# IN THE EUROPEAN UNION WITH THE GEORGIAN AND ABKHAZIAN LANGUAGES — AIMS, PROBLEMS, RESULTS, AND RECOMMENDATIONS OF THE COMPLETE TECHNOLOGICAL SUPPORT OF THE GEORGIAN AND ABKHAZIAN LANGUAGES

Konstantine Pkhakadze Merab Chikvinidze Giorgi Chichua Shalva Malidze David Kurtskhalia Constantine Demurchev Nodar Okroshiashvili Beso Mikaberidze

Center for Georgian Language Technology, Georgian Technical University Kostava str. 77, 0160 Tbilisi, Georgia. gllc.ge@gmail.com

#### Abstract

Together with Meta-Net's publications "Europe's Languages in the Digital Age" and "Strategic Research Agenda for Multilingual Europe 2020" this paper is mainly based on the long-term projects "Technological Alphabet of the Georgian Language" and "Plan-Program for the Complete technology support of the Abkhazian Language" of the center for Georgian language technology of the Georgian technical university. Namely, by this center, On May 17, 2019, a report "In the European Union with Georgian and Abkhazian languages – Aims and Problems of Complete Technology Support of Georgian and Abkhazian Languages" was presented in the center for innovation and high technologies of the Georgian national academy of sciences. Thus, in the paper, which is an extend publication version of the abovementioned report: A reality of high-level danger of digital extinction of the Georgian and Abkhazian languages is proved; Aims, problems and results of complete technology support of the Georgian and Abkhazian languages are overviewed; Recommendations, which have been supported by # 53 protocol from 2019 on 17 May of the center for innovation and high technologies of the Georgian national academy of sciences, are reported.

### 1 Introduction

On 27 October 2017, on the Abkhazian Language Day, during government meeting, the former prime minister of Georgia Giorgi Kvirikashvili "congratulated all and, especially, Abkhazian brothers and sisters on the Abkhazian Language Day. The prime minister considers that state program for protection and development of Abkhazian language will create a solid foundation in the confidence building process. The program provides for a variety of activities. Naturally, a relevant budget will be allocated and I am convinced that the society will be actively engaged in the implementation of this program. I believe that our honest and comprehensive approach towards Abkhazian language and culture will be pivotal in terms of restoring the burnt bridges, which are crucial for our unity."<sup>1</sup>

Thus, the above quoted is a very clear confirmations of the right attitude of the Georgian state towards the Abkhazian language. However, if we take into account article 37 of the "organic law of Georgia on official language"<sup>2</sup> according to which "unified program of the official language" aimed at the protection and development of the Georgian state languages – of the Georgian and Abkhazian languages should has been acted from the February 14, 2016, but, unfortunately, it is not acting till today.<sup>3</sup>

The statements of the president of Georgia Mrs. Salome Zurabishvili, which are quoted below, are also very clear confirmations of the right attitude of the Georgian state not only towards the Abkhazian and Georgian languages, but, in general, towards the Caucasian languages too:

- 1. "Today, as never before, the Abkhazian language and identity need protection."  $^{4}$
- 2. "Georgia's second state language, Abkhazian, is under huge threat today. The State of Georgia, its Constitution protects the Abkhazian language, but it is disappearing. Therefore, as President of Georgia, I feel a special responsibility. In all my official visits, I discuss the grave situation the Russification policy has created for the Abkhazian people, their culture and language."<sup>5</sup>

<sup>&</sup>lt;sup>1</sup>See at the address http://gov.ge/index.php?lang\_id=ENG&sec\_id=463&info\_id= 62686 - State Program for Protection and Development of Abkhazian Language to be Launched.

<sup>&</sup>lt;sup>2</sup>We will very shortly overview this very important law below.

<sup>&</sup>lt;sup>3</sup>For clarity, we emphasize that we have no doubt about the sincerity of the former prime minister of Georgia in relation to the Abkhazian language. Thus, this convinces us that he was not completely informed by his surroundings on this issue, since on May 27, 2016 we provided extensive and comprehensive information on this issue in an official letter, where we asked him to accelerate an activation of the "unified program of the official language" in a timely manner.

<sup>&</sup>lt;sup>4</sup>See at the address https://www.president.gov.ge/eng/prezidenti/inauguracia. aspx - 18 December 2018, President Salome Zourabichvili's Inauguration Day.

<sup>&</sup>lt;sup>5</sup>See at the address https://www.president.gov.ge/eng/pressamsakhuri/ siakhleebi/saqartvelos-prezidentma-saqartvelos-moqalaqeebs-da.aspx - 26 May, 2019, Independence Day.

- 3. "Defense of the state language is my constitutional obligation and is both a high responsibility and a great pride. Language is the state's core. A unified state cannot exist without a language. Language unifies society and its second characteristic is that it carries a nation's culture. Culture cannot exist without language. Today's globalization and technologies create large challenges for languages."<sup>6</sup>
- 4. "The Georgian scientific community requests the senior officials to establish a coordination center for Georgian and Caucasus studies in order to preserve Caucasian languages and cultures in the age of globalization."<sup>7</sup>

Together with above quoted completely right views of the President of Georgia Mrs. Salome Zurabishvili, it becomes more unclear why "unified program of the official language" aimed at protecting of the Georgian and Abkhazian languages is not acting till today.<sup>8</sup>

In addition, the fact, that Georgian state has really very responsible and right attitude to the Georgian and Abkhazian languages, is also proven by the fact, that the organic law of Georgia on official Language was approved finally on 22 July 2015.<sup>9</sup> For more clarity, below, we have quoted the  $37^{\rm th}$  and  $4^{\rm th}$  articles of this law:

"article 4 – Status of the official language: 1. In accordance with Article 8 of the Constitution of Georgia, the official Language of Georgia is Georgian and the official language of the Autonomous Republic of Abkhazia is Georgian as well as Abkhazian. 2. The State ensures the protection of the constitutional status of the official language throughout the whole territory of Georgia. 3. The State continually facilitates the preservation and exploration of the Kartvelian Languages and dialects, which is one of the most significant preconditions for maintaining viability and resilience of the official language."

<sup>&</sup>lt;sup>6</sup>See at the address https://www.president.gov.ge/eng/pressamsakhuri/ siakhleebi/%E2%80%8Bsalome-zurabishvili-mshobliuri-ena-chveni-identob.aspx - May 31, 2019, at the conference "State language for civic integration and human capital development in Georgia".

<sup>&</sup>lt;sup>7</sup>See at the address https://www.president.gov.ge/eng/pressamsakhuri/ siakhleebi/salome-zurabishvili-unda-gaaqtiurdes-mushaoba-saqa.aspx - 12 October, 2019, President Salome Zourabichvili on Intensify Scientific Studies of Georgia and Caucasus and Provide Objective Information to International Community.

<sup>&</sup>lt;sup>8</sup>If the "Unified program of the official language" was be acting, we would be having today very well grounding for the implementation of very important researches for the protecting and developing all other Caucasian languages too, need of which is fully rightly emphasized the President of Georgia Mrs. Salome Zurabishvili.

<sup>&</sup>lt;sup>9</sup>See at the address https://matsne.gov.ge/en/document/view/2931198? publication=3

"article 37 – Unified Program of the Official Language: 1. The Department of Official Language shall submit the unified programme of the official language to the Government of Georgia for approval upon the recommendation of the Experts Committee of the official language; 2. Public authorities designated by the Government of Georgia shall ensure implementation of the unified programme of the official language. 3. The unified programme of the official language aims to: a) meet maximum requirements of persons who are willing to learn the official language; prepare methodological and educational resources in order to teach the official language and to raise linguistic culture; teach the official language in accordance with contemporary requirements, and to introduce a bilingual teaching programme for the groups of linguistic minorities; b) conduct researches in the field of the structure, history and functional features of the official language in a consistent manner; ensure a unified lexicography of the contemporary Georgian language and provide a full set of norms and terminology standards of the Georgian literary language; c) provide complete technological support of the official language; create contemporary and comprehensive digital database of linguistic data (texts); develop search, analytical-operative and translation software. 4. The unified programme of the official language is funded by the State Budget of Georgia. 5. The State shall enhance the attraction of investments for the implementation of the unified programme of the official language."

