsu Hideyuiki, Yoshimi Fukumura. Design of a Multilinguistic Problem Based Learning Learning Environment in the Metaverse. Proceedings of 2nd International Symposium on Aware Computing, Taiwan, IEEE, pp. 298-303, 2010.
5. Katsuko T. Nakahira, Ryosuke Taguchi, Rodrigo N. R, Hideyuki Kanematsu, Farujami Sahar, Yoshimi Fukumura. Implementation of Learning System for Multilingual PBL. IEICE Technical Report AI2010-57, p.81-86, 2011.
6. R. Taguchi, Katsuko T. Nakahira, H. Kanematsu and Y. Fukumura. A Multilingual Problem-Based Learning Environment for Awareness Promotion. Proceedings of the 16th International Symposium on Artificial Life and Robotics, 150-153, 2011.

Anatoly ZHOZHNIKOV

*Director, New Information Technologies Centre,
North-Eastern Federal University
(Yakutsk, Russian Federation)*

North-Eastern Federal University: Strategic Action Lines to Promote Cultural Diversity of Indigenous Communities of the Russian North in Cyberspace

The cultural and linguistic diversity of the Russian North's small indigenous peoples has been an object of increasingly close attention of late. Attesting to this is the ever-growing number of forums focused on the topic.

The 1st International Conference on Linguistic and Cultural Diversity in Cyberspace, which took place in the Republic of Sakha's capital, Yakutsk, in July 2008, is among the most important such events held so far, with its final document, the Lena Resolution, acclaimed globally.

Another recent international gathering of note is the symposium "Circumpolar Cultures in the Museums Worldwide: Past, Present and Future," hosted by the Yakut capital July 29 through August 1, 2009. Held under UNESCO's patronage, the event focused on the following issues:

- Dialogue of Subarctic cultures in the context of cultural diversity;
- Preserving and passing on the ethno-cultural heritage of Subarctic communities;
- Museology and the use of information technology;
- Role of the world's museums in enhancing knowledge about circumpolar cultures.

This conference brought together about a hundred delegates from 13 of Russia's regions, as well as from 14 foreign countries. Its final documents include Recommendations on Linguistic and Cultural Diversity:

- consolidating the museum community's efforts to organize expeditions in search of new acquisitions, as well as to process, study and conserve circumpolar heritage objects already held by museums across the world;
- stepping up work to promote knowledge about the diversity and cohesion of circumpolar cultures through a Web links page, to be created by partner nations in association with UNESCO and the International Council of Museums (ICOM).

A new web portal, *Circumpolar Cultures in the Museums Worldwide: Past, Present and Future* (www.arcticmuseum.com), was presented at the forum, along with an eponymous DVD disc and a colourful album released by the UNESCO Sakha Committee and the republic's Northern Indigenous Communities History & Culture Museum.

The North-Eastern Federal University's Innovative Information Technology Centre provided software, content, DVD design, and hosting services for this project. Both the links page and the DVD earned praise from the ICOM administration.

Yakutsk was also the venue of a symposium entitled "Art Education in the Subarctic Cultural Environment," held on November 26-27, 2009. This one covered seven focus areas, including three related to linguistic and cultural diversity:

- Indigenous Subarctic communities: Traditional culture & folklore studies;
- Fine arts at Subarctic institutions of art education;
- Modern trends in creating library & information spaces in an arts-focus academic environment.

Recommendations adopted by these conferences, as well as existing circumpolar heritage collections, are of significant scholarly and cultural value, and they could prove instrumental in efforts to preserve and promote linguistic and cultural diversity in cyberspace.

The international conference "Linguistic and Cultural Diversity in Cyberspace" and its final document have become major reference points for practical efforts in the area. The reports and essays contributed by the forum's Russian and foreign participants are a most valuable source of inspiration, as are all related materials published by the Russian Committee of UNESCO Information for All Programme (IFAP).

Based on the Lena Resolution recommendations, we suggest three strategic tasks to be carried out to pave the way for online representation of the linguistic and cultural diversity of the Russian North's indigenous communities:

1. Representing the scripts of all indigenous languages of the Russian North in computer operating systems;
2. Creating a Web portal with exhaustive information on the cultures and languages of small indigenous peoples of the North, along with opportunities for online communication in indigenous languages;
3. Establishing a related department within the North-Eastern Federal University.

To make it happen, a number of wide-ranging issues will have to be addressed, including political, state, regional, financial, technical, and organizational.

Some of these were considered by the Committee for the Small Indigenous Peoples of the Russian North at the Federation Council (parliament's upper house) during its away session in Yakutsk on November 16–18, 2010.

The session focused on the use of modern information and educational technology in preserving and advancing the native communities' traditional cultures and languages, with Sakha's expertise taken as an example. It approved all related initiatives from the North-Eastern Federal University and incorporated them in its own Recommendations.

Here are some of the decisions made by the gathering:

- supporting the university's idea to provide computer operating systems with a universal keyboard layout that would include the scripts of all indigenous languages of Russia's Subarctic, Siberia and Far East; providing assistance in translating this plan into reality;
- for the Education and Science Ministry to sponsor – financially and otherwise – the North-Eastern Federal University and the Siberian Federal University in their efforts to develop a federal Web links page on the small indigenous peoples of Russia's Subarctic, Siberia and Far East;
- examining the possibility of setting up an indigenous languages translation department at the North-Eastern Federal University, so as to create a scholarly basis for the development of translation and interpretation from these languages.

Sufficient funding is yet to be found to ensure the implementation of all adopted resolutions, though. With this purpose in mind, a North-Eastern Federal University ad hoc group has been working extensively to get the above projects included in the following republican target programmes, intended for the year 2012 and beyond:

- Preserving, studying and advancing the official and state languages of the Republic of Sakha (Yakutia) in 2012 through 2016. This programme will involve studying indigenous language scripts, identifying those not yet represented in computer operating systems, and devising online spelling thesauri and MS Word morphology parsers and spell-checkers for each indigenous language;
- Developing the Republic of Sakha's information and communication infrastructure, where indigenous language scripts and adapted keyboard layouts for computer operating systems could be provided, along with support for the multilingual Web links page in Russian

and English, as well as in 16 of the languages spoken by the Russian North's natives.

The most tangible of efforts have been those made as part of the North-Eastern Federal University's "Programme for Preserving and Advancing the Yukaghir Language & Culture on Digital Carriers and in Cyberspace in 2011-2014" (Project 4.1, Event 2.35).

In the first semester of 2011, five educational DVDs with rich content in the Yukaghir language were released and a basis for a Web links page on the Russian North's small indigenous peoples was created (www.arctic-megapedia.ru). Much of the content compiled for the site is already available online.

The materials featured are intended primarily for school and university students pursuing courses on the Yukaghir language and traditional culture, as well as on the folklore of this and other indigenous communities of the Russian North.

Work is currently underway to engage indigenous communities based in adjacent regions. Maintaining linguistic and cultural diversity in cyberspace and at the www.arctic-megapedia.ru site was a focus of the interregional conference "Continuous Vocational Training as a Prerequisite for Sustainable Community Development," held in April 27-29, 2011, in the Chukchi city of Anadyr.

Conclusions made in the course of related discussions found their way into the forum's final recommendations:

- Developing joint interregional projects to represent scripts of small indigenous peoples of Russia's Northeast in computer operating systems;
- Setting up an ad hoc group for gathering material on the linguistic and cultural diversity of the Russian North's native communities populating the Chukchi Autonomy and posting that material at the North-Eastern Federal University's website (www.arctic-megapedia.ru);
- Elaborating a scientifically justified programme for introducing innovative technology into the traditional economic practices of the Chukchi Peninsula's indigenous population.

In keeping with the Agreement on Cooperation between the North-Eastern Federal University and the Russian IFAP Committee, the Centre to Advance Multilingualism in Cyberspace was established (Chancellor's Order dt. September 1, 2010; Registration No. 93-OD). Its primary aim is to stimulate and coordinate efforts toward sustaining the linguistic and cultural diversity of the Russian North's indigenous communities. The centre has been instrumental in organizing the second international conference "Linguistic and Cultural Diversity in Cyberspace".

Liudmila ZAIKOVA

*Director, Centre to Advance Multilingualism in Cyberspace,
North-Eastern Federal University
(Yakutsk, Russian Federation)*

Establishing the Centre to Advance Multilingualism in Cyberspace under the North-Eastern Federal University

Promoting linguistic diversity in cyberspace has both cultural and political significance. Dissemination of multilingual information on the history, languages and cultures of different nations facilitates mutual understanding and tolerance development.

Each language is a unique repository of the information on its speakers and their culture. Recently, however, the process of languages extinction is becoming rampant. According to pessimistic forecasts, by the end of the XXI century, only about 10% of present-day languages may survive.

Numerous factors threaten the existence of a language: natural disasters, leading to the death of entire peoples; weaknesses of the education system discouraging children to learn school subjects in their native language; the lack of writing to name a few. Any language's extinction is a great loss, as languages reflect historical experience and serve as a tool for socialization, expression and transmission of social and cultural traditions. While facilitating the growth of human knowledge, languages are a means of enhancing self-identification especially important for their speakers.

In recent years a new factor has appeared, namely the rapid development of ICTs and the Internet.

The Internet offers huge opportunities for users in terms of freedom of expression, education and access to information. However, information and services are accessible on the Internet only in dominant languages (about 400 out of the existing 6700 languages).

To help maintain linguistic diversity and its development in cyberspace Centre to Advance Multilingualism in Cyberspace was established under the Department of Information Policy and Communication Technologies of the North-Eastern Federal University named after M. Ammosov (NEFU) by the order of the NEFU Rector Evgenia Mikhailova on September 1, 2010.

The idea of creating such a center in the Republic of Sakha (Yakutia) has arisen at the International Conference "Linguistic and Cultural Diversity in

Cyberspace”, which was held in Yakutsk, on 2-4 July 2008 and became Russia’s contribution to the International Year of Languages.

Many universally renowned participants of the conference, held in such an interesting place as Yakutia, found it the most significant of international forums that gathered in the International Year of Languages at such a high level of organization and content. The Government of Yakutia placed the conference among its three greatest achievements of 2008, on a par with the Year of Family and the Children of Asia international games.

The final document of the conference, known as the Lena Resolution, was one of its major achievements. This landmark document – an appeal of the conference to the host country and the world – emphasizes the importance of preserving languages, in particular, for the implementation of the Recommendations of the World Summit on the Information Society and for universal access to information and knowledge.

The resolution advances initiatives for the development of linguistic and cultural diversity in cyberspace – particularly the organization of a world summit on linguistic diversity in cyberspace under the joint aegis of UNESCO and the World Network for Linguistic Diversity (MAAYA). A set of measures is proposed to preserve minority languages and cultural heritage. The resolution stresses the necessity of drawing relevant programmes, and of improving international, national and local policies in this sphere.

Yakutia’s fruitful activities for the preservation of linguistic and cultural diversity earned enthusiastic appreciation by conference participants. They asked the Government of the Republic of Sakha (Yakutia) to consider the opportunity for establishing in Yakutia an expert centre for the study and development of multilingualism in cyberspace and global exchange of experience in that sphere.

A special two-day seminar on the problems and prospects of establishing such a centre was held in Yakutsk on July 20-21, 2010. The seminar was organized by the Russian Committee of UNESCO Information for All Programme, the Interregional Library Cooperation Centre, the North-Eastern Federal University, the Council on Language Policy under the President of the Republic of Sakha (Yakutia) and the National Library of the Republic of Sakha (Yakutia) with the support of the UNESCO Moscow Office and the Ministry of Culture and Spiritual Development of the Republic of Sakha (Yakutia).

