



**Culture Sector
Communication and Information Sector
Education Sector**

Background document

“Assessing Language Situation and Planning in Relation to the Internet”

**Prepared by Mr Marcel Diki-Kidiri,
LLACAN (CNRS / INALCO), France**

for the expert group meeting
“Towards UNESCO guidelines on Language Policies:
a Tool for Language Assessment and Planning”

30 June – 1 June 2011

Table of Content

Introduction	3
I. Assessment.....	3
1. The different ways to communicate on Internet.....	3
2. The applications.....	5
3. The contents	7
4. Applied samples	11
5. An easy-to-use tool for decision makers.....	12
II. Planning.....	13
1. Language and information policy.....	13
2. The internationalization of domain names.....	13
3. Some key steps to secure a language position in the cyberspace.....	14
III. Recommendations	16
References.....	17

Introduction

This document is a working paper for a group of experts convened by UNESCO to discuss about how Member States should assess their languages situation and then make adequate planning in relation to the Internet.

The study starts by suggesting an easy-to-use tool for a quick assessment of the vitality of any language on the cyberspace in a way that shortcomings can be easily seen. It is then straight forward to decide what to do to improve the situation. Then it addresses the issue of language planning for cyberspace with all its preliminary conditions.

Finally, a set of recommendations are given as a summary of what should be considered as the essential of language planning, at a governmental level, when it comes to promote all the languages spoken in a given country on the cyberspace, taking into account that local situations vary from a country to another and may affect decisions.

I. Assessment

It is a complex operation to measure the vitality of a language on the Internet. We shall attempt to address it as accurately as possible starting from its three main components:

- The different modes of communication while using the Internet
- The applications
- The contents

1. The different ways to communicate on Internet

a) Website

If we want to know whether a given language is present on the cyberspace, the first evidence we would look for is a website entirely or partially written in that language. A website is somehow easy to create and to manage, and it can be visited from anywhere in the world, as long as a connection can be technically established. Therefore, a website is the first tool to be built when there is a need to promote a language on the Internet. Yet, it doesn't mean that there is a community of speakers committed to support this language on the Internet, as a website can quite well be nothing more than a poster on the cyberspace wall.

b) E-mail

E-mail is the most popular and the most used mode of communication on Internet. Because of it, telegram and telex have become obsolete and even fax and postal mails are much less used. But, since mails are protected by confidentiality rights, it is more difficult to search their contents and check precisely in which language messages are written. Furthermore, it needs more technical skill to set and manage a mail server than what is needed to create a website. That is why, in many situations, people use publicly available (or commercial) e-mail services, such as Google, Yahoo, Hotmail, rather than creating an e-mail server in their own language, except when non Latin scripts are needed. Yet, what

is relevant here is that people could send mails to each other in their own language, regardless of the mail services they use. The act of sending mails in a given language implies, in itself, that there is a language community whose members send messages to one another in the same shared language.

c) Fora, social nets, blogs, collaborative environments

Fora, social nets and blogs are intended for a large public with a high level of interactivity. They need to be set in special environments that can manage huge data bases and a very high flow of multimedia data. Such environments lead to a quite different way to conceive websites. To make simple, let us say that websites in the 1990s are mostly made of pages which can be modified only by the webmaster. Visitors have a very limited interactivity with the website. They can sign a guestbook, take part in a newsgroup or a forum, or send a mail to the webmaster, and that's all! With the social networks, the website becomes highly interactive. As (s)he signs up, each final user creates her/his profile, has a blog, an e-mail, a SMS, a gallery of pictures and so forth. (S)He may parameter her/his space, write and post articles multimedia data, and participate to a collaborative work. He can choose her/his friends, add or remove whatever (s)he wants, just like what a webmaster would do. The architecture of websites which are built for social networks is radically different from the old way of conceiving websites. That is why it is called Web 2.0, as it gives a much larger place to the final user. Many of the social networks such as Facebook (which is only one of the many social networks out there) have nearly or more than 100 Millions of users.