All of the above, as whole, makes it very clear that in the rapidly forthcoming digital age the Georgian state take care for protection and development of its state languages – Georgian and Abkhazian languages. This, on the one hand, is obvious truth, but, on the other hand, it must be strongly emphasized that, today, 2020 is almost over and, nevertheless, the most important article of the "organic law of Georgia on official language" – the article 37 is still out of force.<sup>10</sup> Thus, the main reason why we are publishing

<sup>&</sup>lt;sup>10</sup>According to our information, the most active supporters of the organic law of Georgia on the official language together with Georgian national academy of sciences was former prime minister of Georgia Irakli Garibashvili (20.11.2013 - 29.12.2015). In additions to this it must be underlined, that we addressed an official letter to the former prime minister (from October 2012 to November 2013) of Georgia, Mr. Bidzina Ivanishvili (together with us this letter was signed by Lasha Abzianidze and Aleksandre Maskharashvili), in March 2013, where we asserted the reality of the danger of digital extinction of the Georgian language in the rapidly forthcoming digital age and, accordingly, we asked for timely initiation of the necessary measures to protect from this danger the state languages of Georgia. Also, at that time, our friend, famous political figure Pridon Sakvarelidze was member of the parliament committee of education, science and culture (17.11.2012 - 26.10.2015) and in the March of the same 2013 we had a long conversation with him about the dangers of digital death of languages. During this conversation we introduced him the alarming results of META-NET two-years research "Europe's Languages in the Digital Age" as well as the alarming low level of the technological support of

another paper dedicated to the aims of defence of Georgian and Abkhazian languages from the danger of the digital extinctions is that, unfortunately, "unified program of the official language", which was to be put "into force from 14 February 2016", not only does not act, but, moreover, it are not elaborated yet.<sup>11</sup>

But, it should be noted also, that in 2019, with financial and institutional support of Shota Rustaveli Georgian national science foundation, Georgian national academy of sciences and Georgian technical university the First Tbilisi International Summer School "Logic, Language, Artificial Intelligence" was organized, dedicated to the memory of Shalva Pkhakadze.<sup>12</sup> The summer school was aimed to provide the complete technological support of Georgian and Abkhazian languages. – This gives hope, that defense of Georgian state languages is really important for Georgian state and that in the near future according to the article 37 of the "organic law of Georgia on official language" the "unified program of the official language" will be elaborated. – In any case, this is our goal and we will try everything to make it happen in the very near future.

Georgian language. As a result all of these, on July 8, 2013, with the initiative of Fridon Sakvarelidze and with the support of the former chairman of the parliament of Georgia David Usupashvili (2012 - 2016) and former chairperson of the education, science and culture committee of parliament of Georgia Ivane Kiguradze (2012 - 2016) parliamentary conference "Georgian Language – Challenges of the 21<sup>st</sup> Century" was been organized, where we made a report "The Technological Alphabet of The Georgian Language – The One of The Most Important Georgian Challenge of The XXI Century". We say all this in order to make it clear to the readers that the first parliamentary and governmental team of the "Georgian Dream – Democratic Georgia" were working very intensive and fruitful for the aims of defense of Georgian state languages from the danger of digital extinction, which, we repeat, is clearly evidenced by the organic law of Georgia on the official Language finally approved on July 22, 2015 (this fact together the fact that Saakashvili's government completely blocked the ongoing processes for the protection and development of the Georgian language, is an important feature of the European and national nature of both the old and the new government of Georgia).

<sup>&</sup>lt;sup>11</sup>By this we want to say that, the second parliamentary and governmental team of the "Georgian Dream – Democratic Georgia" (this refers to the period from 30 December 2015 to 2 September 2019, when prime ministers of Georgia were Giorgi Kvirikashvili and Mamuka Bakhtadze (On September 7, 2018, we addressed an official letter to him with the aims of defence state languages of Georgia from danger of digital extinction, but there was not any result!)) had severely hampered the rapid progress made in previous years towards the aims of defense of Georgian state languages from the danger of digital extinction, because of which increased clearly the quality of dangers of digital extinction, which are faced Georgian and Abkhazian languages today.

<sup>&</sup>lt;sup>12</sup>See at the address https://geoanbani.com/TbiLLAI/

# 2 Aims and Problems of the Complete Technological Support of the Georgian and Abkhazian languages

### 2.1 The Aims of the Complete Technological Support of the Georgian and Abkhazian languages

The long-term projects "Technological Alphabet of the Georgian Language" (launched from 2012) and "Plan-Program for the Complete Technology Support of the Abkhazian Language" (launched from 2015) are aimed at providing complete technological support of the Georgian and Abkhazian languages. At the same time, a complete technological support of the Georgian and Abkhazian languages implies, first, to construct Georgian and Abkhazian technological alphabets, in other words, this implies to build computer systems almost completely knowing Georgian and Abkhazian languages and, second, this implies to equip these Georgian and Abkhazian computer systems with multilingual translation abilities [1–4].

Our above-mentioned aims are fully in line with the very important European aims of the META-NET,<sup>13</sup> which were first voiced in the guideline style publication "Strategic Research Agenda for Multilingual Europe 2020"<sup>14</sup> of the same center. Truly, in this for us very important publication, to the question "What are language technologies?" is given the following answer: "Language technologies are technologies for automatically analysing and generating the most complex information medium in our world, human language, in both its spoken and written forms (as well as sign language)." In addition to this, to prove that our aims are really fully in line with the very important European aims of the meta-network, we quote here from the paper "Strategic Research Agenda for Multilingual Europe 2020", according to which the META-NET has aimed to the "multilingual European society, in which all citizens can use any service, access all knowledge, enjoy all media and control any technology in their mother tongues. This will be a world in which written and spoken communication is not hindered anymore by language barriers and in which even specialised high-quality translation will be affordable".<sup>15</sup>

<sup>&</sup>lt;sup>13</sup>Multilingual Europe Technology Alliance Network of Excellence, shortly META-NET (http://www.meta-net.eu/), is dedicated to building the technological foundations of a multilingual European information society.

<sup>&</sup>lt;sup>14</sup>"Strategic Research Agenda for Multilingual Europe 2020" (http://www.meta-net. eu/sra) presented by the META Technology Council is a guideline style publication edited by Georg Rehm and Hans Uszkoreit. It was published on December 1, 2012 and is the result of a more then two years discussion between hundreds of experts from research and industry.

<sup>&</sup>lt;sup>15</sup>On the matter also noteworthy is the publication "Language technologies for a mul-

Thus, it is clear from the above, that the aims of the long-term projects "Technological Alphabet of the Georgian Language"<sup>16</sup> and "Plan-Program for the Complete Technology Support of the Abkhazian Language"<sup>17</sup> are in fully line with the most important European aims of the META-NET, as well as that these long-term projects are aimed to reach the perfect solutions of the problems of construction artificial intelligence systems for the Georgian and Abkhazian languages, which are the problems of the top difficulties laid in interdisciplinary area of logic, language, and artificial intelligence.

<sup>17</sup> It should also be noted that since 2015, the long-term project "Plan-Program for the Complete Technology Support of the Abkhazian Language" has been separated from the long-term project "Technological Alphabet of the Georgian Language" as an independent project. We have also to emphasize that it is impossible to construct a computer system, which will have almost complete knowledge of the Abkhazian language without the direct involvement in this researches specialists who are naturally familiar with the Abkhazian language! – Therefore, at this stage, one of the main aims of the "Plan-Program for the Complete Technology Support of the Abkhazian Language" is to ensure the voluminous involvement of Abkhazian sciencyists in the research planned by the project. Accordingly, if today there are not sufficient amount of such scientists among Abkhazians, then we consider it as our duty to speed up their upbringing and training [3, 4, 17].

tilingual Europe" (https://langsci-press.org/catalog/book/106), published in 2018 and edited by Georg Rehm, Felix Sasaki, Daniel Stein and Andreas Witt, in which to the same question – "What are language technologies?" is given the following answer: "In the next IT revolution computers will master our languages. Just as they already understand measurements and formats for dates and times, the operating systems of tomorrow will know human languages." It should be emphasized here that the exact same vision for the computers of the future was first expressed in the publication of the META-NET "Strategic Research Agenda for Multilingual Europe 2020". – Also, we must emphasized here that the exact same vision for the computers of the future was declared by us firstly in 2001 [4,5].