Evgeny Kuzmin, Chairman of the Intergovernmental Council and Russian Committee of the UNESCO Information for All Programme, President of the

Interregional Library Cooperation Centre, made a keynote on the establishment of the Centre to Advance Multilingualism in Cyberspace.

On July 21, 2010 the NEFU and the Russian Committee of UNESCO Information for All Programme signed an agreement on cooperation in the support and promotion of multilingualism in cyberspace.

Special measures taken in Yakutia to preserve and develop the Yakut language deserve attention and study. Yakutia is a unique place in the world from the point of language preservation. It not merely develops the language of its small titular ethnic group – this group, in its turn, works to preserve the languages of the Northern indigenous ethnic minorities: the Evens, Evenkis, Chuckchis and Dolgans. But in the face of all these measures, recent studies indicate a steady decrease in Yakut as a first language among ethnic Yakut speakers.

Establishing the Centre to Advance Multilingualism in Cyberspace aims at weakening the negative and enhancing the positive trends in this field. The Centre's activities should focus on assisting all institutions and organizations that are eager and able to contribute to the maintenance and strengthening, equipping and promoting linguistic diversity in cyberspace.

The Commission of the Russian Federation for UNESCO considers working to promote multilingualism in cyberspace as one of the top priorities, linking it to policies to support the Russian language and save it as a language of international communication.

The project of the Centre's development within the framework of the North-Eastern Federal University Development Programme received approval by Vladimir Putin, Chairman of the Government of the Russian Federation.

The Centre is facing the task of preserving and promoting linguistic diversity in cyberspace as cultural and spiritual heritage of the Republic of Sakha (Yakutia) and the whole Russia.

The Centre's key activities include:

- Initiating, organizing and conducting research in the field of linguistic diversity preservation and development in cyberspace involving NEFU staff as well as other organizations and institutions of education, science, culture, mass media, civil society and private sector.
- Studying and promoting Russian and foreign best practices on the preservation of linguistic diversity and its development in the cyberspace, facilitating its development and use.

- Facilitating the creation of centres of excellence and the development of activities by various institutions and organizations in the sphere of promoting linguistic and cultural diversity in cyberspace.
- Contributing to the enhancement of regional, national and international policies and regulatory framework in the field of culture, education, communication and information for the support and development of linguistic and cultural diversity in cyberspace.
- Providing assistance in the preparation and publishing of scientific, educational, methodological, and other materials to draw attention to the problems of preservation of linguistic and cultural diversity of the Republic of Sakha (Yakutia) and the Russian Federation from students and graduates, teachers, researchers and other staff members of the NEFU and other institutions of education, science and culture, as well as public at large.
- Providing methodological, organizational, information support for the activities by memory institutions to promote multilingualism in cyberspace.
- Facilitating collaborative efforts by organizations and institutions to support linguistic diversity in cyberspace.
- Contributing to organizing and implementing activities to document, preserve and develop the indigenous languages of the Republic of Sakha (Yakutia) and other peoples of the Russian Federation with the use of ICT.
- Assisting the development and dissemination of appropriate free software, fonts, and other technical tools necessary for using various languages in cyberspace.

The Centre to Advance Multilingualism focuses on encouraging and assisting all entities and individuals who can and should contribute to preserving, strengthening, equipping and developing languages, in particular in cyberspace.

Sargylana IGNATYEVA

*Rector, Arctic State Institute of Arts and Culture
(Yakutsk, Russian Federation)*

Arctic Arts & Culture Institute: Creating Online Resources on Russia's Northern Indigenous Communities' Culture

The Arctic State Institute of Arts and Culture (ASIAC) is Russia's only higher educational establishment with a mission to protect the culture of the small indigenous peoples of the North and to pass it on to future generations nationwide through the system of public education.

The institute's strategic goals include implementing study programmes as well as creative and research projects of high academic quality, exploring ways to preserve indigenous communities' traditional culture, and promoting what is known as "living heritage" across Russia and the world using innovative educational, socio-cultural and information technology.

Founded in 2000, ASIAC has since evolved into an important cultural centre, making a difference in community life across the region. While preserving the best of teaching traditions, it seeks to be relevant to the modern-day requirements for national education. The development of a quality management system is currently underway here, with the teaching staff trying new creative approaches and employing innovative teaching techniques.

At ASIAC, conditions have been created for studying mother tongues and foreign languages, with these latter crucial to the institute's integration into the world's academic community. An Arctic dimension has been added to the curriculum so as to raise the students' awareness of local indigenous heritage and foster their commitment to preserving and promoting it, including through public education institution.

ASIAC aims to raise public awareness of the value system and the traditions of northern indigenous communities' spiritual culture, making use of modern information technology. IT is crucial to its ambition of growing into a major centre of education, science and culture in North-Eastern Russia, whose rich intellectual and information resources could benefit the nation and the world at large. It is with this goal in mind that the institute works to expand its online resources, including the academic e-library. The library is comprised of databases featuring all kinds of material essential for learning, including digitized textbooks, teaching aids, methodology essays, and so on. A Web interface has been developed for the convenience of authorized users' access to those of the resources whose use should be confined to the local network.

ASIAC's study programmes will soon be enriched with a new modular on indigenous Arctic heritage. A number of original multimedia teaching aids have been developed especially for this modular, including Yuri Sheikin's "Musical Culture of the Peoples of the North" and "Musical Culture of Siberia's Reindeer-Breeding Communities."

Its IT-related achievements include the creation of three links Websites with original content on culture and the arts of the subpolar region's indigenous inhabitants: "Yakutia's Memory," "Kuyaar," and "Circumpolar Civilization in World Museums: Past, Present, Future."

"Yakutia's Memory" (<http://www.sakhamemory.ru>) is a collaborative project designed by ASIAC in association with Yakutia's Republican Library and Archives, the Yakut State Museum of History and Culture, and the Republican Heritage Film Archives. The project is aimed at facilitating preservation and access to valuable documents on the arts and culture in the Republic of Sakha (Yakutia), including manuscripts, printed books, film footage, photographs, and audio recordings. It incorporates the following six subprojects:

- Yakutia: Literary monuments;
- Bibliography of the Republic of Sakha (Yakutia);
- Voices of the Century: Music and Audio Heritage of the Republic of Sakha's Indigenous Communities;
- Yakutia Archives: Documents of Value;
- Yakutia: Film Chronicles;
- Yakutia: Photographic Chronicles.

"Kuyaar" web site (www.kuyaar.ru) provides a window on the cultural diversity and education of Yakutia's small indigenous peoples. It is run in partnership with UNESCO's Moscow Office, and benefits from technical assistance by the North-Eastern Federal University's Centre for Innovative Information Technology. The ambition of this multilingual site, available in Russian, English, Yakut and Evenki, is to contribute to the promotion and further development of Yakutia's indigenous communities' cultural diversity, creative activity and communication by spreading knowledge about their heritage. The idea was suggested by the republic's Deputy Prime Minister Feodosia Gabysheva and supported by the Minister of Culture and Spiritual Development, Andrei Borisov. Content for the Website www.kuyaar.ru is selected with a view to promoting the culture and folklore of Yakutia's indigenous peoples, generalizing the experience of their cultural practices in

education, systematizing research in their arts and culture, and accumulating creative ideas and intellectual resources to preserve their cultural diversity.

“Circumpolar Civilization in World Museums: Past, Present, Future” (<http://arcticmuseum.com/>) has also been developed in association with UNESCO. This project shows how, despite the harsh climatic conditions, northern Russia’s subpolar communities have managed to survive and develop highly distinctive cultures and lifestyles. They are part of a civilization referred to by the research community as “circumpolar” and dating back thousands of years.

Here are some of the proposals concerning further efforts to support and promote the northern indigenous peoples’ languages, arts and culture:

1. In order to improve native subpolar communities’ access to information resources, the elaboration and implementation of the Programme for Promoting Indigenous Cultural and Linguistic Diversity should be recognized as a priority of Russia’s Information Society Development Strategy (this task will be hard to fulfill without serious government support in the form of a target programme);
2. Setting up an association of developers of online resources on indigenous peoples’ cultural and linguistic diversity and building an integrated e-library based on digitized content from Yakutia’s Republican Library, complete with a common electronic catalogue;
3. Recommending to the Yakut government that it place a state order for programmes to train, retrain and upgrade specialists in culture-related IT at the Arctic Institute of Arts and Culture, in line with the new, third-generation national educational standards.

SECTION 3. CREATING POSITIVE ENVIRONMENT FOR THE PROMOTION OF LINGUISTIC AND CULTURAL DIVERSITY

Dieter W. HALWACHS

*Professor, University of Graz
(Graz, Austria)*

Linguistic Diversity, Dominated Languages and the Internet: The Case of Romani

The internet in its still open and democratic mode is the easiest way for speakers of dominated languages to present their languages to the wider public. Nowadays all languages of the world whose speech communities have access to modern technology can be found on the internet. However, their virtual presence differs in function, mode, style and type. With respect to the latter, a distinction has to be made on the basis of the difference between societal multilingualism and individual plurilingualism.

- Multilingualism on the internet is a static, additive phenomenon which presents dominated languages above all alongside dominant languages. Functionally this type has to be described as emblematic and symbolic. Because of its static and additive character, multilingualism on the internet is measurable or rather countable. However, the meaningfulness of such statistics remains highly questionable.
- Plurilingualism on the internet is a dynamic interactive phenomenon which presents languages in their actual use. Functionally this type has to be described as integrative and communicative. Although plurilingualism on the internet shows the actual language use, it is, above all because of its dynamic and transient character, only qualitatively describable and not countable.

This paper discusses this fact and the resulting differences in the case of Romani, a dominated language of Europe.

1. Romani

Romani is the language of the *Roma*, *Sinti*, *Calé*, and of all other ethnic groups in Europe who speak or spoke an Indic language. These population groups are collectively called “Gypsies”, a term mostly used derogatively.

From a genetic point of view Romani belongs to the Indo-Aryan sub-branch of the Indo-Iranian branch of the Indo-European language family. It is one of the Indo-Aryan diaspora languages that are almost exclusively spoken outside of the Indian subcontinent. Romani has been spoken in Europe since the Middle Ages.⁸⁵ Although its genetic affiliation is Indo-Aryan it is an integral part of the linguistic and cultural plurality of Europe.

From a structural point of view, Romani may be described as a heterogeneous cluster of varieties with a homogeneous core – a common morphology and a common lexicon – but without any generally accepted homogenizing standard. The latter is often perceived as a shortcoming but is characteristic of the vast majority of the languages of the world. The same applies for the characterisation of Romani as a variety cluster. Heterogeneity, or rather plurality of dialects and varieties, is a typical feature of the languages of numerically large and geographically widespread speech communities.

The sociolinguistic situation of Romani reflects the sociopolitical and sociocultural status of its speakers. Romani has to be described as a primarily oral, functionally restricted, dominated, stateless diaspora language with no monolingual speakers. For most Roma their respective Romani variety is reduced to intra-group communication, and thus limited to the private domains. Romani primarily functions as an intimate variety in the social microcosm. Adult Romani speakers are always plurilingual and use the language(s) of the respective majority population(s) for inter-group communication in the public domains and increasingly also in the domains of everyday life. Consequently, Romani speakers may be described as non-monolingual.