As to collaborative environments, such as SPIP, TikiWiki, or Wikipedia, they provide platforms where people can come and contribute to a collective project or initiate new ones. Wikipedia, the most outstanding of them, contributes substantially in spreading local knowledge in multiple languages and is the natural successor or at least evolution of dictionaries and encyclopedias.

So, if a language is used in the forum, the blogs, the social networks and the collaborative environments, this language is definitely very active on the Internet.

d) Interactive real time messages

By "Interactive real time messages" we mean short messages systems which are interactive and in real time, instant messaging (e.g. MS-Messenger, Skype, Google or Yahoo messaging) IP telephone, video conference systems, webcam videos, as they all allow communication on the Internet. Although it is very widespread, this mode of communication is mostly used by professionals or for some particular contents. Its specific tools are usually¹ not available in less spread languages. Therefore, only languages which are pretty well established in the cyberspace are actually used here.

e) Online streaming

In the realm of the ICT, of which the Internet is part, online streaming, including sound and/or video, but also encompassing files sharing and information feeds, is one of the latest technologies to become more and more popular. Information feeds, such as RSS

¹ We talk here about instant messaging tools, not about other kinds of tools such as open access softwares (e.g. Firefox) which are regularly localized and translated in many languages.

feeds which are now even more popular because of mobile phones, since they allow the continuous reception of information and this maybe in many languages. As usual, only very widespread languages that benefit from powerful production companies are used here long before all the others.

In order to evaluate the vitality of a language, each mode of communication is given 1 point. The more the language is used in different modes of communication, the higher its vitality is ranked. Its vitality level matches the number of the modes of communication it is used in.

Table 1. Communication mode on Internet

Level of Vitality	Index level or number of modes of communication used.
Dynamic	5
Steady/active	4
Receptive	3
Adaptable	2
Minimal	1
Inactive	0

2. The applications

These are programs that allow the user of Internet to work or to communicate. This human activity sustains the language in which it is carried out. Each application has a user interface that allows the user to communicate with the computer in order to monitor his work. All warning, help or notice messages the application may have to send to the user are part of the user interface as well as the control or browsing buttons which show their functions clearly written on them. It is assumed that the language in which these informations are given is familiar to the user. The more the language is spread among a large community of speakers, the more the localization of applications in this language becomes a necessity and easier to justify. Given equal features, most people would prefer softwares that let them work in their native language.

Hence, the bare availability of an application in a language is already a first significant indication of that language's vitality. Then, of course, it is often necessary to check if there are also such help documents as user guides, reference manuals, handbooks, written in the same language than the user interface. One should also try if possible to look at statistics for downloads to certify the actual use of these localized versions of the applications, etc. Taking into consideration all the above checking, the number of applications available and really in use in a given language is a reliable index of the vitality of that language in the cyberspace, and more generally, in the digitalized world.

a) Browser and search engine

While a browser searches your own computer disc locally, a search engine is the emblematic applications of the access to the Internet since it is the vessel that let someone travel from website to another across the cyberspace as you look for informations or data to download. It is a basic application designed for a wide public. Therefore, it is often the first software to be localized in a new language. As search engine, Google is very active in collecting information about languages but many others are

available (Chinese engines, etc). Yahoo, Google and Bing have also localized pages for search engines, and allow searches filtered by language. Yet, some experts like Daniel Pimienta (Funredes / Maaya) find their performance somehow disappointing when used intensively.

Hence, the initiative taken by Funredes to develop the DILINET project, a much more reliable search engine to observe and report on languages used by Internet users in all of their online activities.

b) E-mail software

Though e-mailing is the most popular and intensively practiced activity on the Internet, an e-mail software is a specialized tool since it is designed essentially to send, receive and manage mails and address books. Even when it is part of a browser (webmail) an e-mail application doesn't do more than what it is able to do alone. As mentioned earlier, many people are quite satisfied with public or commercial e-mail services available in a language other than theirs, even if they use them to send e-mails in their own language. That is the very reason why, the availability of a localized e-mail server in a given language, specially a less spread one, is peculiar enough to be significant of the vitality of that language. The huge popularity of e-mailing makes localized e-mail applications, when they are effectively in use, a good indicator of vitality.