<sup>&</sup>lt;sup>16</sup>The long-term project "Technological Alphabet of the Georgian Language" was elaborated on the basis of the aims [5-8] and results [9-16] of state priority program "Free and Complete Programming Inclusion of a Computer in the Georgian Natural Language System" of the Ivane Javakhishvili Tbilisi state university, which, in turn, was elaborated by K. Pkhakadze in 2000-2002 years with the same aims i.e. with aims of construction computer system almost completely knowing Georgian language. This state priority program was leading by K. Pkhakadze in 2002-2007 years at the Ilia Vekua institute of applied mathematics, in 2008-2009 years at the open institute for the Georgian language, logic and computer and in 2009-2010 years at the Saint Andrew the first-called Georgian university of the patriarchate of Georgia. - To make it more clear to the reader the aims and content of this state priority program developed almost twenty years ago, we are quoting from Pkhakadze's article "Globalization, Georgian Language and State Priority Program - Free and Complete Programming Inclusion of a Computer in the Georgian Natural Language System" published in 2005: "the aims outlined here are to equip computers with the general lingual-logical thinking skills. This means to build thinking machines, that in future will be full-scale intellectual partners of humans. Thus, the Georgian thinking machine is the necessity without which our state, linguistic and cultural existence in the forthcoming age of thinking machines may be even more marginalized, than we can imagine today" [7].

At the same time, and it is also very clear, that the solutions of these problems have especially high importance for cultural future of the Georgian and Abkhazian languages, because of only in the case of perfect solutions of these problems it will be possible to defense cultural future of Georgian and Abkhazian languages in rapidly forthcoming digital age and, accordingly, to enter in the European union, in more general, in the future cultural digital world with technologically completely supported Georgian and Abkhazian languages.<sup>18</sup> – It is very clear that in the digital age – in the age of such computers, which will have an almost complete knowledge of languages, those languages, which do not have such lingual computers, will be completely marginalized from the global cultural processes!

Thus, to summarize, it can be said, that complete technological support of the Georgian and Abkhazian languages – in other words – the long-term projects "Technological Alphabet of the Georgian Language" and "Plan-Program for the Complete Technology Support of the Abkhazian Language" aimed, first, to protect Georgian and Abkhazian languages from danger of the digital extinction and, second, to enter Georgian state in the European Union with completely supported Georgian and Abkhazian languages, that, we think, in whole, is necessary for the preservation of Georgian and Abkhazian cultural identity in today very rapidly forthcoming digital age.

<sup>&</sup>lt;sup>18</sup>In addition to full support in democratic and anti-occupation movements one of the main reasons why we are so keen in the enter in the European Union is an already very strictly defined language policy of European Union, which is based on eco-linguistic views and directly aimed on the defense of the different languages and identities. – Below are given two fragments from "Strategic Research Agenda for the Multilingual Europe 2020" done by META technology council as short confirmations of the already mentioned: 1. "Everybody must have the chance to communicate efficiently in the enlarged EU. This does not only affect those who already are multilingual but also those who are monolingual or linguistically less skilled. The media, new technologies and human and automatic translation services can bring the increasing variety of languages and cultures in the EU closer to citizens and provide the means to cross language barriers. They can also play an important role to reduce those barriers and allow citizens, companies and national administrations to exploit the opportunities of the single market and the globalising economy. Faced with the globalising online economy and ever-increasing information in all imaginable languages, it is important that citizens access and use information and services across national and language barriers, through the internet and mobile devices. Information and communication technologies (ICT) need to be languageaware and promote content creation in multiple languages." 2. "The Council of the European Union ... encourage[s] the development of language technologies, in particular in the field of translation and interpretation, firstly by promoting cooperation between the Commission, the Member States, local authorities, research bodies and industry, and secondly by ensuring convergence between research programmes, the identification of areas of application and the deployment of the technologies across all EU languages."

### 2.2 The Problems of the Complete Technological Support of the Georgian and Abkhazian Languages

Problems of complete technological support of the Georgian and Abkhazian languages are mainly conditioned, first, by difficulties of constructing technological alphabets of the Georgian and Abkhazian languages, in other words, by difficulties of constructing computer systems almost completely knowing the Georgian and Abkhazian languages, second, by the absence of a unified Georgian research center, which will be targeted to solve these very difficult and very important problems, and, third, by low quality of the support of Georgian and Abkhazian languages with the language technologies and resources. Because of that, that from these three the first two are without doubt [17–20], below, we will try to prove the third.<sup>19</sup>

Thus, below, on the basis of the very alarming results of the META-NET's two-year research "Europe's Languages in the Digital Age"<sup>20</sup> and, also, META-NET's press-release "At Least 21 European Languages in Danger of Digital Extinction – Good News and Bad News on the European Day of Languages"<sup>21</sup> we will try to prove, that Georgian state languages – Georgian and Abkhazian languages are in the high level danger of digital extinction in the rapidly forthcoming digital age.

On 20 September, 2012 – on the European Day of Languages, on the basis of the alarming results of the study "Europe's Languages in the Digital Age" was published a very alarming press-release "At least 21 European languages in danger of digital extinction – good news and bad news on

<sup>&</sup>lt;sup>19</sup>It is also clear and, accordingly, it is not controversial, that for countries with such a small population as our country, the most optimal and effective way to solve such problems is the timely formation of the unified research center targeted to solve these very difficult and very important problems. For more clarity, the fact, that, for one side, in July 22, 2015 the Georgia law on official language was finally approved, and, for the second side, till today i.e. during four years, there is not been elaborated the "unified program of the official language", makes very clear that in our country the wrong attitudes prevail over in these very important issues during past four years.

<sup>&</sup>lt;sup>20</sup>A all European study conducted by META-NET in 2010-2012, "European Languages in the Digital Age" (http://www.meta-net.eu/whitepapers/overview), also known as the White Paper Series, was published in 32 volumes and includes 31 European language. These languages are the Basque, Bulgarian, Catalan, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Greek, Hungarian, Icelandic, Irish, Italian, Latvian, Lithuanian, Maltese, Norwegian (bokmål), Norwegian (nynorsk), Polish, Portuguese, Romanian, Serbian, Slovak, Slovene, Spanish, Swedish, and Welsh languages.

<sup>&</sup>lt;sup>21</sup>META-NET Press Release "At Least 21 European Languages in Danger of Digital Extinction - Good News and Bad News on the European Day of Languages" (http://www.meta-net.eu/whitepapers/press-release) reports on the state of 30 European languages with respect to Language Technology and explains the most urgent risks and chances. The series covers all official EU Member State languages and several other languages spoken in Europe.

the European day of languages", according to which "a new study by Europe's leading Language Technology experts warns", that "most European languages are unlikely to survive in the digital age".

This two-years "study, prepared by more than 200 experts and documented in 30 volumes of the meta-net white paper series, assessed language technology support for each language in four different areas: automatic translation, speech interaction, text analysis and the availability of language resources. A total of 21 of the 30 languages (70%) were placed in the lowest category, "support is weak or non-existent" for at least one area by the experts. Several languages, for example, Icelandic, Latvian, Lithuanian and Maltese, receive this lowest score in all four areas. On the other end of the spectrum, while no language was considered to have "excellent support", only English was assessed as having "good support", followed by languages such as Dutch, French, German, Italian and Spanish with "moderate support". Languages such as Basque, Bulgarian, Catalan, Greek, Hungarian and Polish exhibit "fragmentary support", placing them also in the set of high-risk languages."