Romani is a language that, until recently, was exclusively passed on orally and did not exist in a written form. It has not developed a generally accepted written standard and, as a consequence, prescriptive norms. This linguistic situation reflects the sociopolitical situation of the Roma: politically, economically and culturally marginalised as well as ethnically stigmatised, discriminated against, and persecuted up to genocide, the Roma could only survive in small groups which led to the geographical and social heterogeneity that still exists today. Consequently, Roma have never been in a position to build large political-economic structures or to demand their share of political and economic power. Considering that the development of standard varieties generally follows the development of political and economic power structures, it is clear why Romani has not developed such a variety. Furthermore, the outlined sociohistorical situation explains why Romani is labelled a stateless diaspora language.

⁸⁵ Accessed 2012-02-24.

In the course of the emancipation of the Roma as a European ethnicity, Romani became one of the primary features of a transnational Roma identity. This process, which started in the 1970s, is mainly based on a common history and culture and consequently aims for a common language. This resulted in various attempts to standardise Romani, which have more or less failed so far. However, concrete measures to codify individual varieties and expand their functions from mainly oral use in informal situations into formal written usage have been successful to some extent. This expansion of Romani into formal domains resulted in literate forms of Romani and *inter alia* also in its use on the internet.

2. Romani on the Internet

The use of Romani on the internet is best described by the examples of various web pages which were set up as a result of the emancipation of the Roma as an ethnolinguistic community of Europe. Romani is used on websites of:

- national and international Romani NGOs,
- international NGOs and organisations,
- public and private Romani media.

2.1. Online Multilingualism

Romani is a dominated language. Romani speakers are always plurilingual and mainly use the respective dominant language in formal written domains. Consequently, there are no monolingual Romani web pages on the internet. Romani texts primarily accompany texts in dominant languages which is in line with the additive type of linguistic diversity on the internet.

2.1.1. National NGOs on the Web

Almost as a rule, the language of the state in which the respective NGO is active dominates its web pages. Some websites are almost exclusively in the dominant language as for instance the pages of the Austrian cultural association of Roma, the *Kulturverein Österreichischer Roma* <<http://www.kv-roma.at/>>.⁸⁶ Only the title of the association's journal *Romano Kipo* 'Romani picture' and the link to an archive of radio broadcasts in Romani – *Roma Sam* 'We are Roma' – are in Romani. All other texts are exclusively in German, on the website as well as in the journal. The same applies for the web pages of the official Romanian centre for Romani culture, the *Centrul Național de Cultură a Romilor – Romano Kher* <<http://romanikultura.ro/>>. The URL *Romani Kultura* 'Romani culture' and

⁸⁶ Accessed 2012-02-24.

the extension of the centre's name *Romano Kher*, 'Romani house', and some words in the headlines of events such as *khetanes*, 'together', in *Seminarul "Khetanes"*, the title of a seminar organised by the centre, are in Romani. All other texts are in Romanian.⁸⁷

Websites of Roma NGOs that are geared towards multilingualism often intend to offer versions of their web pages in Romani. But in many cases this is only hypothetical. If at all, only a few pages are in Romani. Some examples include:

- Besides the name *Romanothan*, 'Romani Place/World', the website of the Romanian NGO *Romanothan* <<http://www.romanothan.ro/>> is almost exclusively monolingual Romanian in its written form. However, it offers a video language course to get an impression of Romani which is bilingually announced in written form and uses oral Romani alongside Romanian.
- The website of the Barcelona/Spain-based *Unión Romani* <<http://www.unionromani.org>> has a button symbolising the Romani flag for the Romani language version which in actual fact is only a repetition of the Spanish version with some Romani words in the navigation menu.
- *Rom News Network* <<http://www.romnews.com>> from Hamburg/Germany claims to offer information in English, French, German, Romani, and Spanish. Yet only one version in a mixture of German and English is available.
- The organisation *Dženo*, 'People', <<http://www.dzeno.cz>> produced pages in Czech, English and Romani until mid-2010. However, the full range of information was only available in Czech. Some information in English was offered until 2009, the last Romani entry dates to 2007 which coincides with the end of the funding period for translations into Romani.
- The Vienna/Austria-based NGO *Romano Centro* <<http://www.romano-centro.org>> offers the option to switch between German, English, and Romani. Whereas many pages are available in English, there is only one page in Romani: notably the presentation of the organisation with the menu headings: *Über uns – About us – Pa amende*.⁸⁸
- *Romani CRISS – Centrul Romilor pentru Intervenție Socială și Studii/Roma Center for Social Intervention and Studies* <<http://www.romanicriss.org>>, a Romania-based NGO, only offers English translations of much of the information given on the website. *Romani CRISS* alludes to the homophonous *Romani kris* 'Romani

⁸⁷ Accessed 2012-02-25.

⁸⁸ Another symptom of the low status of Romani in the web presentation of this organisation is the fact that the heading of the Romani text still remains German: *Das sind wir.../That's us ...*

court of justice’, a strategy in line with using Romani exclusively for headings, titles and names.

- The use of Romani only in the name of the organisation also applies for the web presentation of the Kumanovo/Macedonia-based NGO *Daja*, ‘Mothers’, <<http://daja.org.mk>>. Apart from the official denomination, the entire website is in Macedonian with some texts translated into English.

Such bilingual websites using English in addition to the respective dominant language of the country where the NGO is active demonstrate not only the importance of English as an international language but also the functional restrictions of Romani as a dominated language outlined in chapter 1. The written style in public political domains of a state is fully covered by the dominant state language. Thus the scope of all other languages spoken on the territory of the respective country is limited to some other formal written domains, for instance education, and mainly to oral informal domains of everyday life and the private sphere. To spread information beyond the national level and to produce an effect on the international level, the use of an international language is indispensable. Therefore, English in its function as the global *lingua franca* is used alongside the national language. The use of Romani in web presentations of Romani NGOs is more or less mere symbolism. It is used for identity flagging, for indicating cultural independence, for maintaining ethnocultural status, etc., but almost never for communicative purposes.

2.1.2. International Websites

Supranational organisations and Roma NGOs acting primarily on the international level as a rule use English in its globally dominant function as the primary language for their web presence. Examples of such web pages are listed below:⁸⁹

- The web presentation of the *European Roma Information Office* – ERIO <<http://erionet.org>> is exclusively in English. This international Brussels/Belgium-based NGO, which is almost exclusively run by Roma, aims to be the sole legitimate representation of Roma to the *European Commission* and the *European Parliament*.
- The *European Roma and Travellers Forum* – ERTF <<http://www.ertf.org>> based in Strasbourg/France is an international umbrella organisation of Roma associated to the *Council of Europe*. On its website only the name is trilingual: *European Roma and Travellers Forum* – *Europako Forumo e Romenge thaj e Phirutnenge* – *Forum Européen des Roms et des Gens du Voyage*. The French denomination is in line with the tradition of language use in the European supranational context.

⁸⁹ Accessed 2012-02-25.

-
- The web presentation of the *Council of Europe* <<http://www.coe.int/>> is traditionally bilingual in English and French. As a rule, web pages dealing with Roma matters also offer documents in Romani; see among others the web presentations of the project *Education of Roma Children* and the department for *Roma and Travellers*.⁹⁰
 - The Budapest/Hungary-based *European Roma Rights Centre* – ERRC <<http://www.errc.org>> is an internationally-oriented NGO which was initiated by *Gadže* ‘non-Roma’ with the aim of including activists of Roma background into the struggle for human rights for Roma in all parts of Europe and beyond. Recent activities in the Russian Federation have resulted in documents in Russian on the ERRC’s website. Translations into Romani have been presented almost from the very beginning of the ERRC’s web activities. A selection of such documents is available at <<http://www.errc.org/romani>>.

The use of Romani on international websites is obviously linked to the ethnic background of the majority of persons involved or dominating the respective organisation. On the websites of international NGOs run and dominated by Roma, the use of Romani seems to be a rare exception. On the websites of organisations initiated and led by *Gadže*, Romani is almost as a rule present on web pages in the form of translations. At first sight this correlation – the more Roma dominate the less Romani is used – seems contradictory. On the background of functionality and the need for symbolism this contradiction dissolves.

Gadže include Romani into the web presentations of organisations they dominate or participate in for various but highly interlinked reasons: Romani is often used to express support for the self-organisation and emancipation process of the Roma that, to a large extent, aims for socio-cultural equality which is most obviously symbolised by language. Furthermore, the use of Romani insinuates sympathy for the political movement and commitment to assist the Roma in changing their sociocultural and sociopolitical situation. However, the use of Romani might also be interpreted as an act of legitimisation for the involvement of *Gadže* into the on-going emancipation process of the Roma. Consequently the online use of Romani by *Gadže* might be interpreted as the linguistic aspect of political correctness. As for Roma, the symbolic aspect of Romani use on international websites seems to be of minor or even no importance. The functional aspect dominates and, consequently, the internet presentation of international NGOs dominated by Roma and the information provided via the web is almost exclusively in English.

⁹⁰ See <<http://www.coe.int/t/dg4/education/roma>> and <<http://www.coe.int/t/dg3/RomaTravellers>>; for the latter texts in Romani see http://www.coe.int/t/dg3/romatravellers/archive/documentation/otherlanguages/default_en.asp; accessed 2012-02-25.

2.2. Online Plurilingualism

As indicated by mentioning the online video language course on the website of the Romanian NGO *Romathan*, Romani gains communicative importance in oral language use and when the web is used interactively. This is *inter alia* also demonstrated by the use of Romani on and via the web pages of the Swedish *Radio Romano – Nevimata thaj aktualitetura pe romani chib*, ‘Roma radio – news and topicalities in the Romani language’. The web pages of *Radio Romano* are embedded into the website of *Sveriges Radio/Radio Sweden* <<http://sverigesradio.se>> and are accessible via the menu item *Språk* ‘language’ in the general menu of the website. The specific menu for navigation on the web pages of *Radio Romano* is provided in Romani. Written information on the pages is presented in Romani and to some extent also in Swedish. The archived broadcasts offered on the website are primarily in Romani with Swedish and other languages, mainly English, only used by and with interviewees with no competence in Romani. In contrast to the websites described so far, which primarily target the speakers of dominant languages for political purposes, these web pages above all aim to inform plurilingual individuals with competences in Romani, Swedish and also English. Instead of an additive multilingualism with symbolic function, language use in cases like *Radio Romano* has to be described as integrative plurilingualism with a primarily communicative function.

The communicative function also prevails in numerous Roma chat rooms which are characterised by interactive language use and, as it is almost the rule in chat rooms used by speakers of dominated, non-standardised languages, by orate style as well as spontaneous writing.⁹¹ As indicated by the texts on the welcome page of <http://www.zigeuner.de/01_roma-chat.htm> in German and Romani, *Herzlich Willkommen im Roma-Chat!/T’aven saste taj bachtale*, ‘A warm welcome to Roma Chat/Be healthy and happy’, the chat rooms hosted by the website are plurilingual. Actual language use in online chats of ethnic groups usually comprises all languages of the repertoire of the respective speech community. Language mixing or rather linguistic hybridisation in the form of spontaneous loans and communicative code switching are as common as changes in the primary language by situational switching.⁹² If the primary language in a chat is a dominant language, a national language or an international *lingua franca*, Romani is often restricted to symbolic functions – see the welcome formula presented above – of identity flagging, and the expression of solidarity and affiliation. In

⁹¹ Whereas the dichotomy between oral and written refers to the form of communication which is defined by the respective acoustic and visual channels, the dichotomy between orate and literate refers to style. Both orate and literate style occur in written as well as in oral form: orate style in written form means spoken language written, literate style in oral form means written language spoken, etc.