e) Office Suite

An Office Suite (e.g. MS-Office, OpenOffice.org, etc.) is a package of office applications that include mostly a word processor, a spreadsheet, a presentation application, a drawing tool, and a database system. These applications are very frequently used to elaborate data for publication on websites. They are tools for producing contents, even though they are basic applications for office works. Among all the applications included in an office suite, word processor is the first one would like to localize in a new language, because it makes it possible to write texts in that language. Yet, most office suites are built with a strongly integrated architecture, so that it is much easier to localize the whole set rather than each of its parts separately. The work is, nevertheless, more complex and requires a higher qualification than for simply using the software. Therefore, provided, as said above, the availability of user and technical manuals as well as spell checking dictionaries etc are also taken into account, the availability of an office suite in a given language facilitates considerably the production of digitalized data in that language, hence indicating its vitality.

d) Website creators

Softwares for creating websites are specialized tools for professionals or skilled amateurs even though anyone can easily purchase them anywhere. Users of these softwares usually speak one of the most widespread languages in the world. So, there is very little need of localizing them in less spread languages. Therefore, the more a language is lively on the cyberspace, the more website creators in that language may be desired. Hence, the availability of this kind of applications in a language indicates clearly that this language has a rather high level of vitality in the cyberspace.

e) Advanced tools for producing contents

This category includes softwares to process and manage pictures, videos, movies, softwares to create games, tutorials, etc. All of them are highly specialized applications designed for highly skilled users. There is usually very little need, if any, to localize them in less spread languages. Therefore, if any of them is found in a language, this definitely means that language is very dynamic in the cyberspace.

In order to measure the vitality of a language with the applications, each of the applications' types is given a value of 1 point. The more there are different types of applications available in a language, the higher the language's vitality is ranked. So, its vitality level matches the total number of types of applications available in it.

Table 2. Applications

Level of Vitality	Index level or number of types of applications available.
Dynamic	5
Steady/active	4
Receptive	3
Adaptable	2
Minimal	1
Inactive	0

3. The contents

To analyze contents is the best and straightforward way to measure and evaluate the vitality of a language in the cyberspace. Unfortunately, it is virtually impossible to have a global view of the whole content of all the data exchanged or shown on the Internet, mainly because these data are growing so fast that they have already saturated the actual architecture of the Internet, while the most powerful browsers available nowadays are just unable to retrieve reliable data which would allow to compare the relative situations of the used languages.

Nevertheless, it is possible to collect partial but sufficient informations to have a relatively good idea of the situation of a language in the cyberspace. An evidence of this is given, on one hand, by the *Language Observatory Project* led by Yoshiaki Mikami (University of Technology of Nagaoka, Japan) and, on the other hand, by the works of Funredes Association (San-Domingo) led by Daniel Pimienta. Both of them are members of MAAVA, a worldwide network for language diversity.

Beside the hopefully transient technical barriers, another one is rather ethical. How to identify the language used in an e-mail without accessing anonymously to this mail since the contents of mails are normally confidential? The preservation of the mails' confidentiality is one of the main reasons why the observation of the whole Internet content is out of reach, at least for the time being.

Therefore we suggest to take advantage of all the contents the public can access to, namely contents published on websites (e.g. web pages, blogs, forum, etc.) as a basis on which the vitality of a language could be determined. In fact, two indexes linked to the content are both necessary and sufficient to do that : the first one reports on the presence

of a language as part of the content itself, whereas the other one, tells how much the language is used as a vehicle for different kind of contents.