The results of this META-NET's two-year study, "European Languages in the Digital Age" are summarized in the below given Table 1, which we took from a for us very important paper "Strategic Research Agenda for Multilingual Europe 2020".<sup>22</sup>

To make more visible the results presented by the table 1, we rated "excellent support" at 5 points, "good support" at 4 points, "moderate support" at 3 points, "fragmentary support" at 2 points, "weak or no support" at 1 point and, after, on base of table 1, we generated below given table 2 named by us as "State of language technology support for Georgian, Abkhazian and 31 European languages in four different areas in points and percents".

In this table we assessed language technology support for the Georgian and Abkhazian languages on base of results of the long-term projects "Technological Alphabet of Georgian Language" and "Plan-Program for Complete Technology Support of the Abkhazian Language" [1,3,5], which are overviewed in next paragraph of this paper.

Thus, the table 2 built by us below together with the previous assessments of the META-NET makes it clear that European languages, which

<sup>&</sup>lt;sup>22</sup>However, we have added the Welsh language to the table on the basis of the results of paper "The Welsh Language in the Digital Age" (http://www.meta-net.eu/ whitepapers/volumes/welsh) prepared by Jeremy Evas. At the same time, we underline, that according to the Europeans themselves, this table shows alarming differences in technology support between languages. English is ahead from the other languages. A few larger languages have a good technology support, when smaller or very small languages have substantial gaps. Many languages lack basic technologies and language resources. Others have these ones, but lack basic semantic methods and tools. Therefore, it is clear, that most of European languages are in danger of digital extinctions.

| Support  | Excellent | Good    | Moderate                 | Fragmentary                         | Weak/none                           |
|----------|-----------|---------|--------------------------|-------------------------------------|-------------------------------------|
| Machine  |           | English | French,                  | Catalan, Dutch,                     | Basque, Bulgar-                     |
| Transla- |           |         | $\operatorname{Spanish}$ | German, Hungar-                     | ian, Croatian,                      |
| tion     |           |         |                          | ian, Italian, Polish,               | Czech, Danish,                      |
|          |           |         |                          | Romanian                            | Estonian, Finnish,                  |
|          |           |         |                          |                                     | Galician, Greek,                    |
|          |           |         |                          |                                     | Icelandic, Irish,                   |
|          |           |         |                          |                                     | Latvian, Lithua-                    |
|          |           |         |                          |                                     | nian, Maltese,                      |
|          |           |         |                          |                                     | Norwegian, Por-                     |
|          |           |         |                          |                                     | tuguese, Serbian,                   |
|          |           |         |                          |                                     | Slovak, Slovene,                    |
|          |           |         |                          |                                     | Swedish, Welsh                      |
| Speech   |           | English | Czech,                   | Basque, Bulgarian,                  | Croatian, Ice-                      |
|          |           |         | $\operatorname{Dutch},$  | Catalan, Danish,                    | landic, Latvian,                    |
|          |           |         | Finnish,                 | Estonian, Galician,                 | Lithuanian, Mal-                    |
|          |           |         | French,                  | Greek, Hungarian,                   | tese, Romanian,                     |
|          |           |         | German,                  | Irish, Norwegian,                   | Welsh                               |
|          |           |         | Italian,                 | Polish, Serbian,                    |                                     |
|          |           |         | Portuguese,              | Slovak, Slovene,                    |                                     |
|          |           |         | Spanish                  | Swedish                             |                                     |
| Text An- |           | English | Dutch,<br>Energy         | Basque, Bulgarian,                  | Croatian, Esto-                     |
| alytics  |           |         | French,<br>German,       | Catalan, Czech,<br>Danish, Finnish, | nian, Icelandic,<br>Irish, Latvian, |
|          |           |         | Italian,                 | Galician, Greek,                    | Lithuanian, Mal-                    |
|          |           |         | Spanish                  | Hungarian, Nor-                     | tese, Serbian,                      |
|          |           |         | Spanish                  | wegian, Polish,                     | Welsh                               |
|          |           |         |                          | Portuguese, Ro-                     | VV CIBII                            |
|          |           |         |                          | manian, Slovak,                     |                                     |
|          |           |         |                          | Slovene, Swedish                    |                                     |
| Language |           | English | Czech,                   | Basque, Bulgarian,                  | Icelandic, Irish,                   |
| Re-      |           |         | Dutch,                   | Catalan, Croatian,                  | Latvian, Lithua-                    |
| sources  |           |         | French,                  | Danish, Estonian,                   | nian, Maltese                       |
|          |           |         | German,                  | Finnish, Galician,                  | ,                                   |
|          |           |         | Hungarian,               | Greek, Norwegian,                   |                                     |
|          |           |         | Italian,                 | Portuguese, Roma-                   |                                     |
|          |           |         | Polish,                  | nian, Serbian, Slo-                 |                                     |
|          |           |         | Spanish,                 | vak, Slovene                        |                                     |
|          |           |         | Swedish                  |                                     |                                     |

Table 1: State of language technology support for 30 Europeanlanguages in four different areas.

summary point are less to 10 point (see below given table) i.e. Hungarian, Polish, Czech, Galician, Catalan, Portuguese, Swedish, Romanian, Slovenian, Slovak, Greek, Finnish, Danish, Bulgarian, Basque, Norwegian, Estonian, Serbian, Irish Croatian, Icelandic, Latvian, Lithuanian, Maltese, and Welsh languages are appreciated as languages under danger of digital extinction.

This fact together of the fact, that language technology support of the Georgian and Abkhazian language is less than technology support of almost any aforementioned European languages, which are in danger of digital extinction, proves clearly a reality of the high-level danger of digital extinction

| Language       | speech   | automatic | text   | language | summary points /           |
|----------------|----------|-----------|--------|----------|----------------------------|
|                | interac- | transla-  | analy- | re-      | percent                    |
|                | tion     | tion      | sis    | sources  | -                          |
| 1. English     | 4        | 4         | 4      | 4        | 16(-4) = 80% (-20%)        |
| 2. French      | 3        | 3         | 3      | 3        | 12(-8) = 60% (-40%)        |
| 3. Spain       | 3        | 3         | 3      | 3        | 12 (-8) = 60% (-40%)       |
| 4. German      | 3        | 2         | 3      | 3        | 11 (-9) = 55% (-45%)       |
| 5. Dutch       | 3        | 2         | 3      | 3        | 11 (-9) = 55% (-45%)       |
| 6. Italian     | 3        | 2         | 3      | 3        | 11 (-9) = 55% (-45%)       |
| 7. Hungarian   | 2        | 2         | 2      | 3        | 9(-11) = 45%(-55%)         |
| 8. Polish      | 2        | 2         | 2      | 3        | 9(-11) = 45%(-55%)         |
| 9. Czech       | 3        | 1         | 2      | 3        | 9(-11) = 45%(-55%)         |
| 10. Galician   | 2        | 1         | 3      | 3        | 9(-11) = 45%(-55%)         |
| 11. Catalan    | 2        | 2         | 2      | 2        | 8 (-12) = 40% (-60%)       |
| 12. Portuguese | 3        | 1         | 2      | 2        | $8\ (-12) = 40\%\ (-60\%)$ |
| 13. Swedish    | 2        | 1         | 2      | 3        | 8(-12) = 40%(-60%)         |
| 14. Romanian   | 1        | 2         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 15. Slovenian  | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 16. Slovak     | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 17. Greek      | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 18. Finnish    | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 19. Danish     | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 20. Bulgarian  | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 21. Basque     | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 22. Norwegian  | 2        | 1         | 2      | 2        | 7(-13) = 35%(-55%)         |
| 23. Estonian   | 2        | 1         | 1      | 2        | 6(-14) = 30%(-70%)         |
| 24. Serbian    | 2        | 1         | 1      | 2        | 6 (-14) = 30% (-70%)       |
| 25. Irish      | 2        | 1         | 1      | 1        | 5(-15) = 25%(-75%)         |
| 26. Croatian   | 1        | 1         | 1      | 2        | 5(-15) = 25%(-75%)         |
| 27. Icelandic  | 1        | 1         | 1      | 1        | 4(-16) = 20%(-80%)         |
| 28. Latvian    | 1        | 1         | 1      | 1        | 4(-16) = 20%(-80%)         |
| 29. Lithuanian | 1        | 1         | 1      | 1        | 4(-16) = 20%(-80%)         |
| 30. Maltese    | 1        | 1         | 1      | 1        | 4(-16) = 20%(-80%)         |
| 31. Welsh      | 1        | 1         | 1      | 1        | 4(-16) = 20%(-80%)         |
| 32. Georgian   | 0.75     | 0.75      | 0.75   | 0.75     | 3(-17) = 15%(-85%)         |
| 33. Abkhazian  | 0.25     | 0.25      | 0.25   | 0.25     | 1 (-19) = 5% (-95%)        |

Table 2: State of language technology support for Georgian, Abkhazian and 31 European languages in four different areas in points and percents

of the technologically insufficiently supported Georgian and Abkhazian languages in the rapidly forthcoming digital age.