⁹² As most chat rooms require registration to guarantee privacy, no specific examples are presented in this paper.

Romani-dominated chats elements of all other languages in the repertoire of the respective speech community are prone to be inserted into utterances in Romani which fully corresponds with natural communicative behaviour and defines the interaction as plurilingual in the true sense of the word.

This kind of communicative behaviour is reflected on the website of *Radio-Romani-Mahala – Kosova-Mitrovica-Ibar*: <http://www.beepworld.de/members97/romani-mahala/index.htm>. This Germany-based website with online radio broadcasts and the possibility to chat during the radio broadcasts has been established by members of the diaspora community originating from the *Romani Mahala*, ‘Roma quarter’ of *Kosova Mitrovica*.⁹³

*Mitrovica [...] hosted one of the largest Mahalas in Kosovo before June of 1999. Between 6000–7000 Roma lived there. The primary southern Mahala on the banks of the Ibar River is now a gutted ruin. With 650 destroyed homes, it is the largest still-destroyed residential area in Kosovo.*⁹⁴

The welcome lines of *Radio Romani Mahala* present phrases in Serbian, Albanian, Turkish and Romani – *Dobro nam došli dragi gosti/Mir sē vini/Selam Alekim/Šukar alen, lače avilen* – thus reflecting the linguistic reality in the former homeland: Romani as the language of everyday life and in-group marker, Turkish as the former dominant and culturally highly valued language, Albanian as the dominant language of the region and the immediate neighbourhood and Serbian as the dominant language of the adjacent neighbourhood and of the former Yugoslavia. In addition to this, the navigation bar of the introductory page also uses the new dominant language of the country the Roma who run *Radio Romani Mahalla* live in: German as well as English, the *lingua franca* of the internet. The following excerpt from the navigation bar demonstrates this repertoire extension:

- Video Igre Gratis⁹⁵ ‘Free video games’ Serbian
- Mitrovica Mahala News English
- Fotoalbum ‘Photo album’ German
- Romane Gila Gratis ‘Free Romani’ Romani
- Ramazan Iftar Vakt.... ‘(Evening) Meal after sunset in Ramadan’ Turkish

The chat is frequented by young people living in different European countries and beyond. Many of them only know about life in the *Romani Mahala* of *Kosova*

⁹³ Accessed 2012-02-26.

⁹⁴ Accessed 2012-02-26. Source: <http://www.balkanproject.org/roma/mitrovica.shtml>.

⁹⁵ *The lexeme gratis, ‘for free’, is an element of German as well as of colloquial Serbian and of Romani.*

Mitrovica through reports and stories of their parents and grandparents. They were not yet born or small children when they left their countries of origin. Nevertheless, the Romani dialect of *Kosova Mitrovica* is the primary means of communication in the chat and is used in the plurilingual mode described above: orate style, spontaneous writing, and linguistic hybridisation by making use of all common languages in the repertoires of the interlocutors.⁹⁶

2.3. Summary

The difference between additive multilingualism and integrative plurilingualism outlined in chapter 1 of this paper has become obvious from the examples discussed in sections 2.1 and 2.2:

- In additive multilingual settings Romani only functions as an auxiliary language. It is presented in a literate written form with predominant symbolic functions. The communicative aspect of websites that provide translations into Romani is mainly covered by the respective dominant languages, the national language of the country where a Roma NGO is based and active, or English on the web pages of supranational organisations and NGOs with international aims.
- In integrative plurilingual settings Romani usage often equals that of other languages. It is used in an orate style and communicative functions prevail. As neither Romani speakers nor the internet are monolingual in Romani, other languages are used in domains they cover and/or are embedded into Romani in interactive online communication. Compared to the more or less static mode of Romani usage in additive online multilingualism, the mode of interaction in integrative plurilingual settings is dynamic and, consequently, highly transient.

The differences described so far are in no way aspects of simple two-way dichotomies but features marking the two extremes of a gradual field between ethnic symbolism and communicative functionalism against the background of the sociolinguistic situation of Romani as expressed by the arrows in the following summary table:

⁹⁶ I am highly indebted to Daniele Viktor Leggio for information about Radio Romani Mahala. He is currently writing a PhD dissertation on the linguistic and sociolinguistic aspects of the chat on this website at the University of Manchester. He has recently published a description of the Romani dialect upon which this online interaction is based: Leggio, D.V. 2011. *The dialect of the Mitrovica Roma*. *Romani Studies* 5, Vol. 21, No. 1: 57-113.

FUNCTION of Romani	STYLE of Romani use	TYPE of Romani use	MODE of interaction
<i>symbolic</i> ↑↓ <i>communicative</i>	<i>literate</i> ↑↓ <i>orate</i>	<i>additive</i> ↑↓ <i>integrative</i>	<i>static</i> ↑↓ <i>dynamic</i>

3. Conclusion

This limited but to some extent representative cross-section of Romani usage on the internet reflects the repertoire of Romani speech communities as well as the functions of Romani in relation to those of the other languages used in web presentations of Roma NGOs, international organisations and public as well as private Romani online media. As the sociolinguistic situation of Romani is by no means unique and asymmetrical relationships between languages are common and natural, the case of Romani allows for generalisations of language use against the background of linguistic diversity on the internet.

The use of a language on the internet is more or less based on its functional scope. International English as the global *lingua franca*, naturally, is the dominant language of the internet. Transnational languages such as Portuguese, Russian, and Swahili cover their areas of usage in the same way as national, regional, or minority languages do. On the basis of a functional approach each domain is characterised by dominant and dominated languages, a situation that reflects communicative reality or rather functionalities, as well as the status and prestige of languages. However, status and function are not a dialectic pair. Status prevails over functionality and not the other way around. Therefore, web pages translated into dominated languages which lack the necessary functionalities only contribute to symbolic additive multilingualism which can be easily transformed into statistics. However, such data are not meaningful at all. Only dynamic, transient and consequently unmeasurable or rather uncountable language use which is based on communicative integrative plurilingualism reflects the reality of linguistic diversity on the internet.

Fikret BABAYEV

*Head, Cultural Policy Department,
Ministry of Culture and Tourism of the Azerbaijan Republic
(Baku, Azerbaijan Republic)*

Azerbaijani State Policy Toward the Development of Intercultural Dialogue and the Preservation of Cultural Diversity

The dialogue between cultures and civilizations is among the most topical and popular themes in the world. Many countries arrange roundtables, conferences and forums on this dialogue. Every such meeting certainly deserves support and encouragement – at least, because it is in itself a dialogue.

Regrettably, all these events have only token influence on the actual situation. Any news agency's information appearing on the TV or online proves that point. Prejudice, intolerance, ethnic purges, war and genocide reign in this world. Conflicts and confrontations are snowballing.

However, ever more political, religious and community activists, researchers and people-in-the-street join discussions on the intercultural dialogue with every passing year to advance hypotheses, scenarios and initiatives. Thus, certain scholars are sure that the available contradictions between civilizations inevitably lead to antagonisms and clashes, the opportunities for a dialogue are dwindling, and the multicultural community has no future. There is another opinion – that all nations of the world should emulate the Western model to prevent confrontation. Doubtless, Western ideas of freedom and democracy, high living standards and rapid development are praiseworthy. Despite all that, we see quite well that these patterns cannot be accepted by the entire world – mainly due to the current amount and distribution of resources, as well as to geopolitical, historical and cultural factors.

The intercultural dialogue is no longer a problem of cultural studies alone. It has crossed their limits to penetrate big politics. Globalization is underway. Some are enthusiastic about it while others consider it a disaster. Be that as it may, we are all aware of sweeping changes in our life. Probably, each of us should see that we are not only our countries' citizens: we are also members of a family named "humanity". Globalization has made us all dwellers of a vast united space in which dialogues between persons, states and nations should base on ethics.

The political and economic laws that rule the present-day world are deplorably far from the principles of justice, and the idea of global ethics might appear utopian. However, let us look back at the past, and we will see that many breakthroughs of the current civilization started with ideas and actions that were in their time also regarded as utopian.

All over the history of Azerbaijan, it was populated by dozens of tribes and ethnic and religious communities. They contacted each other for millennia to promote information exchanges between cultures and civilizations. The indigenous population consisted of Turkic, Caucasian and Persian ethnic entities. After Azerbaijan was incorporated into the Russian Empire, and later the USSR, it was flooded by thousands of people of diverse nationalities and ethnic backgrounds. They were mildly assimilated, just as the indigenous population with its psychological flexibility. Azerbaijanis have always been able to put up with aliens just the way they were. They never tried to forcefully adapt newcomers to their own customs and mentality.

Shortly after the establishment of the Azerbaijani Democratic Republic, the ideas of equality and ethnic diversity were in the focus of government attention. The Declaration of Independence of Azerbaijan, made public on May 28, 1918, said: “The Azerbaijani Democratic Republic guarantees civil and political rights within its boundaries to all citizens, irrespective of ethnicity, religion, social status and gender” (Article 4), whereas Article 5 vouched “vast possibilities for free development” to all ethnic entities in the republic. There were schools with tuition in many ethnic languages all over the country. The ethnic press, cultural centres, theatres and educational institutions prospered.

In 1920-1991, Azerbaijan was part of the Soviet Union, and so not an international legal entity. Its independence was again proclaimed in 1991. The sovereign Republic of Azerbaijan made its first intergovernmental agreements on the protection of ethnic minorities’ rights with the other former Soviet constituent republics because millions of people found themselves outside their ethnic boundaries in the Soviet era.

The new Constitution of Azerbaijan, endorsed in 1995, proclaimed equal rights of all its peoples irrespective of language, religion, race and cultural traditions. Article 44 says: “1. Every citizen shall have the right to retain his/her ethnic identity. 2. No one shall be forced to change his/her ethnic identity.” This statement fully complies with the Framework Convention for the Protection of National Minorities, whose Article 5, Clause 2, says: “Without prejudice to measures taken in pursuance of their general integration policy, the Parties shall refrain from policies or practices aimed

at assimilation of persons belonging to national minorities against their will and shall protect these persons from any action aimed at such assimilation.”

Article 45, Clause 2, of the Constitution of Azerbaijan confirms: “1. Everyone shall have the right to use his/her native language. 2. No one shall be deprived to the right to use his/her native language,” while Article 21 emphasizes that “the Republic of Azerbaijan shall guarantee the free use and development of other languages spoken by the population.” This statement finds practical confirmation by the presence of more than twenty schools with Russian as the only language of tuition, and approximately 300 schools with several languages of tuition. President Ilkham Aliev said: “Azerbaijan treats ethnic Russians traditionally well. Not a single Russian-language school has been closed. On the contrary, we provide conditions for the best possible Russian language studies.”

There are schools with tuition in Modern Hebrew in Baku and Red Borough, Kuba District. Jewish history and traditions are on the curricula there. A Jewish educational centre opened in a gala in Baku’s Hatai District on October 4, 2010. The decisive role in its construction belonged to the Heidar Aliev Foundation.

Several schools teach the fundamentals of the Ukrainian, Tatar, Lezgian and other languages. The Baku Slavic University, one of the best-respected universities in Azerbaijan, is known for Slav language studies and linguistic research, while the Baku State University has a faculty of Modern Hebrew.

The Samed Vurgun State Russian-Language Theatre successfully works in Baku and frequently hosts guest performances by stage companies from Russia and other CIS countries. The Lezgian- and Georgian-language state theatres also thrive. There are press outlets and cultural and educational centres working in many languages spoken in Azerbaijan.