3.1. Presence of the language in the content

This index is based on the five following clues:

a) Information on the language

Information on a language, with or without samples of texts in that language, is the least required for taking that language as being present in a website's content. Such information is usually given in another language with higher vitality that is used as working language. For instance, a website in French that gives informations about the Arawak language, and eventually some texts in Arawak. This indication witnesses that the Arawak language is present on that website, but doesn't mean that it has a dynamic situation in the cyberspace.

b) Grammar and literature

When beginners start learning a new language, its grammar, texts or even literature, comes usually fairly early into the picture. Although these data are relatively easy to prepare or to compile in order to be published on a website, they are substantial enough to give the language a greater presence online than data of the former clue.

c) Language lessons

To publish a full set of online lessons in a language (not only an advertisement strip intended to collect subscriptions for offline lessons) is a way to make that language available to more people and to widen the number of its potential learners. It is then an indication which clearly points to the vitality of the language rather than to its decay.

d) Dictionaries, encyclopedia (Wikipedia) terminologies

Much more than the wordlist which may be published on a website right away on the first day, the elaboration of a substantial dictionary, an encyclopedia or lexicons of specialized terminologies requires a lot of work and many contributors. So, when such works are found in a language on the Internet, they definitely witness the effective vitality of that language.

e) Language of the user interfaces

On a website, quite a lot of elements are messages intended for the user such as buttons of redirection or browsing, captions of research areas, contents of help pop-up bubbles, etc. The website authors write these messages in a language which, therefore, becomes the working language of that website.

Hence, it is assumed that this language is familiar enough to everyone who comes and visits the website and may use its services and contents. The language of user interfaces is, therefore, a living language in the cyberspace.

As for the other indexes, each of the clues stands for 1 point. The total of the clues relevant to a given language matches the index level of the presence of that language in website content.

Table 3. Presence of the language in the content

Level of Vitality	Index level or number of relevant clues for language presence in content.
Dynamic	5
Steady/active	4
Receptive	3
Adaptable	2
Minimal	1
Inactive	0

3.2. Language as vehicle of diverse types of contents

It is assumed that a language which supports a very large variety of types of contents is expected to be much more dynamic than the one which carries a limited number of types of contents. The diversity of the types of contents can be very large. We have selected five very common of them to shape the needed 1 point clues. Under each type of content we gather 3 or 4 topics which are significant for that type. Since the diversity of the types of contents is, in itself, the index of the language vitality, a language which is found carrying topics all related to only one type would have an index of 1 point. The index of vitality matches the total number of the types of content supported.

a) Communication

- News, online newspapers, media
- Advertisement, adverts
- Directories, practical informations,

b) Society

- Associations, social life
- Religion and philosophy
- National and international policy
- Public services

c) Learning and playing

- Training, education, literacy
- Games, sports and other entertainment
- Music and video

d) Science

- Sciences of technologies
- Human and social sciences
- Science of life and nature

e) Economy

- Economy, finance
- Firms, employment
- E-Business

Table 4. Language as vehicle of diverse types of contents

Level of Vitality	Index level or number of clues for language as vehicle of content
Dynamic	5
Steady/active	4
Receptive	3
Adaptable	2
Minimal	1
Inactive	0

4. Applied samples

4.1. Evaluation of a language.

The Sango (natively written *Sängö*) language is found on several websites and in e-mails, but the three other modes of communication are not yet used. Therefore, its index for "Internet modes" is 2. There is no software localized in that language. Accordingly, its index for "Applications" is 0. On the websites where Sango is present, the involved data include informations on the language, a substantial dictionary which is frequently updated, a complete basic course in Sango, some literature (poetry, texts from novels) and there is at least one website where Sango is used as one of the two working languages of the interface. So, the Sango index for "Presence in the content" is 5. Contents in Sango deal with topics related to Communication, Society and Learning and playing, though not intensively. This gives us an index of 3 for "Vehicle of Contents".

Table 4. Evaluation of a language: the case of Sango

Level of Vitality	Level	Internet mode	Applications	Presence in C.	Vehicle of C.
Dynamic	5			x	
Steady/active	4				
Receptive	3				x
Adaptable	2	x			
Minimal	1				
Inactive	0		x		

It can be observed that to promote Sango language on the Internet, the priority is to develop or to localize softwares in that language. This will help widening the modes of communication in use and the types of contents addressed.