In order to make the reality of the threats already presented by us regarding the Georgian and Abkhazian languages even more credible, below are quoted from "Strategic research agenda for multilingual Europe 2020". – Only quick view at them will make clear a very high responsible attitude of the Europeans themselves are experiencing about fate of their languages in very rapidly forthcoming digital age:

Denmark: "If we have the ambition to use the Danish language in the technological universe of the future, an effort must be made now to maintain

and further develop the knowledge and expertise that we already have. Otherwise we run the risk that only people who are fluent in English will profit from the new generations of web, mobile and robot technology which are up and coming." – Sabine Kirchmeier-Andersen (Director of the Danish Language Council).

Portugal: "Language technology is of utmost importance for the consolidation of Portuguese as a language of global communication in the information society." – Pedro Passos Coelho (Prime Minister of Portugal).

Czech Republic: "META-NET brings a significant contribution to the technological support for languages of Europe and as such will play an indispensable role in the development of multilingual European culture and society." – Ivan Wilhelm (Deputy Minister for Education, Youth and Sport).

Greece: "Further support to language technologies safeguards the presence of Greek language and culture in the digital environment". – George Babiniotis (Minister of Education, Lifelong Learning and Religious Affairs).

European Commission: "Having worked on automatic media analysis for many years and in tens of languages, we are painfully aware of the lack of text analysis tools and resources in most languages. META-NET's analysis is very accurate. Language Technology is a key enabling ingredient for future generations of IT. Languages for which no tools and resources will exist soon will not participate in the next major technological developments." – Ralf Steinberger (Joint Research Centre, IPSC - GlobeSec - Open-Source Text Information Mining and Analysis, Ispra, Italy).

Estonia: "If we do not implement the development plan for language technology or do not cooperate with other countries in the same direction, in the future Estonian will be marginalised in information society." – Development Plan of the Estonian Language 2011–2017.

France: "META-NET provides an invaluable contribution to the development of a genuine European strategy in support to multilingualism, based on existing technologies while encouraging the development of new innovative technologies." – Xavier North (Délégué Général à la Langue Française et aux Langues de France).

Malta: "The technology support for the Maltese language should serve our language to be continuously cultivated, used and placed on the same level as other languages." – Dolores Cristina (Minister for Education and Employment).

Lithuania: "Conserving Lithuanian for future generations is a responsibility of the whole of the European Union. How we proceed with developing information technology will pretty much determine the future of the Lithuanian language." – Andrius Kubilius (Prime Minister of the Republic of Lithuania).

Ireland: "Language technology is no longer a luxury for most European

languages – it is now essential to their survival as viable means of expression across the whole range of areas from business to the arts, and this is as much the case for Irish as any other European language." – Ferdie Mac an Fhailigh (CEO, Foras na Gaeilge).

Slovenia: "It is imperative that language technologies for Slovene are developed systematically if we want Slovene to flourish also in the future digital world." - Danilo Türk (President of the Republic of Slovenia).

Iceland: "Language technology is an essential tool in a variety of linguistic research, and supports the official Icelandic policy of promoting the national language in all aspects of communication." - Guðrún Kvaran (Chair of the Icelandic Language Council).

Netherlands and Flanders (Belgium): "It remains extremely important that citizens can use their native language in all circumstances, including when they deal with modern ICT and leisure devices. But usually English speaking people are the first to benefit from such an evolution. Not only does this pose a danger of reducing the overall functionality of a language (and an impoverishment of an entire culture), but also it threatens those groups in society that do not master the universal language. Therefore, R&D programmes that support the local language are needed. Also, in the future, the Dutch Language Union will continue to emphasise this issue." – Linde van den Bosch (General Secretary of the Dutch Language Union, 2004-2012).

Poland: "Language technologies are more and more present in our everyday life. For their presence to be rational and functional, for it to serve the needs of the economy, as well as the social and cultural life well, further large-scale work in this area is needed." - Michał Kleiber (President of the Polish Academy of Sciences).

Luxembourg: "This is a European challenge of enormous importance!" – Roman Jansen-Winkeln (CTO, Belingoo Media Group)

Germany: "Europe's multilingualism and our scientific expertise are the perfect prerequisites for significantly advancing the challenge that language technology poses. META-NET opens up new opportunities for the development of ubiquitous multilingual technologies." – Annette Schavan (Minister of Education and Research)

UK: "The work of META-NET is an important step towards a future in which Language Technology will be all around us, allowing us to collaborate, conduct business and share knowledge with friends and colleagues, whether or not we speak the same language." – David Willets (Minister of State for Universities and Science, Department for Business, Innovation and Skills).

Latvia: "For such small languages like Latvian keeping up with the ever increasing pace of time and technological development is crucial. The only way to ensure future existence of our language is to provide its users with equal opportunities as the users of larger languages enjoy. Therefore, being on the forefront of modern technologies is our opportunity." – Valdis Dombrovskis (Prime Minister of Latvia).<sup>23</sup>

The above passages are texts of 2012 year, and these texts make it clear how Europeans themselves were concerned about the fate of their languages in the digital age in 2012! – This fact together the fact that technological support of the Georgian, even more, of the Abkhazian, lags far behind from the European languages that are in danger of digital extinction, clearly shows the reality of the high-level danger of the digital extinction, which is faced to Georgian, even more, Abkhazian language!

In order to make even more understandable the validity of the abovementioned opinions of high-ranking European politicians and experts, and also in order to make even more credible the reality of the high-level threats of digital death facing Georgian and Abkhazian languages today, below, we have quoted a short fragments from above already overviewed press-release: 1. "The results of our study are most alarming. The majority of European languages are severely under-resourced and some are almost completely neglected. In this sense, many of our languages are not yet future-proof." -The author of the quote Prof. Hans Uszkoreit, coordinator of META-NET, scientific director of the German Research Center for Artificial Intelligence, editor of the META-NET publication "Europe's Languages in the Digital Age". 2. "There are dramatic differences in language technology support between the various European languages and technology areas. The gap between big and small languages still keeps widening. We have to make sure that we equip all smaller and under-resourced languages with the needed base technologies, otherwise these languages are doomed to digital extinction." - The author of the quote Dr. Georg Rehm, researcher of the German Research Center for Artificial Intelligence, co-editor of the META-NET publication "Europe's Languages in the Digital Age".

These expert views of Hans Ushkorait and Georg Rehm, a well-known

<sup>&</sup>lt;sup>23</sup>Here, in the interests of the question, we decided to once again bring an excerpt from the speech of the President of Georgia, Salome Zurabishvili, on 31 May, 2019, at the conference "State Language for Civic Integration and Human Capital Development in Georgia", where she says: "Defense of the state language is my constitutional obligation and is both a high responsibility and a great pride. Language is the state's core. A unified state cannot exist without a language. Language unifies society and its second characteristic is that it carries a nation's culture. Culture cannot exist without language. Today's globalization and technologies create large challenges for languages." – on this background, or with this completely correct and acceptable position of one of the highest political figures of the Georgian state, it is completely incomprehensible the fact that "Unified program of the official language", which are aimed for the protection of the Georgian state languages – Georgian and Abkhazian languages is not working, but, even more, it is not yet elaborated!