Further harmonization of interethnic relations and prevention of encroachments on ethnic minorities’ rights are an essential part of Azerbaijani government policies.

The Ministry of Culture and Tourism is the principal government agency responsible for the implementation of programmes and policy of the promotion of intercultural dialogue at the national and local levels. It closely cooperates with the Ministry of Education, the Ministry of Youth Affairs and Sport, the Ministry of Foreign Affairs, the Ministry of Justice, the State Committee for Work with Religious Organizations, the State Committee for Family, Women and Children

Affairs, the National Academy of Sciences, the Republican Copyright Agency, and other involved organizations.

The Ministry of Culture and Tourism is implementing practical programmes to preserve and develop the cultural values of ethnic minorities and groups in Azerbaijan. These programmes envisage cooperation with ethnic minorities' cultural centres and communities, their stage companies' guest performances in and outside Azerbaijan, and ethnographic and art exhibitions. The ministry promotes the creation and performance of art works developing ethnic customs and traditions, and provides folk costumes, musical instruments and other equipment for amateur performing companies.

The Ministry of Culture and Tourism goes on implementing the Cultural Diversity in Azerbaijan project, aiming to provide information about ethnic and cultural diversity and promote the public awareness of the necessity to develop and encourage the intercultural dialogue. Ethnic minorities' cultural associations are attracted to take part in the project.

The "Azerbaijan, My Native Land" festival of ethnic minorities' arts was held in December 2006 on the UNESCO Cultural Diversity programme and the Cultural Diversity in Azerbaijan project. Organized by the Ministry of Culture and Tourism, it included concerts by 800 performers representing over forty companies based in Baku and 14 districts of Azerbaijan. Practically all ethnic entities in the republic were involved in the festival, which won enthusiastic public response.

The presentations of the *Atlas of Traditional Azerbaijani Music* and the *Album of Azerbaijani Ethnic Music*, both prepared under the supervision of Professor Tariel Mamedov, were timed to the festival, just as a republican conference on The Cultural Heritage of Ethnic Minorities and Contemporaneity, which gathered top-notch researchers and cultural activists. A major exhibition of ethnic minorities' traditions and material culture was held at the Museum Centre exhibition hall.

A second "Azerbaijan, My Native Land" festival followed in October 2008, and a third in June 2011. They included research conferences, photo shows, and routine and gala concerts that involved several thousand participants from almost all ethnic minorities and groups resident in all parts of Azerbaijan.

With its efficient government policy of promoting languages and cultures, Azerbaijan is steadily turning into a universally recognized exporter of the experience of international support of cultural diversity, and Baku has

developed into a transnational centre of intercultural dialogue. It was not for nothing that the city was proclaimed Capital of Islamic Culture for 2009.

Azerbaijan is the world's only country to be member of the Council of Europe and ISESCO at once. The Ministry of Culture and Tourism has launched the Baku Process to promote the mutual rapprochement of great cultures and civilizations. The city hosted many ambitious international events on its honourable mission. They opened with the conference of Ministers responsible for Culture of the Council of Europe on "Intercultural dialogue as a basis for peace and sustainable development in Europe and its neighbouring regions" (December 2-3, 2008), at which a number of Muslim nations were represented. In October 2009, Baku was the venue of the 6th Islamic Conference of Culture Ministers, to which representatives of the Council of Europe and several European countries' Ministers of Culture were invited, reciprocally.

The World Forum on Intercultural Dialogue, which Baku hosted on April 7-9, 2011, honourably contributed to practical dialogue. It was supported by UNESCO, the UN Alliance of Civilizations, the Council of Europe, its North-South Centre, ISESCO, and Euronews.

Ethnic, social and religious instability is sweeping the world. Former Yugoslavia was the scene of sanguinary interreligious clashes in the 1980s-90s. They were accompanied by ethnic purges that took thousands of innocent lives. Tribal clashes in Africa last for years. Sudan has split on religious grounds. Conflict has been smouldering for centuries between the Muslim majority and the Coptic minority in Egypt. Lebanon is also split on the religious principle. Tensions are building up in the prosperous Bahrain. Even Europe, long famous for tolerance, has recognized that multiculturalism is a failure.

Azerbaijan, with over twenty ethnic entities, has found its own solution of the nationalities issue. Despite a problem-laden international situation, with a lasting conflict in Karabakh and with 20% of the country's territory occupied, the country's leaders laid the basis of a subtle and reasonable nationalities policy with equal terms for all ethnic and religious groups. According to the latest population census, there are more than 178,000 Lezghins in Azerbaijan, 176,000 ethnic Russians, plus Avars, Tats, Tatars, Kurds, Ingilois, Georgians, Jews, Talyshes, and others. The Framework Convention for the Protection of National Minorities, which Azerbaijan ratified even in 1995, proclaims "effective protection of national minorities and of the rights and freedoms of persons belonging to those minorities, within the rule of law, respecting the territorial integrity and national sovereignty of states", which is of great importance in this problem-laden time of ours. It also guarantees "to persons belonging to

national minorities the right of equality before the law and of equal protection of the law". Deserving special attention in this respect is ethnic minorities' active participation in public services of all levels and representative bodies of Azerbaijan. There are Lezghins, Avars, Russians, Jews, Kurds and people of other ethnic backgrounds among the members of the Milli Mejlis.

Azerbaijan has a unique historical and ethnological situation: the offspring of its aboriginal population of many centuries ago still retain their languages and traditions – suffice to mention the Udin, whose majority live in the village of Nij in the Gabala District. With small exceptions, they profess Christianity and speak their native language in everyday situations. The same concerns the Shakhdag ethnic group, which consists of the Khanalyg, Budug and Kryz.

The Azerbaijani nation has always taken pride in the multitude of peoples, religions and cultures in its land. The Azerbaijani think wisely that every citizen of their country is a member of one family whose duty it is to promote the development of his or her salient features and protect unique ethnic and cultural qualities. Our nation's humanism finds material proof in the small ethnic minorities surviving through centuries. History has given the world an inimitable example of peaceful coexistence of the many ethnic and religious groups. For many centuries, the Azerbaijani people have instilled patriotism in the nation without recurring to coercive assimilation, and developed mutual respect and intolerance of nationalism in whatever form. Multiculturalism is my country's present-day reality, which shows that time and patience can make an alloy of many peoples and religions.

Nikolai PAVLOV

*Administrator, Sakha Wikipedia
(Yakutsk, Russian Federation)*

Role of Modern ICT in Keeping Sakha Language Afloat

Continuity: General Considerations

The Sakha language, better known as Yakut, branched off from its parent language, proto-Turkic, about two thousand years ago. Despite Mongol and Tungus-Manchurian influences, it has preserved its basic structure to this day, and, according to experts, is, perhaps, the closest of all modern languages to Proto-Turkic.

The current processes of globalization and cultural homogenization dramatically undermine the world's ethnic diversity while also stepping up contact between languages, which may enrich or debilitate one another as a result.

This leads to many smaller communities losing touch with their roots and finding themselves on the sidelines; the situation further aggravates social inequality, causing public discontent. There is a risk of subversive forces hijacking that protest sentiment, especially strong among the young.

This is why it is so important to sustain linguistic and cultural continuity from one generation to the next. Few people would deny that language is an important (perhaps even the most important) element of any ethnic culture. Every particular language has its own ways to encode basic notions of life, and human language at large is a means of identity building as much as a product of social development.

How a Minority Language Could Be Preserved?

One method consists in linguistic standardization (including of neologism-building rules) and the creation of databases from documents and text corpora.

Another is about the preservation of living small languages as means of everyday communication. This can be achieved by expanding their functional scope and, more importantly, by ensuring their generational continuity.

Instilling Love for Mother Tongue in Younger Generation

Let us begin by outlining children's personality and worldview development milestones. For the sake of convenience, we will rely on the commonly accepted ages & stages model.

Preschool education marks the most important developmental stage, one that paves the way for worldview formation and for the encoding of notions and actions with the help of words and phrases. There are two major types of preschool learning. One is about home-based education, usually provided by parents and grandparents, with the older family members serving in this case as the main transmitter of culture for the young.

The other type is applicable to cases when both parents work outside home. A prominent role in children's early education is then played by nursery schools and kindergartens. Starting from two or three years of age, kids whose parents work fulltime have to spend most of the day with preschool teachers and caregivers.

The next stage is primary and secondary education. In that period, children tend to spend less time with their teachers, so it would be only natural to expect a resurgence in parental influence. In reality, though, the role of parents shrinks further because, on the one hand, child-teacher contacts gain in intensity, with the teacher becoming a major transmitter of knowledge, while on the other hand, the child's inner circle expands and so, too, does the range of his/her information sources.

Then comes tertiary education (vocational schools, colleges, and universities). It is widely believed that by the time they pass on to this phase, most teenagers will have almost reached adult maturity. So the role of tertiary education establishments in personality molding is, more often than not, negligible.

A community's status quo cannot be preserved unless a community member gives birth to and raises a child. For simple demographic reproduction, one woman (family) should give birth to 2.4 children (given infertility incidence, infant mortality, etc.). In smaller language communities these days, there are quite a few parents who believe that teaching the native language to their children is no longer relevant and may even be harmful, so they choose to concentrate on a national language instead. Because of this trend, a native speaker population in small language communities will take an increasingly high birth rate to reproduce itself.

Looking at the modern-day Yakutia, we can see that preschool education here is rarely provided in the language of its indigenous inhabitants. The same is true of any other non-Russian ethnic region or community in Russia. So children who spend much of the day in a nursery-school environment are just bound to grow up without knowing their mother tongue.

The second developmental phase (through high school) further consolidates the alienation from their mother tongue of children from smaller language communities (in most of the country's non-Russian ethnic regions, schools providing education in native languages are few and far between or virtually non-existent). This process is exacerbated by a lack of native-language content

in a schoolchild's information environment (advertising, entertainment, games, extracurricular activities, mass media, the Internet, etc.).

And at the tertiary stage, exposure to one's mother tongue in small language communities is often reduced to a bare minimum.

Possible Conclusions

As the basic speech and language skills are developed at the preschool stage, when children are still too dependent to make their own decisions, it is necessary to convince parents of the importance of passing mother-tongue knowledge on to their offspring – for him or her to grow into a wholesome personality and a community member competitive enough and resistant to harmful influences.

It is necessary to provide every kind of support for native language learning at preschool education establishments, opening new, modern nursery schools and kindergartens.

The right to study one's native language and in one's native language must also be ensured in primary and secondary education. There is a need to create an appropriate native language environment and to provide ample supply of various type of content in the language.

Yakut Language: Current Situation

To better understand what should be done to expand the Yakut language's functional scope, let us try to overview its present-day status. Yakut is better off than many other minority languages spoken in Russia.

It is not that the Yakut have numerical superiority over fellow non-Russian communities. But unlike other indigenous groups, many of the Sakha people not only speak their language, but write in it, as well.

This creates a demand for Yakut-speaking journalists and authors while also ensuring wide enough circulation for Yakut-language print media, large audiences for electronic media, and relatively high book publishing numbers. Basically, the Yakut language owes its high profile in modern-day community life to the prominence given to it on the school curriculum.

But despite the broad circulation of Yakut periodicals (the newspaper *Kyym*, according to statistics from Russia's National Circulation Service, in 2009 set a record among the non-Russian newspapers, outstripping periodicals in Tatar, Bashkir and other minority languages) and the high number of young authors writing in Yakut (at last year's 17th young writing talent conference, the auditorium was packed beyond capacity), we should not let ourselves become complacent. If, making the most of today's level of ICT development, it is not put on a par with functionally more advanced languages, but continues

to service a limited number of areas of human activity (such as agriculture, arts, and household practices), young community members will be reluctant to study it themselves and to teach it to their children. And the Yakut language's relevance may then rapidly dwindle as a result.