4.2. Comparing two or more languages

As shown for the Sango language, any language can be evaluated with the above criteria. Languages which are strongly installed and dynamic on the Internet will surely rank at level 5 for all the indexes. As a matter of fact, the majority of the world languages are not strongly installed in the cyberspace and particularly on the Internet. The evaluation grid proposed here allows to evaluate the situation of any of them and to compare different language situations.

In the table 5 hereafter, the language L1 reflects the situation of the Sango language as described above, while L2 is another language which uses 3 modes of communication on the Internet (let us say, website, e-mail and forum), has two localized applications (a browser and an Office Suite), doesn't have a complete course online but is present in all the other types of contents, and finally, vehicles data related to all the types of contents except Economy.

With 3 indexes out of 4 higher, language L2 is noticeably better positioned on the Internet than language L1.

Table 5. Comparing situations of 2 Languages (L1 and L2) on the Internet

Level of Vitality	Level	Internet Mode	Applications	Presence in C.	Vehicle of C.
Dynamic	5			L1	
Steady/active	4			L2	L2
Receptive	3	L2		L1	L1
Adaptable	2	L1	L2	L1	L1
Minimal	1	L1	L1	L1	L1
Inactive	0	L1	L1	L1	L1

If, in the frame of a language policy, it is chosen to promote an equal use of the two languages on the Internet, then what is to be done appears fairly straightforward:

- 1) The priority is to create or localize applications in L1 with the aim of bringing the index up to at least level 2 ;
- 2) To widen the modes of communication on the Internet for L1 in order to reach at least level 3 ;
- 3) To create data in at least one extra type of contents in L1 in order to bring its index up to level 4 ;
- 4) To create a complete set of lessons in L2 in order to bring up to level 5 its index of presence in the content.
- 5) To keep on localizing softwares in the two languages, aiming at least at level 3.
- 6) To aim level 5 for the three indexes that have not yet reached that level at this point, namely: Applications, Internet Modes and Vehicle of Contents.

5. An easy-to-use tool for decision makers

We must bring into attention that we have chosen criteria of evaluation which are both relevant enough to allow an efficient evaluation of any language situation with respect to the Internet, and easy enough to understand and to manage. Nevertheless, these criteria give a static view of the language situation at the moment of the evaluation. In order to get a dynamic perception of the vitality of a given language on the Internet, it is necessary to repeat the static evaluation at regular periods and compare the results. The cyberspace is a world where technologies evolve quickly. A language has to be constantly receptive and sufficiently dynamic to take advantage of all the technological innovations. The criteria of evaluation we have chosen here will undoubtedly be affected by them.

For instance, the list of topics we selected to build up the types of contents for index "Vehicle of Contents" can greatly evolve depending on the practices in fashion. Furthermore, new investigation tools may be invented which can analyze the whole content of the Internet, including e-mails and secret messages (diplomatic, industrial, strategic, etc.) then a quite different method can be built to measure language vitality on the basis of advanced statistics that shows their real use. The complexity of the required parameters should no longer be a problem for a team of experts working in a specialized institution. Yet the grid of evaluation proposed here still can be useful as it can give a first level assessment of a language situation, generally sufficient for making decisions, especially where nothing has ever been done for less spread languages.

II. Planning

1. Language and information policy

The promotion of languages on the Internet at governmental level is usually embedded in a more global language and information policy. The government of Quebec, for instance, has developed the concept of government online, a framework for boosting together e-administration, information society and e-democracy.

All the public services and informations are made available in French to every citizen on the governmental website. French is used systematically as part of the governmental policy also backed by linguistic laws. In many multilingual countries, such as South Africa which has 11 official languages and intend to promote all of them, languages can be thus enhanced by the government through linguistic laws and an intensive use of the Internet. Yet the technology of Internet itself is not quite ready, until recently, to support all official languages, leave alone all the languages of the world! Many languages are written with a non-Latin script and many others use an extended version of the Latin alphabet which is not taken into account by the current Internet governance technology.

Over the last decade, it has gradually become obvious that, for Internet to remain a really global network, it must be able to handle multilingual scripts and alphabets! Hence fore, an important first step has been done by ICANN toward the internationalization of domain names.