European experts in artificial intelligence and language technoligies together with all above already said, we think, is quite enough for any Abkhazian and Georgian who thinks about the fate of their languages to deeply understand that the low technological suport of Georgian, even more, Abkhazian language is a very significant challenge, which, in turn, fully substantiates, that "State Program for the Protection and Development of Georgian and Abkhazian Languages", in other words, "unified program of the official language", which is already legalized by the organic law of Georgia on official language has the high national importance and, accordingly, there is urgent need to develop it in very near future!

## 3 Results of the Complete Technological Support of the Georgian and Abkhazian languages

As we have already mentioned, from technological support point of view, the Georgian and Abkhazian languages alarmingly lag behind compared to almost any of those European languages, which, according to the well-known research "Europe's Languages in the Digital Age" done by Meta-Net with financial support of Euro Structures are under the danger of digital extinction in the rapidly forthcoming digital age.<sup>24</sup>

This clearly indicates the necessity of the overcoming this lagging as soon as it is possible. Thus, the long-term projects "Technological Alphabet of the Georgian Language" and "Plan-Program for Complete Technology Support of the Abkhazian Language" of the Scientific-Educational Center for the Georgian Language Technology<sup>25</sup> are aimed at reducing this alarming lagging and, therefore, these projects are directly engaged with the aims of saving the Georgian and Abkhazian languages from the danger of digital extinction and, also, with the aim of joining the European Union, in more general, in the future cultural world with the technologically completely supported Georgian and Abkhazian languages, in other words, with the Georgian and Abkhazian technological alphabets – with intellectual computer systems knowing Georgian and Abkhazian languages almost completely and perfectly.

<sup>&</sup>lt;sup>24</sup>Compared to Georgian and Abkhazian languages, other Kartvelian and Caucasian languages are even more backward from that European languages, which are under the danger of digital extinction. Therefore, below, we have also briefly reviewed the technological systems developed by us for Mingrelian, Chechen, Kabardian and Lezgian languages, which system are unique in the sense that there are not any other such type systems for these languages.

<sup>&</sup>lt;sup>25</sup>The Center for Georgian Language Technology acting with the aims of defence Georgian state languages was established at Georgian Technical University in 29 December, 2010 with direct support of the rector of university Academician Archil Prangishvili.

The fact, that language technology support of the Georgian and Abkhazian languages is really much less than technology support of European languages, is also proved by below listed language technology systems, which are been created within above shortly overviewed long-term projects.

Namely, these language technology systems are created on the basis of that new methods of Pkhakadze-Chikvinidze, Pkhakadze-Chichua, Pkhakadze-Malidze, which are elaborated on the basis of Logical Grammar of the Georgian Language K. Pkhakadze and, also, on the base of different today well-known tools and platforms within already successfully completed and till now ongoing subprojects of the long-term projects "The Technological Alphabet of the Georgian Language" and "The Plan-Program for the Complete technology support of the Abkhazian Language" [1–4,21–27]. They are:

- FR/362/4-105/12 project "Foundations of Logical Grammar of Georgian Language and its Application in Information Technology" (Project supervisor – K. Pkhakadze; Project was funded by the Shota Rustaveli National Scientific Foundation);
- AR\_№048-13 project "Internet Versions of a Number of Developable (Learnable) Systems Necessary for Creating The Technological Alphabet of the Georgian Language" (Project supervisor – K. Pkhakadze; Project was funded by the Georgian Technical University);
- DO/308/4-105/14 project "In the European Union with the Georgian Language, i.e., the Doctoral Thesis – Georgian Grammar Checker (Analyzer)" (PhD student – M. Chikvinidze, Scientific Supervisor – K. Pkhakadze; PhD was jointly funded by the Shota Rustaveli National Scientific Foundation and Georgian Technical University);
- DO/305/4-105/14 project "In the European Union with the Georgian Language, i.e., the Doctoral Thesis – Georgian Speech Synthesis and Recognition" (PhD student – G. Chichua, Scientific Supervisor – K. Pkhakadze; PhD was jointly funded by the Shota Rustaveli National Scientific Foundation and Georgian Technical University);
- AR\_122/4-105/14 Project "One More Step Towards Georgian Talking Self-Developing Intellectual Corpus" (Project supervisor – K. Pkhakadze; Project was jointly funded by the Shota Rustaveli National Scientific Foundation and Georgian Technical University);
- PHDF-18-1228 project "In the European Union with Georgian and Abkhazian Languages, i.e. the Doctoral Thesis – Elaboration of the New Developing Tools and Methods of the Georgian Smart Corpus

and Improvement of Already Existing Ones" (PhD student – Sh. Malidze, Scientific Supervisor – K. Pkhakadze; PhD was funded by the Shota Rustaveli National Scientific Foundation);

- PhD thesis "Methods and Tools for the Automatic Intellectual Classification of Georgian Texts" (PhD student – C. Demurchev, Scientific Supervisor – K. Pkhakadze);
- PhD thesis "Georgian-Mathematical Automated Multilingual Semantic Translator" (PhD student – N. Okroshiashvili, Scientific Supervisor – K. Pkhakadze).
- PhD thesis "Formalism and Applications of Georgian Language Processing by Machine Learning Methods" (PhD student – B. Mikaberidze, Scientific Supervisor – K. Pkhakadze).

Thus, below is listed Georgian and Abkhazian language technology systems, which are cretead within above listed sub-projects [2,3,17–27]:

### In the area of speech interaction:

- 1. Trial-applied Georgian spoken browser. This system, which is inbuilt in the Georgian universal smart corpus, is unique in the sense that there is not any other such type Georgian system;
- 2. Trial-applied voice managed Georgian reader. This system is unique in the sense that there is not any other such type Georgian system;
- 3. Trial voice managed Abkhazian reader. This system is unique in the sense that there is not any other such type Georgian system;
- 4. Trial-applied mobile and internet versions of the Georgian spoken assistant for speech disorder persons. – These systems are unique in the sense that there are not any other such type Georgian system;
- 5. Trial versions of the adapted Georgian Internet, Wikipedia, and Computer. – These systems are unique in the sense that there are not any other such type Georgian systems.

#### In the area of automatic translation:

 Trial self-developing Georgian-Mathematical translator. – This system, which is inbuilt in the Georgian universal smart corpus, is unique in the sense that there is not any other such type Georgian system;

- 2. Trial self-developing two-way Georgian-English-German semantic translator based on the mathematical language. – This system, which is inbuilt in the Georgian universal smart corpus, is unique in the sense that there is not any other such type Georgian system;
- 3. Trial-applied internet and mobile versions of Georgian multilingual voice lexicon. – These systems, which are based on google translate, are unique in the sense that there are not any other such type Georgian systems;
- 4. Trial internet and mobile versions of hybrid Georgian multilingual spoken assistant and google translate. – These systems, which are based on google translate, are unique in the sense that there are not any other such type Georgian systems;
- 5. Trial internet and mobile versions of multilingual textual and voice messages between Georgian smart papers. These systems, which are inbuilt in the Georgian universal smart corpus and based on google translate, are unique in the sense that there are not any other such type Georgian systems.

#### In the area of text analysis:

- Trial-applied versions of the automatic processing texts and websites for automatic creation of the self-developing corpuses. – These systems, which are inbuilt in the Georgian universal smart corpus, are unique in the sense that there are not any other such type systems in Georgia not for Kartvelian not for Caucasian languages;
- 2. Trial-applied and trial versions of taggers, descriptors and generators for Georgian N, A and V type words. – These systems, which are inbuilt in the Georgian universal smart corpus, are unique in the sense that there are not any other such type Georgian systems;
- 3. Trial-applied Georgian self-developing orthograph checker. This system, which is inbuilt in the Georgian universal smart corpus, is unique in the sense that there is not any other such type Georgian system;
- 4. Trial Abkhazian self-developing orthograph checker. This system, which is inbuilt in the Georgian universal smart corpus, is unique in the sense that there is not any other such type Abkhazian system;

- Trial Georgian self-developing syntax checker. This system, which is inbuilt in the Georgian universal smart corpus, is unique in the sense that there is not any other such type Georgian system;
- Trial versions of the Georgian text logical analyzer and question answerer. – These systems, which are inbuilt in the Georgian universal smart corpus, are unique in the sense that there are not any other such type Georgian systems;
- 7. Trial versions for automatic generating and testing of the Georgian logical tasks and analogies with voice managed tools. – These systems are unique in the sense that there are not any other such type Georgian systems;
- 8. Trial-applied version of the Georgian printed and scanned texts recognition. – This system is not unique, because that there is existed some such type trial-applied systems for Georgian language constructed on the base of different well-known platforms, but our approach is differed from all others, because of our system is constructed independently from other existed platforms;
- Trial version of the Abkhazian printed and scanned texts recognition. – According to our information, this system is unique in the sense that there is not any other such type Abkhazian system;
- 10. Trial system of Georgian smart paper and smart Journal. These systems, which are inbuilt in the Georgian universal smart corpus, are unique in the sense that there are not any other such type Georgian systems;
- 11. The trial-applied system for Georgian texts classification. This system, which is constructing as inbuilt tool in the Georgian universal smart corpus for the aims of logical, intellectual, and thematical classifications of the texts of the corpus, is unique in the sense that there is not any other such type Georgian system.