This is why it is so important for the Republic of Sakha's indigenous peoples – the Yakut, the Evenki, the Even, and the Yukaghir – to have their languages represented in cyberspace and to expand their use in areas like science and technology, as well as in the activity of various public institutions.

Role of ICT

In the modern-day world, *accessibility of information and communications technology* is the key to language use expansion.

Being accessible to Yakutia's urban population (at least in comparison with neighboring regions), ICT remains out of reach for most of the republic's rural communities, who constitute the main source of traditional culture and language knowledge.

It is highly unlikely, though, that the local Internet and mobile phone providers will lower the prices of their services any time soon, nor that they will invest in the purchase of new hardware. In Yakutia's still narrow ICT market, such a behavior would defy the sheer logic of economics. So alternative solutions have to be found. One way of going about this problem would be to foster competition (for instance, by providing preferential treatment for innovative wireless communications technology) or to adopt regional government programmes aimed at reducing the digital gap.

Providing users with quality information. Users (especially young ones) should be provided with a terminology base (in ICT and related areas) and relevant software, along with informative and entertaining content in their native language. A child/parent/teacher searching for some fairy tale or flash movie in the language should have a choice of several sites offering the required content.

There is also a need for various computer-based teaching aids and training/simulation tools, as well as for publicly accessible databases on various subject areas, such as law.

In addition, indigenous inhabitants should be provided with opportunities for online communication in their native language with fellow counterparts. It seems like a good idea to create forums for specific expert communities, such as finance specialists, accountants, etc. And then again, there should be an excess supply of appropriate content and communication tools.

Terminology. There is a need for the republic's Cabinet and parliament to adopt regulatory acts that would sustain the homogeneity of the Yakut terminology base.

Language standards are usually developed either by the native speaker community (this is a historical, usage-based method) or by professional linguists, who elaborate norms based on language evolution laws and then offer them to the speaker community for approval.

In our globalization era, no community can afford to spend decades or centuries developing usage norms. Unless the Yakut language is seriously upgraded within the next few years and relevant terms are developed for it with the help of linguists and sectoral specialists, future generations of native speakers stand little chance of passing the “relay baton” of their native culture over to posterity. As a result, the language will lose its clout even in those areas where its current positions are quite strong.

Yakut-language software. It is necessary to provide software programmers worldwide with an opportunity to include the Yakut language in their developments. To make this happen, we should team up with the Unicode consortium and have the Yakut language and its basic notions introduced into the Unicode Common Locale Data Repository (CLDR) and repositories with similar functions.

Relying on that base, a producer of computer/mobile phone software could independently develop Yakut language support programmes. This would give to the developer a competitive edge without entailing any additional costs. There will then be no need to remake already existing software programmes, creating a special Yakut-language shell for each.

It is necessary to rapidly design a spellchecker for the Yakut language, as well as a programme for computer-assisted translation of text corpora from Yakut into major world languages and vice versa. The North-Eastern Federal University is already working toward that goal. The early results include a thesis on computer modeling of the Yakut language, completed and defended in 2005 by V. Migalkin.

Providing Yakut-language content. This problem could be resolved by engaging public organizations and businesses. One example of a public organization’s involvement is Wikimedia. Sakha Wikipedia has been living quite a vibrant life of late, with volunteers contributing encyclopedic articles on various subject matter. A second, newly-launched project – Sakha Wikisource – is an e-library with texts in the Yakut language.

The Sakhainternet company offers an example of business involvement. Its website Ykt.ru – the republic’s most visited these days – switched over to the Unicode in 2010, making its resources usable for communication in the Sakha language.

Public organizations and independent businesses launching such projects should be provided with every kind of support and encouragement.

Another example of mutually beneficial partnership is one between Wikimedia's Amir Aharoni and the Hebrew University of Jerusalem. Thanks to Aharoni's efforts, a professor of the Economics Department now asks his students to write Wikipedia entries instead of conventional course papers. The students get additional grade points while Hebrew-speaking Internet users get 20 new articles on economics.

And in Kazakhstan, the Wikipedia writer known as Asel has sent a letter to the republic's culture minister, asking him to tend to Wikipedia's Kazakh-language segment. Fueled by the government's directive support, Kazakh Wikipedia is facing a real boom now.

Technical problems with scripts. Script availability is a rather specific issue, to be considered separately from the more general problem of ICT use. The adoption by the international community of the Unicode system has facilitated software support for most of the world's languages, including Yakut. It has received individual codes for the five letters unique to its alphabet. There is no need now to readjust other coding systems to user needs. But there are problems yet to be solved. One is the reluctance of script producers to include the atypical letters of Russia's minority language alphabets into their products. Which is hardly surprising, though, given the lack of economic incentives.

At long last, Russia's government has started placing orders for the development of integrated scripts supporting the Unicode and the alphabets of all of the federation's constituent ethnicities. The Russian company Paratype has developed the first digital scripts (PT Sans and PT Serif) and presented them to the public. There are just two such scripts so far, and their number is unlikely to increase any time soon. In the Windows operating system, some of the default commercial scripts support letters by many alphabets, but not all, and not to an equal degree. Such scripts are also preinstalled in software programmes produced by Apple and ones based on the UNIX/LINUX platform.

Conclusion

Every culture is valuable and unique, and the disappearance of any single one will be an unrecoverable loss for the entire human race. Multilingualism's cultural significance has global proportions.

Linguistic pluralism is necessary and desirable because traditions are a strong factor of stability in human society. A person fully exercising his or her rights in the native language and culture feels a full-fledged community member and, hence, can counter modern challenges more effectively. In other words, the preservation of languages and attention to the problems of ethnic communities are not just cultural issues, but also safeguards for a secure future of humanity.

Victoria GOLBTSEVA

*Chief Engineer, North-East Complex Research Institute,
Russian Academy of Sciences (Far East Branch)
(Anadyr, Russian Federation)*

Role of Modern Technology in Preserving Linguistic and Cultural Diversity of Indigenous Chukchi Peninsula Communities

According to a 2003 census, 14,600 of the Chukchi Peninsula's inhabitants are representatives of small indigenous peoples, such as the Chukchi, Koryaks, Eskimo, Chuvans, Evens, Evenki, and the Yukaghirs. Younger generations in such communities are facing the challenge of preserving the cultural heritage and linguistic diversity of their ancestors.

With every passing day, the local indigenous communities have fewer elders who could pass on their impeccable command of the language, their expertise in reindeer breeding, sealife hunting, medicinal herbs' application, and in traditional dressmaking, as well as their knowledge of folklore, rituals, traditional dance and festivals and, importantly, that of the history of their ethnicity and its most outstanding representatives.

What should be done to preserve all that wealth and to hand it down to posterity?

This problem wasn't addressed in earnest until the late 20th century, when the Russian ethnographers V. Leontyev and I. Krupnik embarked on a quest for possible solutions. They collected stories from indigenous sea hunters of the Bering Strait and had them published.

Then the Programme for the Preservation of Cultural and Linguistic Diversity of the Eskimo and Chukchi came along, as an unforced result of collaborative efforts between these two renowned experts and the authors of this report. With no official status, this programme is being run as kind of a public movement, which engages relatives and fellow community members of V. Golbtseva and V. Nuvano and representatives of other ethnic communities. Also involved are research fellows from the following institutions:

- Arctic Studies Center, National Museum of Natural History (Smithsonian Institution, Washington, D.C.);
- Regional natural and ethnic park Beringia (Providence, Chukchi Peninsula, Russia);

-
- Chukchi Peninsula Eskimo society Yupik (Providence, Chukchi Peninsula, Russia);
 - Ethnic & cultural public movement Chychetkin Vettav (Anadyr, Chukchi Peninsula, Russia).

The programme's guidelines include identification, studying, preservation, and distribution (paper, digital and word-of-mouth) of content related to:

- fundamentals of the Eskimo and the Chukchi languages, including professional vocabulary, traditional knowledge, and related industrial & cultural expertise, customs and rituals;
- cultural landscape structures of sea hunters and reindeer breeders, including the system of traditional settlement and economic practices overland and at sea, as well as ethnic toponyms;
- history of the various communities, biographies of their most prominent members; anthropologic categorization of the Asiatic Eskimo and Chukchi.

The Programme's four sections are described below.

Section 1. Working with the elders

Recording narratives and theme interviews, along with videotaping the narrators and their traditional economic activities. Materials prepared in the course of that work are preserved in digital form and subsequently published (see L. Bogoslovskaya, I. Slugin, I. Zagrebin, I Krupnik: Introduction to Sea Hunting, Heritage Institute Publishers, Anadyr, 2007).

Of particular value are the original drafts and drawings of traditional hunting gear, primarily harpoon tips and baydaras (open skin boats) and the descriptions of their production techniques, illustrated with video footage. These drafts are nowadays used in building baydaras in many communities of the Providence and the Chukchi regions.

The first ever dictionary of related terminology in the Chaplino Eskimo language has been released, with Russian equivalents. This is the result of a collaborative effort by Andrei Ankalin, a sea hunter from the Sireniki settlement, and the ship designer Sergei Bogoslovsky, who has built – to a traditional design – a model of an Eskimo baydara-anyapik, a small boat employed in ice hunting for sealife; it has a wooden skeleton secured by leather straps, and is upholstered with walrus skin.



*Going to sea on an Eskimo baydara built in the Sireniki community.
© Photo by N.Perov.*

Inspectors of the Beringia park, A. Apalyu (of the Yanrakynnot village) and A. Borovik (from Novoye Chaplino) have photographed with digital cameras the entire cycle of hunting for grey and Greenland whales and of on-shore game processing.

The preservation of traditional whale and pinniped hunting as well as of food culture traditions is a key challenge facing the Chukchi Peninsula's indigenous population. Specialists in medical anthropology have proved that the increased incidence in iron-deficiency anemia and myopia, the high mortality rate among cancer patients, and the appearance of various kinds of "urban" disease, hitherto unknown locally, are all a direct consequence of the shift of the local population's diets toward industrial foods. Only traditional staples, including the flesh and fat of winnipedes and whales, can ensure good health and longevity for the modern and future generations of the Eskimo and coastal Chukchi, researchers say.

Characteristics have been identified for 17 local winds and their seasonal roles in sea hunting off the Chukchi Peninsula's Uelen coast have been described (see V. Golbtseva, V. Vetrovoi: *The Wind System of Uelen: Natural Schedule and Traditional Knowledge//Chukchi Peninsula: Sustainable Use of Natural Resources and Environmental Security*, Magadan, 2008, Issue 12, pages 145-

161). Sea hunting off Uelen is considered to be more challenging than elsewhere. Here, it is done not just on the shore ice belt (known locally as “tukvek”), but also on drifting ice (“rochgygelyk”), which involves the dangerous crossing of the rift between the two (that gap is referred to in the local language as “tylyagyrgyn”).

Golbtseva and her fellow villagers have also compiled an electronic dictionary of Uelen sea hunters’ sea ice terminology. It contains more than 200 entries, with words and phrases in the Chukchi language and their Russian equivalents. Each type of ice is provided with a detailed description from the temporary perspective, specifying the time of its formation and its seasonal changes. In the dictionary, one can also find rules of conduct on drifting ice and the sea hunters’ mutual aid code.