2. The internationalization of domain names.

On November 16th, 2009, at a meeting in Seoul, South Korea, the Internet Corporation for Assigned Names and Numbers (ICANN) launched its First Track Process that allows the use of non-Latin characters in domain names from start to finish.

This historic move has given the go-ahead to countries, nations, and territories all over the world to apply for Internet extensions in their languages and in their scripts such as Arabic, Chinese, Cyrillic, Ethiopic, Greek, N'ko, Hindi, Japanese, Korean and Tifinagh, i.e. more than 100,000 characters. The implication of the move is great, as it can open the web to more people around the globe. Now, it should be emphasized that it is up to each country to take advantage of this opportunity to enhance and promote its national and local languages on the Internet and in the cyberspace in general. Up to now, thirty

countries including Serbia (.cpб), Morocco (المغرب, “al-Maghreb”) and Algeria (الجزائر, “al-Jazairi”) have passed the approval procedure.

The internationalization of domain names is undoubtedly a big step, yet it is just the top of the iceberg. It is not only the domain names that have to be internationalized but also, all the necessary tools to take advantage of them, such as search engines, browsers, e-mail servers, office suites, operating systems, and so on, as mentioned here above. However, since many practical considerations are involved, policies are necessary guidelines for good legislation and regulation. Any language policy with respect to the Internet implies, at least, the existence of consistent general communication frameworks that take in the Internet infrastructure.

3. Some key steps to secure a language position in the cyberspace

3.1. Political decisions

The very first political decision to be taken by a government should be to define a clear vision of what the nation should look like in, for instance, twenty years as a fully connected, educated and multilingual nation after political actions have been carried out in fields like language, information, communication and education. Then a global plan with several phases, each of them having its aims well worked out, can be set up to detail the governmental policies in those fields. In many countries, and mostly in many less-developed ones, this represents a very important stir that requires a permanent unflinching commitment of all partners, supported by governmental will. Modern communications are driven forward by rapid changes and innovation, therefore a firm hand on the policy rudder is needed to keep a country on the track of sustainable and equitable development so that all can fairly benefit from the advantages of the ICTs.

3.2. Communication frameworks

A dense framework of wireless communication should be built all over the country in order to cover both rural and urban areas. This framework includes the Internet infrastructure, but of course, it is designed for all information and communication technologies! It can also be based on different technologies such as satellite broadcasting, optic fiber or local wireless connections. It is important to establish frameworks that allow free and fair competition in the market. A market-driven solution would let cable owners constitute monopolies which force the competition to be at the level of services. Alternatively, a community-driven network, based on open access principles, would essentially allow all operators fair access to the broadband networks. As Habumuremyi and Finlay² point it out,

“The disastrous effects of cable monopolies have already been seen in countries like South Africa, which for years has struggled to bring down the cost of Telkom-controlled Internet access. In an open access network prices would be cost-based and all providers subscribed to common interconnection tariffs”. Incentive measures should be taken to help establishing several and competitive service providers, along with guidelines for controlled prices and easy Internet access for all, especially in developing countries

² Habumuremyi E. and Finlay A. “Rwanda’s policy vacuum could mean trouble for broadband” APCNews. <http://www.apc.org/en/news/rwanda-s-policy-vacuum-could-mean-trouble-broadband>

where this activity is rather rare. This entails a commitment to generalize easy access to and intensive use of computers in every region of the country, opening more opportunities for local communities to contribute to the promotion of their native languages on the Internet. In both cases, a national independent regulatory body should be set up to supervise, boost and channel the growth of the Internet usage in the country.