#### In the area of language resources:

 Trial-applied version of Georgian universal (i.e. multilingual and multimodal) smart (i.e. self-diveloping, interactive, intellectual, communicative) corpus (http://corpus.ge/). – This corpus,<sup>26</sup> which is first self-developing, interactive, intelectual, multilingual

<sup>&</sup>lt;sup>26</sup>In spite of the fact, that by now this corpus is only step toward the complete Georgian Universal Smart Corpus, It is the largest Georgian corpus (For today, and today is 19 Juli, 2020 it contains 331 782 593 word-tokens, among them 6 104 732 are different ones).

and multimodal Georgian corpus, is unique in the sense that there is not any other such type corpus for Georgian;  $^{27}$ 

- 2. Trial-applied version of the self-developing Georgian corpus, which is inbuilt in the Georgian universal smart corpus as separately cluster. – It develops its own volume and intellectual (ortografic, syntactic, logic, semantic, statistic) abilities itself. Thus, this corpus is unique in the sense that there is not any other such type Georgian corpus;
- 3. Trial version of the self-developing Abkhazian corpus, which is inbuilt in the Georgian universal smart corpus as separately integrated cluster. – It develops itself its own volume and intellectual (ortografic, statistic) abilities. Thus, this corpus is unique in the sense that there no any other such type Georgian corpus;
- 4. Trial versions of the self-developing Chechen, Kabardian, Lezgian and Mingrelian corpuses. – These corpuses are first selfdeveloping corpuses for these Kartvelian and Caucasian languages. They develop themself their own volumes and intellectual (statistic) abilities. Thus, these corpus are unique in the sense that there are no any other such type Kartvelian and Caucasian corpuses;
- 5. Trial parallel corpuses of Georgian-English and Georgian-German words. – They can develop themself on the basis of internet sources. Also, they have the opportunity to rate a quality of the translation and in the case of a translation error users may correct the incorrect translations. Thus, these corpuses are unique in the sence that there are not any other such type Georgian corpuses;
- 6. Trial versions of the Georgian and Abkazian titrated speech data. - The corpuses, which are constructed for the aims of improve quality of our Georgian and Abkhazian TTS and STT systems, conatin segmentators and generators for the Georgian and Abkhazian titrated speech data. Thus, these trial corpuses are unique

<sup>&</sup>lt;sup>27</sup>The deal is that, this Georgian universal smart corpus, from one side, is a corpus, but, at the same time, on the basis of it inbuilt technology systems is equipped with self-developing, interacting, translating, analyzing, and communicating abilities, is also a laboratorial prototype of the united Georgian state internet smart network. We believe, that such type state networks in future – in forthcoming digital age, in other words, in the age of computers, which will have almost complete knowledge of languages, must be property of the all various different states and not only of a private company or only a few different states. For more clarity: For us it is very clear, that the Georgian technological alphabet i.e. the computer system almost completely knowing Georgian language must be property of the Georgian state.

in the sense that there are not any other such type Georgian and Abkhazian corpuses;

- Trial tools for constructing Chechen, Kabardian, Lezgian and Mingrelian titrated speech data. – These tools are are unique in the sense that there are not any other such type tools for these Kartvelian and Caucasian languages;
- 8. Foundations of Logical Grammar of Georgian Language. This is a first mathematical theory – mathematical grammar for the Georgian language. This theory, on the basis of which it is constructed more part from above listed systems, from one side, is based on the Shalva Pkhakadze's Notation Theory, and, from second side, it is based on the natural mathematical specifics of Georgian language.

Thus, the above listed unique technologies and resources for the Georgian and Abkhazian languages show us that if we have only trial-applied and trial systems for Georgian and Abkhazian languages, for English as well as for many other European languages are already prepared the same type applied and trial-applied systems with sufficiently high quality! – This is the reason why in the table 2 we have appreciated the technology support of Georgian language by 15%, and of Abkhazian language by 5%. – These facts prove reality of very high-level danger of digital extinction of the Georgian and Abkhazian languages in the rapidly forthcoming digital age in full.

At the same time, the fact that the above listed language technologies and language resources were elaborated by a very small group with a very small funding, argues that it will be possible for our country to provide complete technological support of Georgian and Abkhazian languages in the case of appropriate state care on these issues, which, in our opinion, should be firstly reflected in the timely formation of the "Research Institute for Technological Development and Cultural Defense of Georgian State Languages".

# 4 Conclusion or Our Main Recommendations for the aims of Complete Technological Support of the Georgian and Abkhazian languages

As it was already mentioned, on May 17, 2019, a report "In the European Union with the Georgian and Abkhazian languages – Aims and Problems of Complete Technology Support of Georgian and Abkhazian Languages" was presented at the center for innovation and high technologies of the Georgian national academy of sciences by the center for Georgian language technology of the Georgian technical university.

Taking into consideration the aims of defense of the state languages of Georgia – Georgian and Abkhazian languages from danger of digital extinction and, also, taking into consideration the arguments presented in this paper above, by the #53 protocol, the Center for innovation and high technologies of Georgian national academy of science agreed with our recommendations made by us at this presentation fully. Therefore, and it is very natural, that the same recommendations in shorted way is presented below as our main recommendations.

Thus, for the aims of Complete Technological Support of the Georgian and Abkhazian languages we are recommending:

- 1. The timely formation the "Research Institute for Technological Development and Cultural Defense of Georgian State Languages";
- 2. The timely define "Technological development and cultural defense of Georgian state languages" as a one of main priority of the Shota Rustaveli National Science Foundation of Georgia;
- 3. To held the annual conferences "Georgian state languages in the digital age: cultural and technological aspects" twice a year (on April 14 in Georgian language day, and on October 27 in Abkhazian language day). This will allow us constantly observe dynamics and results of processes current with aims of the cultural defense and technological development of the Georgian state languages.

In addition we underline, and it is very clear, that by the taking these above listed recommendations into account, first, it will become very clear that in the forthcoming digital age Georgian state really take care on the protecting the Georgian and Abkhazian languages from the rapidly increasing danger of the digital extinction, and, second, it will also become very clear to Abkhazians, which reside in occupied Abkhazia, that for the Georgian state is a vital important to preserve Abkhazian language and identity. – Thus, and it is very clear also, that all abovementioned in sum will be very important step into the direction of reconstruction of today's broken bridge of the Georgian-Abkhazian relations. – The fact together with all the above-mentioned recommendations should be considered also as the first important step toward the cultural protection and technological development of all other Kartvelian and Caucasian languages, that, we think, is a matter of the common Caucasian commitment of the Georgia.

The last: In the preamble of the "organic law of Georgia on official language" there is stressed the following: "The Georgian language is a historical and cultural heritage of Georgia and it is essential for its statehood. It is an element of common national identity for all citizens of Georgia. The State of Georgia performs all its functions in this language, supports it and supports the policy for the development and functioning of this language as of the official language of the State". – Thus, here, according to this very important content of the preamble of the "organic law of Georgia on official language", we direct to the Government of Georgia with our main recommendation: For the aim to defense the state languages of Georgia from danger of digital extinction it is vitally necessary immediate elaboration of the "unified program of the official language" with the necessary involvement of Georgian specialists of artificial intelligence and language technologies, which are working as in Georgia, as well as out of Georgia.