Similar e-dictionaries have also been compiled for ice landscapes of Sireniki’s Eskimo inhabitants (A. Nutautie, N. Rodionova; 51 entries), Chaplino’s Eskimo (L. Ainana, L. Bogoslovskaya; 34 entries); the Naukan Eskimo (B. Alpyrgin, E. Dobriyeva; 52 terms); the Chukchi of the Yanrakynnot village (A. Apalyu, E. Gottyrgyn, L. Kutylin, A. Kosyak; 49 entries), and a dictionary of terms denoting various types of snow in the language of the Chukchi inhabitants of the Vayegi village (N. Nuvano, V. Nuvano; 50 entries).

Also, a dictionary of the Chaplino Eskimo language for traditional practices in the use of natural resources (compiled by L. Ainana, L. Bogoslovskaya, P. Aleinikov) has been prepared for print. It contains 1,500 terms and expressions, along with Russian and English equivalents and with the scientific Latin names of animal and plant species.

Section 2. Cultural landscapes of sea hunters and reindeer breeders (settlement layout, nomadic routes, toponymics, etc.)

This is one of the programme’s most challenging parts, requiring large-scale expeditions. While quite a few publications are already available on sea hunters’ cultural landscapes, a comprehensive analysis of cultural landscapes of Chukchi reindeer breeders has only just begun. V. Nuvano is one of the few explorers concerned with the role of landscape in the reindeer breeding culture of Vayegi’s Chukchi inhabitants.

Section 3. Cooperating with public organizations in efforts to preserve languages, folklore and traditional dance.

In the early 1930s, a writing system was created for most of the Chukchi Peninsula’s indigenous groups, paving the way for the rise of print media in community languages. The year 1930 saw the launch of a local newspaper, which would carry one of its articles in the Chukchi and the Eskimo

languages. Another periodical, *Sovetken Chukotka*, then came to the scene, followed by *Murgin Nutenut*. With the advancement of radio and television, shows in community languages began to be broadcast all over the Chukchi Peninsula, even to reindeer breeders' camps.

The local *Krainiy Sever* newspaper these days carries only Chukchi-language translations of Russian news stories, unfortunately. The paper's Website, ks87.ru, runs a special Chukchi-language column, called *Vettav*.

Local radio stations continue broadcasting programmes in the Chukchi, the Eskimo, and the Evenki languages, but the duration of broadcasts has decreased dramatically since the mid-20th century.

The shrinking community-language scope and the decreasing number of indigenous inhabitants speaking their native languages have prompted northern communities to create several public organizations that would work toward preserving their native languages and traditional cultures. The most efficient of these is arguably the *Chychetkin Vettav* (Native Word). Members of this association hold regular meetings, where only the Chukchi language is spoken, and invite over some of the prominent indigenous narrators and dancers, photographing and filming their performances.

Section 4. Photographing and videotaping the physical appearance and traditional activities of indigenous community members. Identification, restoration and digitized conservation of the family photo archives of sea hunters and reindeer breeder dynasties.

This section started off with efforts to study the photographic archives collected by V. Bogoraz and A. Forshtein, with the originals held at the American Museum of Natural History (New York City, U.S.) and the Peter the Great Anthropology and Ethnography Museum, of the Russian Academy of Sciences, St. Petersburg. These are rich collections of photographic images, giving a good idea of previous generations' looks and lifestyles.

Some of the images have been attributed and, with the museums' consent, they were printed in the 2007 manual "Introduction to Sealife Hunting," in the "Tropoyu Bogoraza" collection (2008), and in a book on sled dog breeding, "Nadezhda – a Race Along the Earth's Edge" (2011).

V. Nuvano's family archives have been digitized and retouched. They contain one-off images of the Chukchi jailed after the 1940 uprising and sent to gulags, as well as pictures of the widows and orphaned children of the reindeer breeders persecuted then.



The widows of the Gemavye brothers, executed following the 1940 revolt along with their father: Etteut, grandmother to V. Nuvano, left; Tynatvaal, right.

The village of Vayegi

L. Pautov, late 1970s © Photo archives of V. Nuvano

Conclusion

Despite the significant achievements already made within the framework of the Programme for the Preservation of Cultural and Linguistic Diversity of the Eskimo and Chukchi, there is a need for more efforts toward preserving and studying the dialects of the Chukchi language and its unique gender differences (the male and female phonetic patterns).

No clue has yet been found to explain the phenomenon of scarce cross-borrowings between the Eskimo and the coastal Chukchi languages. The famous sealife hunting culture of the Bering Strait exists in a bilingual environment, yet the two languages, with dissimilar origins and vocabulary, develop each their own separate way.

Funds are now being raised to finance the showcasing of the programme's major achievements in a book series and in a collection of discs. The primary aim, though, is to instill the indigenous cultural and linguistic heritage preservation awareness among younger generations, encouraging them to join in the effort.

Vyacheslav SHADRIN

Research Associate, Institute for Humanitarian Studies
and Ethnic Minorities of the North,
Russian Academy of Sciences (Siberian Branch)
(Yakutsk, Russian Federation)

Yukaghir Language and Culture in Cyberspace

Linguistic Situation

Scholars believe that the Yukaghir people once formed a separate family of related tribes. Russians who explored North-Eastern Siberia at the beginning of the 17th century found survivors of twelve indigenous tribes in the area between the Lena and the Anadyr Rivers. Those tribes all had a common ethnonym, “Odul,” along with various local names such as *alai*, *koime*, and *anaul*.

According to period yasak tribute bills, there were about 6,000 Odul (or Yukaghir) speakers in the early 17th century. Their numbers subsequently dropped, though. Overtime, the remnants of surviving tribes became dissolved in other ethnic communities, with only two groups maintaining their ancestral language and culture – the Yukaghirs of the lower and the upper Kolyma River (some specialists regard them as separate ethnicities speaking related languages).

As of 2002, the total number of ethnic Yukaghirs was 1,509, including 1,097 Yakutia-based. Of these, only 604 (or 40%) spoke the Yukaghir language.

It should be noted here that the Yukaghir language status differs from community to community. Many of the Yukaghir inhabitants based in the tundra have now switched over to either Yakut or Russian while those living in forestland tend to speak Russian more. The tundra-based Yukaghir communities, meanwhile, often opt for multilingualism, using languages such as Yukaghir, Even, Chukchi, Yakut and Russian.

Fig. 1. Findings of a 2002 census

Region	Ethnic Yukaghir	Yukaghir people speaking their native language	Yukaghir people speaking Russian	Yukaghir people speaking other languages	No data available
Republic of Sakha (Yakutia)	1097	310	1046	677	2
Chukchi Autonomy	185	82	152	86	17
Magadan Region	79	13	79	8	-

Importantly, these statistics do not reflect the actual level of language proficiency. The overwhelming majority of the respondents who claimed they speak Yukaghir are not fluent speakers really, with their “command” often based on the knowledge of a limited number of words and phrases.

As of early 2011, fewer than thirty natives were fluent in the Yukaghir language, including five in its forestland dialect.

Main Action Lines for the Revival and Preservation of the Yukaghir Language

The problem of language preservation is closely related to a community’s socio-economic and cultural development. In efforts to preserve the Yukaghir language, particular attention should be paid to the following tasks:

1. Encouraging the use of the native language in a household environment;
2. Teaching the language at pre-school institutions;
3. Shaping language awareness in schoolchildren;
4. Teaching the language to college students;
5. (Re)creating an environment that could stimulate language use;
6. Promoting research on the language;
7. Recording and archiving native speech samples;
8. Promoting folklore and literature in the language;
9. Enhancing the presence of the language and culture in cyberspace.

Only a systemic approach will enable us to cope with the challenge of preserving the Yukaghir language.

Yukaghir Language: Domains of Use

Any natural language is used in the following four domains:

1. Personal;
2. Professional;
3. Public;
4. Educational.

For the Yukaghir language, these domains were closely linked together as long as the traditional lifestyle was maintained. But they split apart in the 20th century, prompting a sharp decrease in the language use. Yukaghir left the public domain in the 1950s and the ’80 saw its use in the personal domain reduced

to a minimum. The subsequent commercialization of traditional Yukaghir occupations also pushed the language away from the professional domain.

The educational domain, nowadays associated primarily with the Internet, is the only one where the presence of the Yukaghir language has been growing.

Initially, Yukaghir speech samples would have been published exclusively in scholarly publications intended for a narrow circle of specialists. The first Yukaghir-language edition came out in 1965. A Yakut publishing house then released verses by Uluru Ado, as part of a book of poems and short stories, *Yukaghirskiye Kostry* (Yukaghir Fires).

Yakutia's government set spelling rules for Yukaghir in an April 28, 1983 decree, but this did not spur the development of book publishing in the language, with as few as seven book titles released through the mid-1990s.

In the '90s, the Yukaghir language appeared in the media, first in the Gevan broadcaster's radio and television shows and, later on, in the Ilken newspaper. Unfortunately, its presence has not expanded much since then. The Ilken, with only six editions remaining per year, is the only surviving periodical in an indigenous language of the Russian North. But its Yukaghir-language content amounts to an annual 2 pages overall. And Gevan's broadcasts on Sakha Channel have now been reduced to two television shows, 30 minutes long each, and four 15-minute radio shows. Total broadcasts in the Even, Evenki and the Yukaghir languages come to just two hours per week.

Yukaghir Language in Cyberspace

The Yukaghir language first appeared on the Internet in 2004 when E. Maslova opened a website featuring Yukaghir-language content. Later on, A. Nikolayeva launched another website, *Kolyma Yukaghir Documents Online*: <http://ling.unikonstanz.de/pages/home/nikolaeva/documentation/index.html>. Ms Nikolayeva then proceeded to post a number of Yukaghir-language materials at Stanford University's webpage.

The National Library of the Republic of Sakha (Yakutia) has contributed its bit by opening a Knigakan section on its home page (<http://nlib.sakha.ru/elib/collections.php?cd=60>), with a Yukaghirika subsection, where library resources in the Yukaghir language are posted online. In 2011, the scholar A. Nemirovsky opened his page in the Live Journal (<http://nyoro-2011.livejournal.com/604.html>), to post ample content on the Yukaghir language and culture.

The Yukaghir schools in the settlements Nelemnoye and Andryushkino have also launched their websites.

Expanding the Outreach of Yukaghir Language and Culture

The compilation of digital teaching aids in the Yukaghir language began in 2003, when 7 CDs featuring examples of Yukaghir folklore were released by the Yakutsk Music & Folklore Museum as part of an international symposium on the folklore of Paleo-Asian peoples.

In 2007, Yakutia's Education Ministry, in association with research fellows from the Russian Academy of Sciences' Institute of Small Indigenous Peoples of the North, implemented a UNESCO project aimed at preserving the Yukaghir language by enhancing the potential of ethnic schools in the republic's Yukaghir communities. As a result of this project, 14 language textbooks were released on CDs and DVDs. The ministry built on that positive experience in subsequent years, compiling five more CDs with Yukaghir speech recordings in 2009-2010.

In 2010, the Yukaghir language and culture specialists L. Zhukova, P. Prokopieva, A. Prokopieva, E. Atlasova, and V. Shadrin, working in collaboration with the North-Eastern Federal University's Center of New Information Technology, developed a programme for the promotion of the Yukaghir language and culture on digital carriers and in cyberspace. Designed for the period through 2014, the programme should become a real breakthrough in efforts to preserve the Yukaghir language, including through the publication of relevant textbooks and teaching aids. That project received wide support from the university's top and became part of its own development programme. It provides, specifically, for the release of several dozen new-generation textbooks, the launch of the links webpage www.arctic-megapedia.ru, and the organization of expeditions into Yukaghir communities to collect ethnographic, linguistic and folklore material. The effort has already yielded its first results – five Yukaghir-language textbooks on CDs and a language & culture links webpage, prepared with contributions from native speakers and tradition carriers. There is every reason to expect that the North-Eastern Federal University will soon become a major centre for the preservation and advancement of the Yukaghir language and culture.