3.3. Language resources and contents

A national language body should be set up with specialized committees for each language, including the lesser-used ones, to initiate researches and conduct every necessary activity which can enhance their use and status. Universities and research centers are the main places where research on the languages and their digitalization may mostly be carried out. Yet, examples from Canada, Spain (Catalunya), South Africa and many other countries show that a national body devoted to the implementation of the governmental vision and language policy is needed to give guidelines and help every partner in his/her contribution so that programs like the following could be set up:

- To take advantage of the Internationalized Domain Name new opportunity to secure domain names in local writing systems and national languages, if the case applies.
- To create and localize softwares and digital resources in the different national languages;
- To build databases such as terminology, literature, digital encyclopedia that cater for the needs of language developers as well as ordinary language users;
- To translate official documents, public service informations, administrative procedures in each national language and spread them online;
- To create online lessons and teaching materials in each of the national language to be used in educational initiatives;
- To carry surveys and measure the evolution of the implementation of the whole language and information policy, and to suggest legislation accordingly.

While the deployment of Internet infrastructure is the bedrock of ICTs, it is worth to remember that what make it an inestimable wealth are its contents, the informations. So any ICT policy should pay a very close attention to an optimal use of the infrastructure by developing as many contents as possible. Multilingual countries have the opportunity to produce much more contents in several languages, provided the producers are trained and empowered enough to do the job.

3.4. Empowering skilled users

In most industrialized countries, it is a truism to say that children are almost born with digital gadgets in hand! The situation is, of course, quite opposite in developing countries even though it is evolving. In any case, no country with an ambition to develop an ICT-based economy can afford to bypass the empowerment of its population in order to produce skilled users of ICTs. Many methods and ways are open to a governmental action on that line. Let us just mention here-after some of the outstanding ones:

- Initiating or enhancing the study of national languages and computer science in school, while taking into consideration the different levels from primary school to university.
- Training writers, translators and interpreters for all the languages of the country;

- Encouraging non-governmental associations to create clubs of computer-helped learners of any subject, but particularly in literacy, language learning and writers training.
- Creating or enhancing a school of translation and interpretation with national languages included.
- Facilitating the acquisition of low-cost computers for students and school children, and for anybody living in rural areas.
- Deploying Internet access centers throughout the country with a cheap cost to facilitate its use by all.

It is obvious that each of these lines of action has ties with large domains such as education, literacy, telecommunication, law, economy and labor, each of which entails a complete governmental policy.

III. Recommendations

The following recommendations are derived from the above study and aim at helping decision makers at governmental level who wish to design policies for assessing and enhancing language situations with regard to Internet.

- 1) In case it is not yet done, the very first issue to address is to define a clear vision and general outlines of a national language and information policy for the country. This document should be particularly well discussed and agreed upon by experts of different ministerial departments and civil sectors involved (e.g. education, information and communication, culture, economy and finance, NGOs and civic society representatives) as it will be the foundation of all subsequent political decisions in this field. It should answer the following question: What the nation should look like in, for instance, twenty years as a fully connected, educated and multilingual nation after political actions have been carried out in fields like language, information, communication and education?
- 2) An assessment of the situation of all the national languages including the lesser-used ones should be done to give the government objective informations to improve decision making in its language policy.
- 3) The next step is to elaborate language laws, if necessary, to give each language a status such as official, national, regional, or language for teaching, which implies a pragmatic definition for each of these qualifications. As definitions may change from a country to another, it is worth suggesting that an official language is expected to be the one in which official documents and laws are written. In many countries what is called a “national” language is the same as an “official” one, yet, where they are different, a national language is defined as an indigenous vehicular language. A regional language could be the one spoken in a large part of a country or even a cross-boundary one. A genuine language policy must clarify what the government intend to do with each of these languages in its vision. For instance, which language(s) must be taught or used for teaching in schools, which ones must be used on trial, and so forth. Laws should be taken to back actions in favor of the different languages including the less-used ones.