And finally, we underline one more time, that a computer, which will have almost complete knowledge of a language, must be property of that national state in confine of which this language is used! – We believe that this is the only path to healthy and peaceful globalization.

## References

- PKHAKADZE K., CHICHUA G., CHIKVINIDZE M., MASKHARASHVILI A. The Technological Alphabet of The Georgian Language – Aims, Methods and Results, *Reports of Enlarged Session of the Seminar of* VIAM, Volume 27, 48-51, 2013.
- PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., MALIDZE SH., KURTSKHALIA D., DEMURCHEV C., OQROSHIASHVILI N. In the European Union with Georgian and Abkhazian languages – Aims and Problems of Complete Technology Support of Georgian and Abkhazian Languages. Bulletin of the Georgian National Academy of Sciences, (in publishing). 2020.
- MALIDZE SH., PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., KURTSKHALIA D. The final summary of the results of the Phd Project "In the European Union with Georgian and Abkhazian languages, i.e The Doctoral thesis – Elaboration of the New Developing Tools and Methods of the Georgian Smart Corpus and Improvement of Already Existing Ones". Journal Georgian Language and Logic, Vol. 12, pp.62-164, 2019-2020.
- 4. PKHAKADZE K. Free and Complete Programming Inclusion of a Computer in Georgian Natural Language System. *Newspaper "Resonance"*, 242-2362, 2001.

- PKHAKADZE K., IVANISHVILI M. Toward the formal-logical adequate of the natural Georgian language system. Georgian-European University, Paris, *Journal "Iveria*", Vol. VI, 129-148. 2001.
- PKHAKADZE K. Globalization, The Georgian Language and The State Priority Program "Free and Complete Programming Inclusion of a Computer in Georgian Natural Language System". Journal Georgian Language and Logic, vol. 2, pp.1-11, 2005.
- PKHAKADZE K. TSU State Priority Program "Free and Complete Programming Inclusion of a Computer in Georgian Natural Language System" – Aims, Results and Perspectives. Main and Additional text-books in Contemporary Mathematical linguistics, Vol.1, pp.1-18, 2005.
- PKHAKADZE K. On the Linguistic Relation and Logical Declension in the Georgian Language. Journal Georgian language and Logic, Vol.1, pp.19-77, 2005.
- PKHAKADZE K., CHICHUA G., VASHALOMIDZE A., GABUNIA K., ABZIANIDZE L., MASKHARASHVILI A., CHIKVINIDZE M. The Grounding Questions of The Mathematical Theory of The Georgian Language and Thinking and Some Subsystems of The 1st Version of the Voice Managed Georgian Intellectual Computer System. Proceeding of The I International Conference of Chikobava institute of Linguistics on Georgian Language and Modern Technology, pp.60-66, 2009.
- PKHAKADZE K., CHICHUA G., CHIKVINIDZE M., MASKHARASHVILI A. The Short Overview of the Aims, Methods, and Results of the Logical Grammar of the Georgian Language. *Reports of Enlarged Session* of the Seminar of VIAM, Volume 26, pp.58-64, 2012
- PKHAKADZE K. The Technological Alphabet of The Georgian Language – The One of The Most Important Georgian Challenge of The XXI Century. Proceedings of The Parliament Conference The Georgian Language – The Challenge of the 21th Century, pp.98-105, 2013.
- PKHAKADZE K., CHANKVETADZE G., TIBUA L., IVANISHVILI M., LEKIASHVILI L., SOSELIA E. About Main Ideas of Direct Formal-Logical Description of the Georgian Language. *Proceedings of VIAM*, Vol. 53, pp.33-40, 2003.
- 13. PKHAKADZE K., CHANKVETADZE G., TIBUA L., IVANISHVILI M., LEKIASHVILI L., SOSELIA E. About the Main Ideas of the Direct Formal-Logical Description of the Georgian Natural Language System

Through the Examples. Proceedings of the V Tbilisi Symposium on Language, Logic and Computation, pp.129-137, 2003.

- 14. PKHAKADZE K. On the Linguistic Relation and Logical Declension in the Georgian Language. *Journal Georgian language and Logic*, Vol.1, pp.19-77, 2005.
- 15. PKHAKADZE K., ABZIANIDZE L., MASKHARASHVILI A. Georgian Language's Theses. *Reports of the Seminar of VIAM*, Vol. 34, pp.108-121, 2008.
- PKHAKADZE K., CHICHUA G., ABZIANIDZE L., MASKHARASHVILI A. 1-Stage Voice Managed Georgian Intellectual Computer System. *Reports of the Seminar of VIAM*, Vol. 34, 96-107, 2008.
- 17. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., KURTSKHALIA D., MALIDZE SH. The Open Letter to the Georgian Parliament, Government, National Academy of Sciences and Georgian and Abkhazian Society i.e. Key Principles of the Unified program of Complete technological supporting of the official language of Georgia (Georgian, Abkhazian) i.e. In the Future Cultural World with the Completely Supported Georgian and Abkhazian Languages. Journal Georgian Language and Logic, Vol.11 (II Part), pp.121-164, 2017-2018.
- PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., BERIASVILI I., KURTSKHALIA D. Take Care on the Georgian Language. Proceedings Of Georgian National Academy of Sciences, pp.149-158, 2016.
- 19. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., KURTSKHALIA D. The Open letter to the Georgian Society, Or, Already, It Is Really the Time To take Care for the Georgian Language. *Journal Georgian Language and Logic*, Vol.9, pp.131-158, 2015-2106.
- PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., BERIASVILI I., KURTSKHALIA D. A Logical Grammar of Georgian Language: Foundations and Applications – First Part – Preface. Journal Georgian Language and Logic, Vol.9, pp.4-130, 2015-2106.
- 21. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G. The Project "Foundations of Logical Grammar of Georgian Language and Its Application in Information Technology" and Doctoral Themes "Georgian grammar checker (analyzer)" and "Georgian speech synthesis and recognition". *Journal Georgian Language and Logic*, pp.21-36, 2013-2014.

- 22. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., MASKHARASHVILI A., BERIASVILI I. An Overview of the Trial Version of the Georgian Self-Developing Intellectual Corpus Necessary for Creating Georgian text Analyzer, Speech Processing, and Automatic Translation Ssystems. *Reports of Enlarged Session of the Seminar of VIAM*, Volume 28, pp.70-75, 2014.
- 23. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., BERIASVILI I., KURTSKHALIA D. The Aims and First Results of the Project "One More Step Towards Georgian Talking Self-Developing Intellectual Corpus". Proceedings of the I international Conference on Language and Modern Technology, pp.108-112, 2015.
- 24. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., BERIASVILI I., KURTSKHALIA D. In the European Union with Georgian language – the Aims and Basements of the Project "One More Step Towards Georgian Talking Self-Developing Intellectual Corpus". *Reports of Seminar* of VIAM, Volume 29, pp.37-43, 2015.
- 25. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., KURTSKHALIA D., MALIDZE SH. Georgian Universal Smart Corpus and in It Inbuilt Abkhazian, Chechen, Kabardian, Lezgian, and Mingrelian Self-Developing Corpuses – Results, Perspectives. *Reports of Enlarged Sessions of the Seminar of VIAM*, Volume 32, pp.55-58, 2018.
- 26. PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., KURTSKHALIA D., MALIDZE SH., N. OQROSHIASHVILI, DEMURCHEV C. The Trial Versions of the New Developing Tools of the Georgian Universal Smart Corpus. *Reports of Enlarged Sessions of the Seminar of VIAM*, Volume 33, pp.50-53, 2019.
- PKHAKADZE K., CHIKVINIDZE M., CHICHUA G., BERIASVILI I., KURTSKHALIA D., MALIDZE SH. The Georgian Intellectual Web Corpus: Aims, Methods, Recommendation. Additional Publication of the Journal Georgian Language and Logic, pp.4-320, 2017.