FINAL DOCUMENT

Yakutsk Call for Action: A Roadmap Towards the World Summit on Multilingualism (2017)

PREAMBLE

The 2nd International Conference on Linguistic and Cultural Diversity in Cyberspace was held in Yakutsk, Russian Federation, from July 12 to July 14, 2011 in the framework of the Russian chairmanship in the UNESCO Information for All Programme. The event was organized by the Ministry of Culture of the Russian Federation, Commission of the Russian Federation for UNESCO, UNESCO, Russian Committee of the UNESCO Information For All Programme, the North-Eastern Federal University, the Interregional Library Cooperation Centre, MAAYA World Network for Linguistic Diversity, and Union Latine.

The Conference which held several working sessions and plenary meetings gathered over 100 participants from more than 30 countries and all continents. It was attended by leaders and experts of intergovernmental and international nongovernmental organizations, government bodies, agencies and institutions of culture, education, research, information and communications, representatives of private sector, civil society and the mass media.

The conference tackled the three major themes dealing with 1) Instruments for language preservation and promotion in cyberspace, 2) Institutes to promote linguistic and cultural diversity, and 3) Creating positive environment for the promotion of linguistic and cultural diversity in cyberspace.

A roadmap towards the World Summit on Multilingualism has been proposed and approved by the participants.

I. Achievements since the Lena Resolution (Yakutsk, Russian Federation, 2008)

Lena resolution was adopted at the 1st conference on Linguistic and Cultural Diversity held in 2008 proclaimed by the UN the International Year of Languages. It introduced the idea of holding the World Summit on Multilingualism and supported a series of initiatives. Since then, several activities have taken place including the Bamako International Forum on Multilingualism 2009 (BIFM), the creation of the Centre to Advance Multilingualism in Cyberspace under the North-Eastern Federal University

(Russian Federation), the development of DILINET project on indicators of linguistic diversity in cyberspace by MAAYA and partners, and the preparation of the 2nd International Conference on Linguistic and Cultural Diversity in Cyberspace (July 2011, Yakutsk).

II. The roadmap to the World Summit on Multilingualism (2017)

We, the participants and organisers of the 2nd International Conference on Linguistic and Cultural Diversity in Cyberspace, inspired by the Lena Resolution and Bamako commitment on universal multilingualism, invite all interested individuals and organisations concerned in linguistic and cultural diversity support and promotion to participate in the preparation of the World Summit on Multilingualism. We therefore propose the following roadmap.

- 2012: Experts meetings and consultations in various regions of the world
- 2013: Ministerial conference on Multilingualism
- 2013: Proposing a draft resolution on the World Summit on Multilingualism to the UNESCO General Conference. Possible creation of an international commission to prepare a Report on multilingualism
- 2014: Possible adoption of a draft resolution on the World Summit on Multilingualism by the UN General Assembly based on the Report
- 2015-2017: Thematic meetings and regional conferences to prepare for the Summit
- 2017: World Summit on Multilingualism

All the stakeholders (governments, international organisations, private sector, and civil society including the academia and research community) are encouraged to commit themselves in preserving and promoting linguistic and cultural diversity worldwide and invited to contribute to the success of the World Summit on Multilingualism which could lead to the adoption of an international Charter on the governance of linguistic diversity in the world.

LENA RESOLUTION

Final Document of the 1st International Conference “Linguistic and Cultural Diversity in Cyberspace” (Yakutsk, Russian Federation, 2-4 July, 2008)

The International Conference on Linguistic and Cultural Diversity in Cyberspace was held under UNESCO auspices in Yakutsk, Russian Federation, from July 2 to July 4, 2008. This conference represents the Russian Federation’s contribution to the United Nations International Year of Languages.

The Conference was organized by the Government of the Republic of Sakha (Yakutia), the Russian Federal Agency for Culture and Cinematography, the UNESCO Moscow Office, the Commission of the Russian Federation for UNESCO, the Russian Committee of the UNESCO Information for All Programme, the Interregional Library Cooperation Centre, the National Library of the Republic of Sakha (Yakutia), and the Committee of the Republic of Sakha (Yakutia) for UNESCO.

The Conference gathered more than 100 representatives from 15 countries and all continents—leaders and staff members of intergovernmental and international nongovernmental organizations, government bodies, agencies and institutions of culture, education, research, information and communications, business, civil society and the mass media, as well as political and community leaders and leading experts.

The Conference discussed the political, cultural, social, ethical and technological aspects of research and development of linguistic and cultural diversity in cyberspace, and examined issues related to the use of information and communication technologies to safeguard all languages.

The participants held three plenary meetings and five thematic sections: Political, Ethical and Legal Aspects of the Development of Multilingualism in Cyberspace; The Internet and Other Media; The Role of Libraries in the Preservation of Linguistic Diversity; Languages and Education; and The Preservation and Development of Cultural Diversity.

In concluding its work, the Conference agreed on the following resolution:

The participants of the International Conference on Linguistic and Cultural Diversity in Cyberspace (Yakutsk, Russian Federation, July 2-4, 2008),

1. Recalling the Universal Declaration on Cultural Diversity, adopted at the 31st session of the UNESCO General Conference in 2001, which says that

“cultural diversity as a source of exchange, innovation and creativity is just as indispensable for humanity as biological diversity for Nature, and is a treasure shared by the entire human race”,

2. Recalling further the Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace, adopted at the 32nd session of the UNESCO General Conference in 2003,
3. Taking into account the key documents of the World Summit on the Information Society (WSIS)–Declaration of Principles and Plan of Action (first phase in Geneva, 2003) and Tunis Commitment and Tunis Agenda for the Information Society (second phase in Tunis, 2005)—which all emphasize the importance of the preservation of cultural and linguistic diversity and suggest a set of measures necessary to achieve this goal,
4. Emphasizing the need to continue preserving multilingualism because languages are stores of a rich and vast amount of human heritage and knowledge, as well as necessary instruments for social life, the expression and dissemination of social and cultural traditions, self-identification and preservation of human dignity of their speakers,
5. Recognizing that the dissemination of multilingual information on the history, languages and culture of different nations contributes to the promotion of tolerance and mutual understanding and guarantees a peaceful sustainable development of the contemporary civilization,
6. Taking into account that the current absence of certain languages in cyberspace contributes to the widening of the already existing digital information gap,
7. Considering that the already existing information and communication technologies offer new opportunities for the preservation of linguistic and cultural heritage for equal and universal access to knowledge,
8. Bearing in mind the messages of the Director General of UNESCO at the occasions of the International Year of Languages in 2008, the World Day for Cultural Diversity, for Dialogue and Development on May 21, and the International Mother Language Day on February 21,
9. Thanking the Director General of UNESCO for his message at the opening of the current International Conference on Linguistic and Cultural Diversity in Cyberspace,
10. Thanking also the Government and the people of the Republic of Sakha (Yakutia) for their hearty hospitality, which ensured the success of the Conference,

Consider it necessary:

- a. to use every opportunity to promote the preservation of linguistic and cultural diversity in cyberspace and elsewhere;
- b. to stress the importance of implementing the Recommendations of
 - the World Summit on the Information Society (Geneva, 2003, and Tunis, 2005),
 - the International Conference “UNESCO between Two Phases of the World Summit on the Information Society” (St. Petersburg, 2005), in particular Action Line C8 (Cultural Diversity and Cultural Identity, Linguistic Diversity and Local Content), and
 - the thematic meeting on “Multilingualism for Cultural Diversity and Participation of All in Cyberspace”, organized by UNESCO in preparation for the second phase of the World Summit on the Information Society (Bamako, Mali, 2005);
- c. to invite the UNESCO Intergovernmental Council of the Information for All Programme to examine the possibility of creating an expert group on multilingualism in cyberspace;
- d. to suggest to UNESCO and MAAYA (World Network for Linguistic Diversity) to convene a World Summit on Linguistic Diversity in Cyberspace;
- e. to support the ACALAN (African Academy of Languages) and MAAYA initiative for a World Forum on Multilingualism.

Suggest in particular,

- f. to invite all interested organizations and private persons to contribute to the UNESCO website “Languages Matter!”, dedicated to the International Year of Languages, and to the discussion forum “Communication and Information”, organized by UNESCO as a follow-up to the decisions of the World Summit on the Information Society in Action Line C8 (Cultural Diversity and Identity, Linguistic Diversity and Local Content), by placing information on their own projects in support of linguistic diversity;
- g. to invite the national committees of the UNESCO Information for All Programme to elaborate indigenous and minority linguistic programmes promoting linguistic and cultural diversity in cyberspace;
- h. to call the attention of governmental and other organizations regulating television, radio and mobile telephones environments to the necessity of considering the opportunity of using the latest technologies for broadcasting in minority languages during the media switch to digital broadcasting;

-
- i. to continue searching for fair and balanced solutions between the international copyright and intellectual property legislation and regulatory frameworks for the formation of digital libraries responsible for cultural and linguistic diversity;
 - j. to increase support for libraries, museums and archives in the access to and preservation of multilingualism and dissemination of information representing cultural and linguistic diversity;
 - k. to promote the elaboration and dissemination of relevant free/libre open source software, fonts and other technical means necessary for the use of languages in cyberspace;
 - l. to promote the elaboration and dissemination of a unified and universally accessible set of fonts and diacritic marks supporting the languages of all cultural groups in the Russian Federation, along with Unicode enhancements as necessary;
 - m. to continue promoting the activities aimed at documenting, preserving and developing languages, mainly minority languages, through the latest information technologies;
 - n. to continue promoting the elaboration of multilingual thesauri and dictionaries and the inclusion of minority languages into internationally accepted thesauri on cultural heritage such as UNESCO, HEREIN, ICONCLASS, Getty, Garnier etc.;
 - o. to continue the study of ethical aspects of the preservation of linguistic and cultural diversity in cyberspace, and their consideration in the activities aimed at preserving and promoting such diversity;
 - p. to step up activities for the introduction, promotion and enhancement of minority languages in the educational systems and the expansion of their application and use (in the government sector, the sphere of research and art, on television, radio, etc.);
 - q. to suggest to the participating countries the monitoring of language diversity on the Internet;
 - r. to extend international cooperation with all interested organizations including the International Federation of Library Associations (IFLA), the Conference of Directors of National Libraries (CDNL) and the International Indigenous Librarians Forum (IILF);
 - s. to invite the Government of the Republic of Sakha (Yakutia) to consider establishing a competence centre for the study and promotion of multilingualism in cyberspace and for sharing its experiences with the world at large.

Linguistic and Cultural Diversity in Cyberspace

Proceedings of the 2nd International Conference

(Yakutsk, Russian Federation, 12-14 July, 2011)

Compilers: *Evgeny Kuzmin and Anastasia Parshakova*

Translators: *Tatiana Butkova, Ekaterina Komarova and Anastasia Parshakova*

Computer-aided makeup: *Igor Goryunov*

Executive: *Sergey Bakeikin*

Publisher:

Interregional Library Cooperation Centre

2A, bldg. 1, 1st Basmanny Lane, Moscow, 105066

Tel./Fax: +7(499) 267 33 34, +7(499) 263 26 61

www.mcbs.ru