- 4) A national body should be set to supervise and regulate all the activities linked to the implementation of the language and information policy. It should be the technical top connecting point of a vertical architecture of stakeholders involved in the implementation of that policy in their particular sectors of activity.
- 5) The deployment of communication frameworks is likely to take several years until each area in the country is completely connected, but it is very important to initiate and sustain this investment as it is the backbone of all the other operations foreseen in the national political vision. So the sooner it is started, the better.
- 6) Many incentive measures have to be taken to develop and encourage service providers and content production. While a suitable legislation organizes the activities of service providers, it is recommended to create a governmental program to digitalize all official documents, laws, decisions, judgments, administrative procedures, public services, and so forth, to put them online.
- 7) For a language to be used on a permanent basis in the cyberspace, it must quickly develop a huge amount of technical terms in many different fields, starting with the Internet terminology itself. A program devoted to the creation of terminology databases for each language to be promoted should be one of the top priorities of governmental concern. In many countries, this is the job of an important special institution directly linked to in the government sphere.
- 8) More contents can be created by individuals through NGOs such as the association of writers, mass media, cultural clubs, religious institutions, etc. An advocacy program should be launched to get their adhesion to the governmental language policy and help it gain popularity amongst the civic society.
- 9) To empower the language users, it is very important to teach the national languages in school. A multilingual education policy should be carefully elaborated which would be based on mother tongue to start with. Basically, any language which is promoted on the Internet should have enough skilled native speakers to keep it alive in the cyberspace by producing continually renewed contents. The only way to raise the competence of the native speakers is to get them read and write fluently in their own language. This can mostly be done in school, or in long life learning literacy training programs for adults.
- 10) All the languages of the country, including the less-used ones, should be studied as object of research in universities and high level skilled specialists trained in order to serve in the different programs of the governmental language and information policy.

References

1. Books and papers

DIKI-KIDIRI Marcel 2007 *Comment assurer la présence d'une langue dans le cyberspace ?*
Edité par la Division de la société de l'information, Secteur de la communication et
de l'information (Editeur : Claudio Menezes) – Paris, UNESCO. Programme
Information pour tous (PIPT) 80 p.

FANTOGNAN Xavier, 2005, « Une note sur les langues africaines sur la Toile mondiale » dans UNESCO : *Mesurer la diversité linguistique sur Internet* ; pp. 106-110.

MIKAMI Yoshiki, ZAKI ABU BAKAR Ahmed, SONLERTLAMVANICH Virach , VIKAS Om, PAVOL Zavorsky, ZAIDI ABDUL ROZAN Mohd, NAGY JÁMOS Göndri, TAKAHASHI Tomoe 2005, « Diversité linguistique sur Internet : une perspective asiatique » dans UNESCO : *Mesurer la diversité linguistique sur Internet* ; pp. 93-105.

PIMIANTA Daniel, 2005, « Diversité linguistique dans le cyberspace : modèles de développement et de mesure » dans UNESCO : *Mesurer la diversité linguistique sur Internet* ; pp. 13-34.

PRADO Daniel 2005. « Le contexte politique et juridique » dans UNESCO : *Mesurer la diversité linguistique sur Internet* ; pp. 35-41.

PAOLLILO John, 2005, « Diversité linguistique sur Internet : examen des biais linguistiques » dans UNESCO : *Mesurer la diversité linguistique sur Internet* ; pp. 42-91

SCIADAS G. *Découvrir la fracture numérique*. Série sur la Connectivité n° 7. Statistiques Canada, Ottawa, 24 pages (ISSN: 1492-7934, ISBN: 0-662-87226-6).

UNESCO, 2005, *Mesurer la diversité linguistique sur Internet*, un ensemble d'articles signés par John Paolillo, Daniel Pimienta, Daniel Prado et autres, révisé et accompagné d'une introduction de l'Institut de statistique de l'UNESCO, Paris, France, (CI-2005/WS/06 CLD 24822)

2. Websites

750 linguistic laws from many countries can be found here :

<http://www.tlfg.ulaval.ca/axl/Langues/lois-linguistiques-index.htm>

HABUMUREMYI E., FINLAY A. "Rwanda's policy vacuum could mean trouble for broadband" APC News. <http://www.apc.org/en/news/rwanda-s-policy-vacuum-could-mean-trouble-broadband>

South Africa and its 11 official languages online :

<http://www.safrica.info/about/democracy/pansalb.htm>

The Quebec government online concept:

<http://www.msg.gouv.qc.ca/gel/index.html>